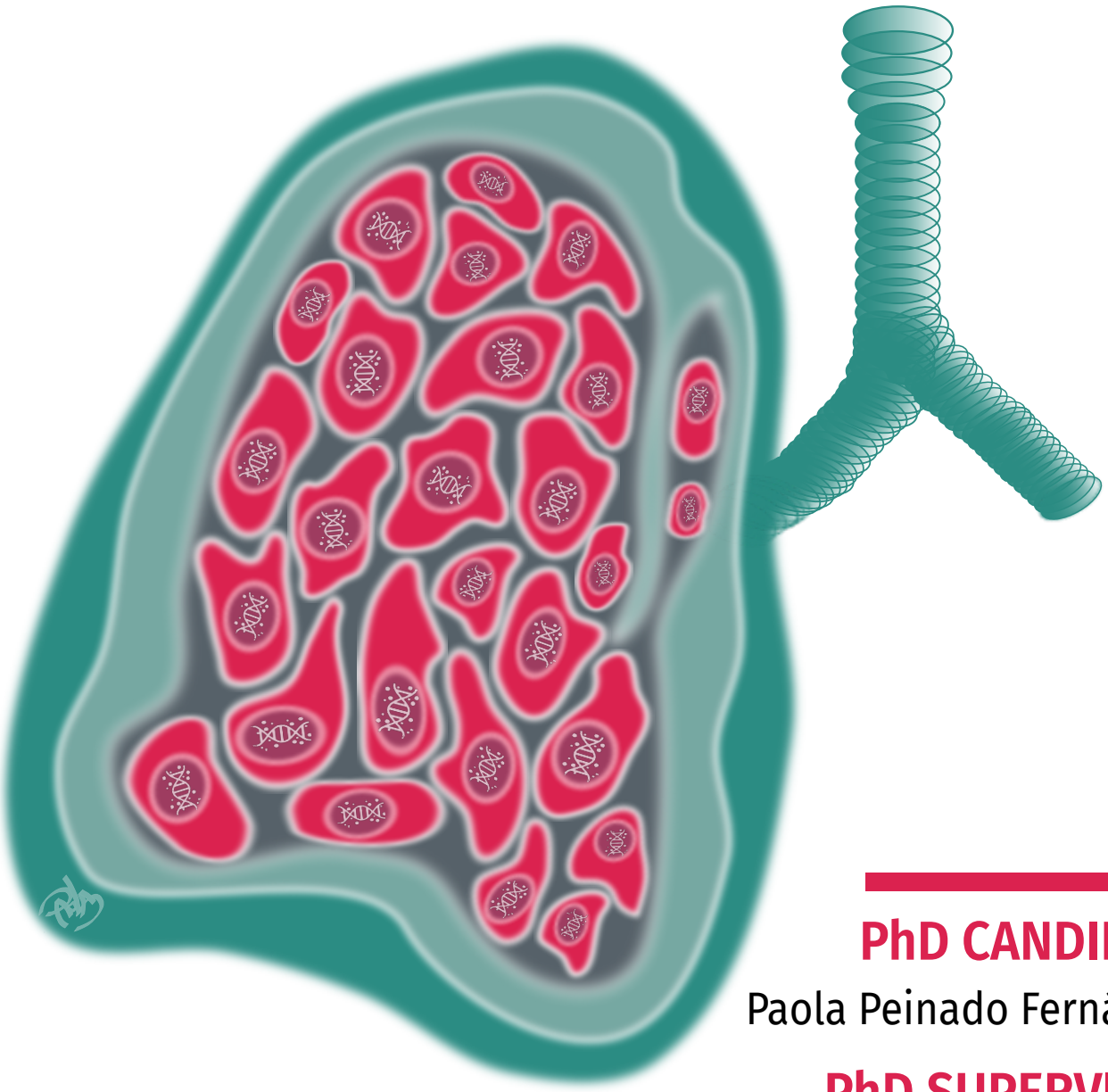
PROGRAMA DE DOCTORADO

Bioquímica y Biología Molecular

IMPLICATIONS of the chromatin remodeling complex SWI/SNF in LUNG CANCER



PhD CANDIDATE

Paola Peinado Fernández

PhD SUPERVISOR

Pedro P. Medina Vico

UNIVERSIDAD DE GRANADA

Genyo

CENTRO PFIZER-UNIVERSIDAD DE GRANADA-JUNTA DE ANDALUCÍA DE **GENÓMICA** E **INVESTIGACIÓN ONCOLÓGICA**

UNIVERSIDAD
DE GRANADA

Departamento de Bioquímica y Biología Molecular I

Facultad de Ciencias

Programa de doctorado en Bioquímica y Biología Molecular (B16.56.1)

Escuela de doctorado en ciencias de la salud

TESIS DOCTORAL

IMPLICATIONS OF THE CHROMATIN REMODELING COMPLEX SWI/SNF IN LUNG CANCER

Paola Peinado Fernández

Tesis doctoral dirigida por:

Pedro P. Medina Vico

Granada, 2021

Genyo
CENTRO PFIZER-UNIVERSIDAD DE GRANADA-JUNTA DE ANDALUCÍA
DE **GENÓMICA** E **INVESTIGACIÓN ONCOLÓGICA**

Editor: Universidad de Granada. Tesis Doctorales
Autor: Paola Peinado Fernández
ISBN: 978-84-1117-605-7
URI: https://hdl.handle.net/10481/78936

A mis padres,
por ser los dos grandes pilares
sobre los que se ha construido
esta tesis doctoral.

"No es el resultado de la investigación científica
que ennoblece a los seres humanos y enriquece su naturaleza,
sino la lucha por entender mientras realiza
un trabajo intelectual creativo y de mente abierta".

–Albert Einstein

SUMMARY

SWI/SNF (SWItch/Sucrose Non-Fermentable) complexes are ATP-dependent chromatin remodelers composed of varying combinations of subunits that, together, fine-tune relevant biological processes, such as transcriptional regulation and genome integrity. Moreover, the subunits of this chromatin remodeler have been identified as major targets of mutations in several tumor types underlying a relevant role in tumorigenesis. However, there is a lack of comprehensive studies of the whole SWI/SNF complex in lung adenocarcinoma (LUAD), and its clinical implications are not yet fully understood. In this Ph.D. thesis, we combine genomic, transcriptomic, and proteomic techniques to identify which SWI/SNF subunits are present in lung epithelial cells, as well as to determine their mutational status and expression levels in LUAD in a more integrative manner. For these purposes, we analyzed data from LUAD primary tumors, normal lung and LUAD cell lines, and external LUAD data from The Cancer Genome Atlas. We found that SWI/SNF complex mutations not only present a high incidence of 41.4% in LUAD patients and 76% in LUAD cell lines, but also the mutational status of the SWI/SNF complex was associated with poorer overall survival of LUAD patients and higher tumor mutation burdens. Furthermore, we observed that the expression of the SWI/SNF complex in LUAD suffers overall repression, which cannot be exclusively explained by genetic alterations. Based on our findings, we propose that SWI/SNF-mutant LUAD tumors could be considered as a clinical subgroup with applications in the prognosis of LUAD patients. In addition, we chose different LUAD cell models to perform various functional experiments to answer two of the open-ended questions related to the role of the SWI/SNF complex in cancer. On the one hand, we studied the function of the SWI/SNF complex in the regulation of the non-protein-coding part of the genome, such as microRNAs, which are small RNAs with key regulatory functions and whose expression is frequently altered in cancer. Specifically, we used a SMARCA4-deficient model of LUAD cells to track changes in the miRNome upon SMARCA4 restoration. We found that the SWI/SNF complex, when SMARCA4 was the catalytic subunit, induced the expression of miR-222 through its direct binding to the enhancer region of this microRNA. Furthermore, we demonstrated that the overexpression of miR-222 phenocopied the tumor suppressor effect of SMARCA4

in LUAD. Thus, these results highlight the importance of considering microRNAs as another layer of regulation that can be controlled by the SWI/SNF complex with relevant implications in tumor development. On the other hand, we analyzed the functional consequences of silencing ARID1A, one of the most mutated SWI/SNF subunits in cancer, in different LUAD cell lines. We discovered that several LUAD cell lines, regardless of their genetic background, developed a dependency on ARID1A that was not observed in normal lung epithelial cells. Interestingly, we found that ARID1A loss in LUAD cells impaired proteostasis, increasing cellular stress and DNA damage that led to cell death. Moreover, we observed that depleting ARID1A behaved as a genotoxic treatment improving the effectiveness of other chemotherapy drugs used for treating lung cancer. In conclusion, our study provides an overview of the variety of implications of the SWI/SNF complex in LUAD, underlining the unexplored potential that this chromatin remodeler has in clinical practice.

RESUMEN

Los complejos SWI/SNF (del inglés *SWItch/Sucrose Non-Fermentable*) son remodeladores de cromatina dependientes de ATP compuestos por combinaciones variables de subunidades que, juntas, ajustan importantes procesos biológicos tales como la regulación de la transcripción y la integridad del genoma. Además, las subunidades de este remodelador de la cromatina se han identificado como principales dianas de mutaciones en muchos tipos tumorales, lo que enfatiza su relevante papel en la tumorogénesis. Sin embargo, faltan estudios exhaustivos sobre el complejo SWI/SNF en el adenocarcinoma de pulmón (LUAD) y sus implicaciones clínicas aún no se conocen del todo. En esta tesis doctoral combinamos técnicas genómicas, transcriptómicas y proteómicas para identificar qué subunidades del SWI/SNF están presentes en las células epiteliales de pulmón, así como para determinar su estado mutacional y sus niveles de expresión en LUAD de una forma más integrativa. Para ello, analizamos los datos de tumores primarios de LUAD, de líneas celulares de pulmón normal y de LUAD, así como los datos externos de LUAD del TCGA (del inglés *The Cancer Genome Atlas*). Encontramos que las mutaciones del complejo SWI/SNF no sólo presentan una alta incidencia del 41,4% en los pacientes con LUAD y del 76% en las líneas celulares de LUAD, sino que también el estado mutacional del complejo SWI/SNF se asociaba con una peor supervivencia de los pacientes con LUAD y con una mayor carga de mutaciones tumorales. Además, observamos que la expresión del complejo SWI/SNF en LUAD sufre una represión global que no puede explicarse exclusivamente por alteraciones genéticas. Basándonos en nuestros resultados, proponemos que los tumores de LUAD con un SWI/SNF mutante podrían considerarse como un subgrupo clínico con aplicaciones en la prognosis de los pacientes con LUAD. Además, elegimos diferentes modelos celulares de LUAD para realizar varios experimentos funcionales con el fin de responder a dos de las preguntas abiertas que existen en relación al papel del complejo SWI/SNF en cáncer. Por un lado, estudiamos la función del complejo SWI/SNF en la regulación de la parte del genoma que no codifica para proteína, como los microARNs, que son pequeños ARNs con funciones reguladoras clave y cuya expresión se encuentra frecuentemente alterada en cáncer. Concretamente usamos un modelo celular de LUAD deficiente en SMARCA4 para realizar un

seguimiento de los cambios en el miRNoma derivados de la restauración de SMARCA4. Descubrimos que el complejo SWI/SNF, cuando SMARCA4 era su subunidad catalítica, inducía la expresión del miR-222 a través de su unión directa a la región potenciadora de este microARN. Además, demostramos que la sobreexpresión de miR-222 copiaba el fenotipo supresor tumoral de SMARCA4 in LUAD. Por lo tanto, estos resultados ponen de manifiesto la importancia de considerar los microARNs como otra capa de regulación controlada por el complejo SWI/SNF con importantes implicaciones en el desarrollo tumoral. Por otro lado, analizamos las consecuencias funcionales de silenciar *ARID1A*, una de las subunidades del SWI/SNF más mutadas en cáncer, en diferentes líneas celulares de LUAD. Descubrimos que varias líneas celulares de LUAD, independientemente de su fondo genético, desarrollaban una dependencia hacia ARID1A que no se observaba en células epiteliales de pulmón normal. Curiosamente, la pérdida de ARID1A en células de LUAD alteraba la proteostasis, aumentando el estrés celular y el daño en el ADN que conducía a la muerte celular. Además, observamos que eliminar ARID1A presentaba un comportamiento similar a los tratamientos genotóxicos, mejorando incluso la eficacia de otros fármacos genotóxicos usados para tratar el cáncer de pulmón. En conclusión, nuestro estudio proporciona una visión general de la gran variedad de implicaciones que tiene el complejo SWI/SNF en LUAD, resaltando el potencial inexplorado que tiene este remodelador de la cromatina en la práctica clínica.

TABLE OF CONTENT

FIGURE INDEX

MAIN FIGURES

SUPPLEMENTARY FIGURES

TABLE INDEX

MAIN TABLES

SUPPLEMENTARY TABLES

ABBREVIATIONS/ACRONYMS

Abbreviation/Acronym	Extended name
ACTB	β-Actin
AGO	Argonaute
ARID1A/B	AT-rich interactive domain 1A/B
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
BAF	BRM/BRG1 Associated Factors
bp	Base pair
CAA	Chloroacetamide
CCLE	Cancer Cell Line Encyclopedia
cDNA	Complementary DNA
ChIA-PET	Chromatin interaction analysis by paired-end tag sequencing
ChIP	Chromatin Immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation sequencing
CRC	Chromatin Remodeling Complex
CSS	Coffin-Siris syndrome
CTCF	CCCTC-binding factor
DDR	DNA Damage Response
DF	Degrees of freedom
DGCR8	DiGeorge Syndrome Critical Region 8
DMEM	Dulbecco's modified Eagle's medium
DNMT	DNA methyltransferase
Doxo	Doxorubicin
DSB	Double-Strand Break
EGA	European Genome-phenome Archive
EGF	Epidermal Growth Factor
ENA	European Nucleotide Archive
ENCODE	Encyclopedia of DNA Elements
ER	Endoplasmic reticulum
Etop	Etoposide
EV	Empty vector
FA	Fanconi's Anemia
FASP	Filter Aided Sample Preparation
FBS	Fetal Bovine Serum
FC	Fold change
FDR	False Discovery Rate
GEO	Gene Expression Omnibus
GEPIA2	Gene Expression Profiling Interactive Analysis
GO	Gene Ontology
GSEA	Gene set enrichment analyses

GSH	Glutathione
GTEx	Genotype-Tissue Expression
H3K27ac	Acetylation of the lysine residue at N-terminal position 27 of the histone H3 protein
H3K4me1	Mono-methylation at the 4th lysine residue of the histone H3 protein
H3K4me3	Tri-methylation at the 4th lysine residue of the histone H3 protein
HCC	Hepatocellular Carcinoma
HCD	Higher-energy collisional dissociation
Hi-C	High-throughput chromosome conformation capture
HR	Homologous Repair
HUGO	Human Genome Organization
IGV	Integrative Genomics Viewer
IP	Immunoprecipitation
IQR	Interquartile Range
Kb	Kilobase
KD	Knockdown
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knockout
LC-MS/MS	Liquid chromatography with tandem mass spectrometry
LFQ	Label-Free Quantification
LOF	Loss of function
LUAD	Lung adenocarcinoma
Mb	Megabase
miRISC	miRNA-induced silencing complex
miRNA	MicroARN
mRNA	Messenger RNA
MRT	Malignant Rhabdoid Tumors
MSI	Microsatellite instability
MSigDB	Molecular Signatures Database
nBAF	Neural BRM/BRG1 Associated Factors
ncBAF	Non-canonical BRM/BRG1 Associated Factors
NER	Nucleotide Excision Repair
NES	Normalized Enrichment Score
NGS	Next-generation sequencing
NHEJ	Non-Homologous End-Joining
NS	Non-significant
NSCLC	Non-Small Cell Lung Cancer
nt	Nucleotide
NT	Not treated
OCCC	Ovarian Clear Cell Carcinoma
OS	Overall Survival
P.adj	Adjusted *p* values
PARP1	Poly-(ADP-ribose) polymerase 1
PBAF	Polybromo-Associated BAF complexes

PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PE	Phycoerythrin
PoN	Panel of normals
PRC	Polycomb Repressor Complex
pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA
PTM	Post-translational modification
PVDF	Polyvinylidene difluoride
ROS	Reactive Oxygen Species
RPA	Replication protein A
RPMI	Roswell Park Memorial Institute
RT	Room Temperature/Reverse Transcription
RT-qPCR	Real-time quantitative PCR
SCCOHT	Small Cell Carcinoma of the Ovary Hypercalcemic Type
SCR	Nonsense scrambled siRNA
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	Standard error
siRNA	Small interference RNA
SMARC	SWI/SNF-related, Matrix-associated, Actin-dependent Regulator of Chromatin
SWI/SNF	SWItch/Sucrose Non-Fermentable
TAD	Topologically associating domains
TCEP	Tris-(2-carboxyethyl)-phosphine)
TCGA	The Cancer Genome Atlas
TET	Ten Eleven-Translocation
Th	Thompson unit
TMB	Tumor mutation burden
TMM	Trimmed Mean of M values
TMT	Tandem Mass Tag
TOP2	Topoisomerase II
TPM	Transcripts Per Million
TRBP	Transactivation-responsive RNA-binding protein
TRC	Transcription-Replication Conflict
UCSC	University California, Santa Cruz
UPR	Unfolded Protein Response
UTR	Untranslated region
VCF	Variant Call Format
WT	Wild type
XPO5	Exportin 5

Introduction

1) INTRODUCTION

1.1. Importance of chromatin dynamics in gene expression

Under normal conditions, every cell of the hundreds of cell types that comprise the human body contains an identical genetic material. The same DNA provides all cells with an adaptive tool to respond to different environmental and developmental conditions thanks to its variety of gene expression patterns. Thus, the regulation of those expression patterns is vital for the cell. Consequently, eukaryotic cells have a highly dynamic storage system for DNA called **chromatin**.

Chromatin is composed of DNA wrapped around an octamer of two copies of each of the histones H2A, H2B, H3, and H4, forming a fundamental repeating unit called **nucleosome** (Luger et al. 1997). That particular organization generates a primary chromatin structure that other architectural proteins can additionally condense (Tremethick 2007). Initially, this characteristic organization was thought to be a system to compact genomic DNA (~2 m for the human genome) to fit into the cellular nucleus (with a diameter of only ~10 μm). However, this packaging still leaves ample unoccupied space within the nucleus. Therefore, instead of solving a physical packaging problem, this organization is crucial for controlling accessibility to genetic information, and it affects many cellular processes that require an unwinding of the DNA, such as replication, DNA repair, and transcription (Luger et al. 2012; Soria et al. 2012; Petesch and Lis 2012).

In the 80s and 90s, many researchers noted that studying chromatin structure was essential for understanding the regulation of our genetic material. These researchers performed pioneering studies that demonstrated how the position of nucleosomes could alter transcription (reviewed in (Grunstein 1990; Felsenfeld 1992)). Later, advances in genomics and high-throughput techniques corroborated that nucleosome-free regions are characteristic of transcriptionally active genes, specifically, of their proximal (promoters) and distal (enhancers) regulatory elements (Jiang and Pugh 2009; Lai and Pugh 2017). This **nucleosome depletion**

ensures DNA accessibility to transcription, replication, and DNA repair machinery.

1.1.1. Mechanisms of regulating chromatin structure

In general, there are **three classes of enzymes that regulate chromatin structure and nucleosome dynamics**: DNA methylation-demethylation enzymes, histone-modifying enzymes, and ATP-dependent chromatin remodelers (***Figure 1***) (Valencia and Kadoch 2019). All of them comprise a group called: **epigenetic factors**. The word "*epigenetic*" means "*in addition to changes in genetic sequence*". This term includes those processes that alter gene activity without changing the DNA sequence and lead to modifications that can be transmitted to daughter cells.

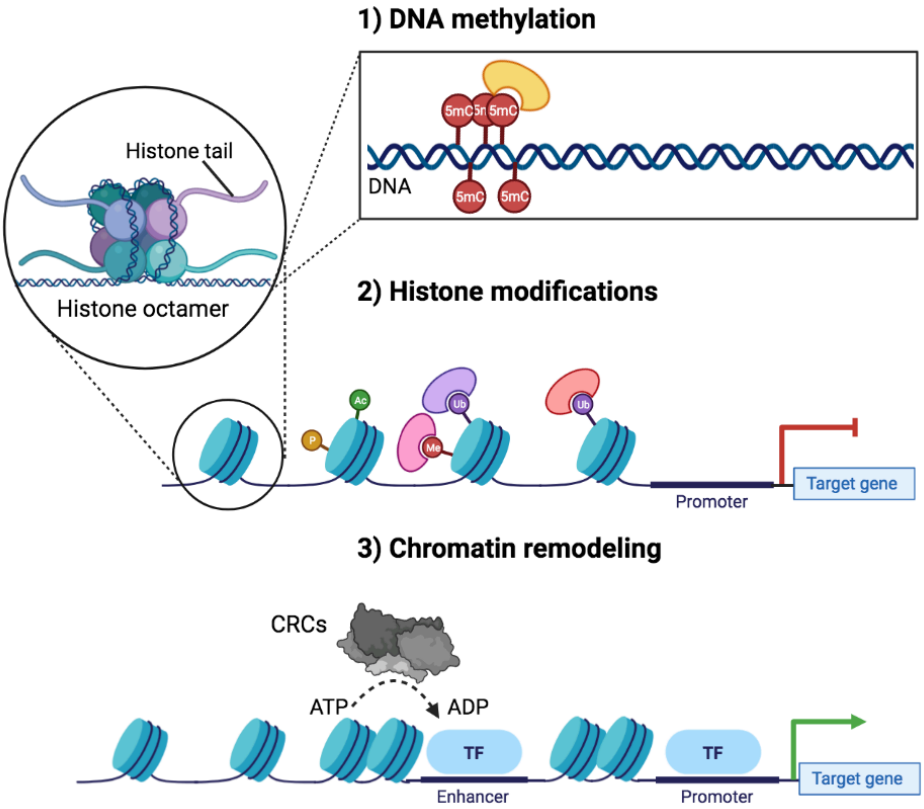


Figure 1: Three levels of regulation of chromatin structure: (**1**) DNA methyltransferases and DNA demethylases add or erase methyl groups in cytosines (5mC). Those marks are associated with transcriptional repression. (**2**) Histone modifications, such as acetylation (Ac), methylation (Me), ubiquitination (Ub), and phosphorylation (P), serve as instructive marks for both gene activation and gene repression. (**3**) Four families of ATP-dependent chromatin remodeling complexes (CRCs) alter chromatin architecture by mobilizing, depositing, or evicting nucleosomes. (Image made with BioRender.com).

1) The first class of enzymes involved in chromatin regulation is composed of DNA methyltransferases (DNMTs) and Ten Eleven-Translocation (TET) family of proteins among other proteins (Schübeler 2015). These enzymes establish or erase a methyl group in cytosines in the context of a CpG dinucleotide. These reactions create DNA methylation patterns that are associated with transcriptional repression. Several studies have shown that highly methylated DNA regions, specifically methylated CpG islands, correlate with high nucleosome occupancy (Collings et al. 2013; Chodavarapu et al. 2010; Collings and Anderson 2017).
2) The second group comprises proteins involved in adding or removing covalent post-translational modifications (PTMs) to histone tails, such as serine phosphorylation, lysine acetylation and ubiquitylation, and lysine and arginine methylation (Tessarz and Kouzarides 2014; Marmorstein 2001). There are 106 different sites of PTMs in the 55 human histone proteins and over 150 histone-modifying proteins identified (Khare et al. 2012). Those histone marks have different outcomes in chromatin dynamics and gene expression.
3) The third class corresponds to large multi-protein complexes called **chromatin remodeling complexes (CRCs)** or chromatin remodelers. These CRCs are typically composed of multiple subunits, and they can promote or distort nucleosome structure and mobilize nucleosomes by the use of energy derived from ATP hydrolysis (Clapier et al. 2017; Becker and Workman 2013). Overall, the role of these chromatin-remodeling complexes is to facilitate accessibility of the transcription machinery to DNA, although other chromatin remodelers are involved in transcriptional repression (Whitehouse and Tsukiyama 2006; Wade et al. 1999).

To date, there are **four families of ATP-dependent chromatin remodelers: SWI/SNF, ISWI, NURD/CHD, and INO80/SWR1** (reviewed in (Hargreaves and Crabtree 2011)). These chromatin remodelers contain a single motor subunit that belongs to the superfamily of ATP-dependent translocases and helicases. These catalytic subunits contain class-specific domains that modulate their activity or mediate binding to DNA or histones (Bracken et al. 2019). The rest of the non-

catalytic subunits of these remodeler complexes provide a variety of additional functionalities, including regulation of the ATPase and facilitating contacts with DNA or other proteins, such as histones, histone chaperones, or transcription factors.

Mostly, a chromatin remodeling reaction has different outcomes (Clapier et al. 2017; Becker and Workman 2013): it can move nucleosomes along the DNA through a sliding mechanism; it can change DNA accessibility without any translocation of the histone octamer; it can disrupt the nucleosome structure; it can mediate the exchange between histone variants; or it can eject the histone octamer (***Figure 2***).

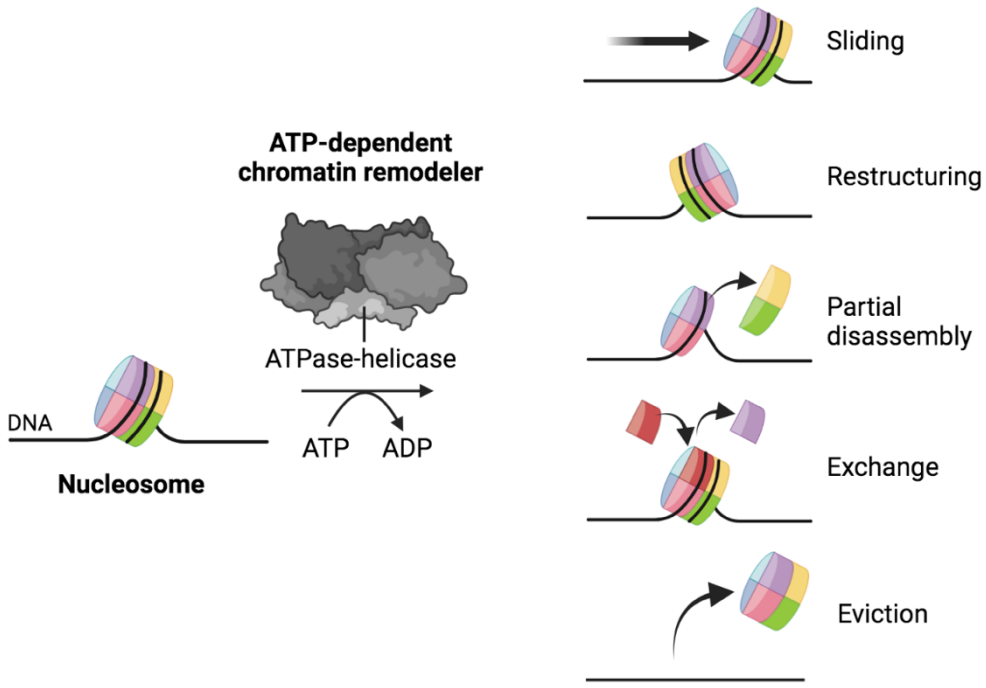


Figure 2: **Overview of the different effects of ATP-dependent chromatin remodelers on nucleosomes**. These chromatin remodelers can move nucleosomes to other positions on the DNA and lead to the exposure of a previously hidden DNA sequence (1. Sliding). They can rearrange DNA around the histone octamer and change DNA accessibility (2. Restructuring). Chromatin remodelers can also evict part of the histone octamer, specifically histone H2A/H2B dimers (3. Partial disassembly). Other remodelers can exchange histone variants within the nucleosomes (4. Exchange). They can also eject the nucleosome from a DNA sequence (5. Eviction). Image made in BioRender.com and adapted from (Bracken et al. 2019).

1.2. Implications of epigenetic plasticity: cellular identity and pathology

Overall, these mechanisms of chromatin regulation establish and maintain gene expression states that are responsible for guiding cell identity transitions. This process is called 'epigenetic plasticity' (Flavahan et al. 2017). Aberrations in this dynamic regulation have been associated with a wide range of common diseases, including aging-related diseases, neuropsychiatric disorders, autoimmunity, and cancer (Flavahan et al. 2017).

1.2.1. Epigenetic alteration, the new hallmark of cancer

Cancer is a heterogeneous group of diseases characterized by abnormal and uncontrolled cell growth. It originates from almost any organ or tissue, and in some cases, tumor cells can invade surrounding parts of the body or even reach other organs in a process known as metastasis. This latter process is responsible for the high lethality of this disease (Massagué and Obenauf 2016).

Nowadays, cancer is one of the leading causes of mortality worldwide, reaching almost 10 million deaths only in 2020 (GLOBOCAN, 2020). Among all tumor types, **lung cancer is currently the deadliest**, causing more than 1.7 million deaths in both men and women around the world (GLOBOCAN, 2020). In particular, lung adenocarcinoma (LUAD) is the histological subtype more represented among lung cancer patients (Bray et al. 2018). However, despite the urgent health problem that it is causing, researchers and clinicians are still facing difficulties to improve its early diagnosis and treatment. For this reason, understanding the biology of this heterogeneous disease is crucial.

Traditionally, cancer was seen as a genetic disease, and therefore, the main goal of cancer research was to study the molecular players involved in ten essential pathways that dictate tumor growth. Those essential pathways, considered as the "**Hallmarks of cancer**", are: self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of cell death, limitless replicative potential, sustained angiogenesis, activation of tissue invasion and metastasis, evasion of

immune control, metabolic deregulation, pro-inflammatory state, and genetic instability and mutation (Hanahan and Weinberg 2011).

However, during the last years, several studies have demonstrated the relevance of an **additional cancer hallmark: Epigenetic deregulation**. Cancer cells present gene expression abnormalities and aberrations of their cellular identity and their responsiveness to internal and external stimuli (Baylin and Jones 2011; Esteller 2008; Feinberg et al. 2016). Large-scale genome sequencing projects have shown that approximately 50% of human cancers harbor mutations in chromatin proteins (You and Jones 2012; Shen and Laird 2013). In fact, after performing whole-genome sequencing of LUAD patients, Imielinski and colleagues proposed the 11th Hallmark of cancer ('Epigenetic or RNA deregulation') to highlight the relevance of epigenetic and splicing regulators in oncogenesis (***Figure 3***) (Imielinski et al. 2012).

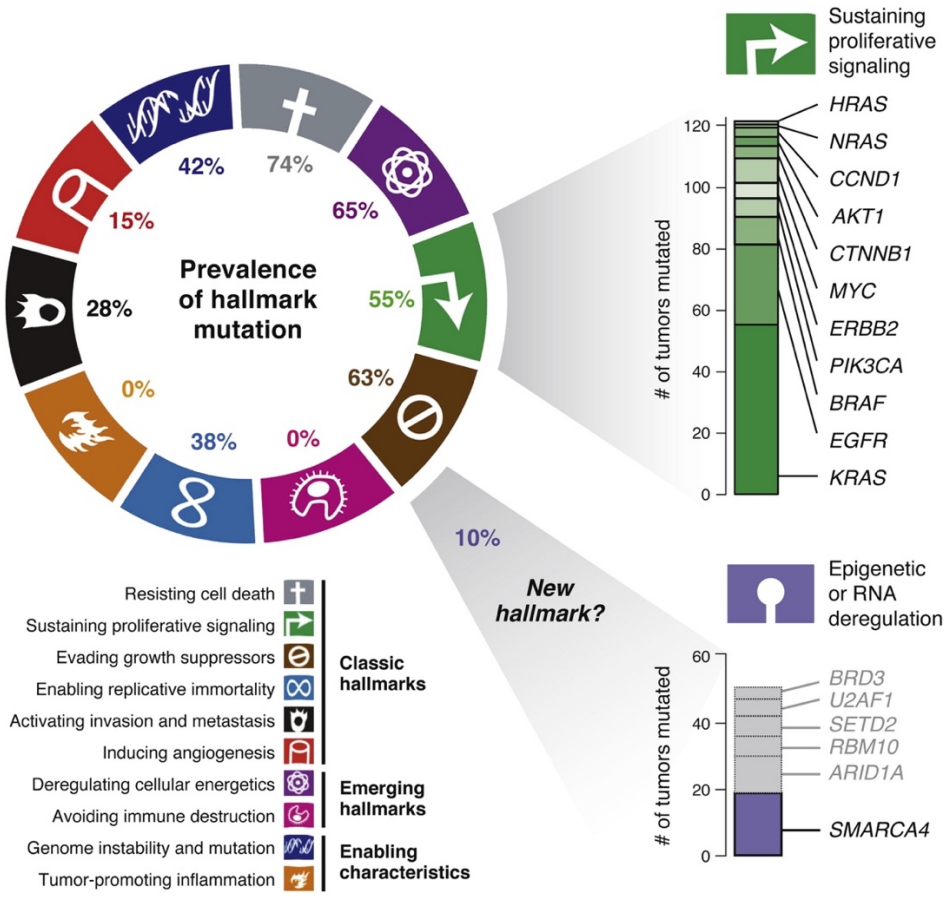


Figure 3: Next-generation cancer hallmarks with data from LUAD patients (Imielinski et al. 2012). Left, Hanahan and Weinberg model of cancer hallmarks (Hanahan and Weinberg 2011) showing the prevalence of mutation of the genes involved in each pathway. Top right, top mutated

genes of LUAD involved in the 'Sustaining proliferative signaling' hallmark. Bottom right, top mutated genes of LUAD with a role in 'Epigenetic or RNA deregulation'.

Nowadays, it is widely accepted that **genetic and epigenetic mechanisms are not separate events in cancer**, and they depend on each other during tumor formation and development. On the one hand, a malfunction of epigenetic regulators, apart from deregulating gene expression, can also cause genetic instability that generates mutations. On the other hand, genetic alterations of epigenetic regulators can result in aberrant activation or silencing of several cellular signaling pathways that lead to cancer (You and Jones 2012) (***Figure 4***).

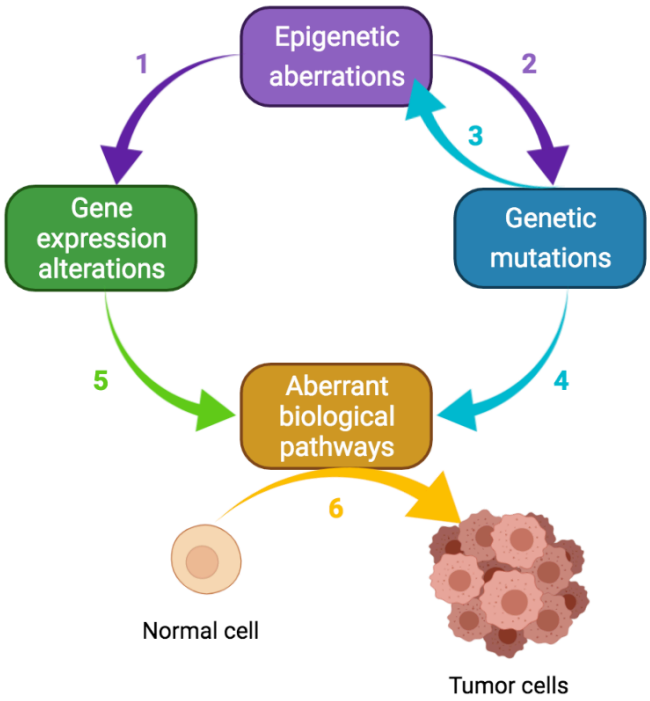


Figure 4: Genetic-Epigenetic interplay in cancer. Alterations in epigenetic regulators, such as the three types described in the section 1.1.1, cause changes in gene expression (**1**) and can also lead to genetic instability and genetic mutations (**2**). Genetic alterations lead to the production of aberrant molecules that can be epigenetic regulators (**3**) or other molecules with important roles in specific pathways (**4**). Different expression levels of relevant molecules for the cell can also impair the normal functioning of the cells (**5**). Overall, the changes that confer a selective advantage for cell proliferation and survival will drive tumor formation. Image made with BioRender.com.

Regardless of their classification as genetic or epigenetic alterations, the scientific community considers as '**drivers**' those events that are essential for tumor formation because they confer on cancer cells selective advantages over neighboring cells (Stratton et al. 2009). The genes that are the target of those driver events are considered **cancer genes**.

Traditionally, cancer genes are divided into two categories: oncogenes and tumor suppressor genes.

A) *Oncogenes*: In their wild-type state, these genes are responsible for controlling cell growth pathways. They may function as growth factors, transducers of cellular signals, and nuclear transcription factors, and they receive the name of 'proto-oncogenes'. In general terms, gain-of-function mutations convert those proto-oncogenes into oncogenes that stimulate cell growth, division, and survival. Amplifications, chromosomal translocations, and epigenetic alterations can also be causes of this aberrant behavior (Kontomanolis et al. 2020).

B) *Tumor suppressor genes*: They correspond to genes that normally prevent unrestrained cellular growth and promote DNA repair and cell cycle checkpoint activation (Kontomanolis et al. 2020). In cancer cells, these genes are inactivated through different mechanisms: loss-of-function mutations, deletions, and epigenetic dysregulation, such as promoter hypermethylation. To fully inactivate these tumor suppressor genes, both alleles have to be altered by a process called the 'two-hit hypothesis' (Knudson 1971). Usually, cancer cells present loss of heterozygosity (LOH) events in tumor suppressor genes through deletions of chromosomes during abnormal mitotic divisions (Liu et al. 2011).

1.2.2. Epigenetic alterations that promote tumorigenesis

Although genetic alterations play a crucial role in oncogenesis, different epigenetic mechanisms can also lead to tumor formation.

A) *Promoter hypermethylation*: Several studies have observed that promoter hypermethylation of tumor suppressor genes is a common feature in cancer (Baylin and Jones 2011). There are genes that are rarely mutated but they lose their functions after they are silenced by promoter hypermethylation. For example, the DNA repair gene O6-methylguanine-DNA methyltransferase (MGMT), the cell cycle regulator p15 or Cyclin-dependent kinase inhibitor 2B (CDKN2B), and the regulator of the RAS oncogene RASSF1A have relevant protective roles for the cells and they suffer this type of inactivation in different

types of cancer (Baylin and Jones 2011). A high-throughput study performed in ovarian carcinoma patients showed that, although the mutations mostly lied on a single gene (TP53), the methylation analysis found that 168 genes presented promoter hypermethylation and that correlated to lower expression levels (Network 2011).

B) *microRNA regulation*: MicroRNAs (miRNAs) are **small non-protein-coding RNAs** of 18–25 nucleotides that bind to target messenger RNAs (mRNAs) and promote either their degradation or their translational repression (Bartel 2004). They constitute an additional epigenetic layer of expression control. MiRNAs are expressed in all major animal model systems, and their variety correlates with organismal complexity. Approximately there are 3000 human miRNAs annotated (Friedländer et al. 2014; Kozomara and Griffiths-Jones 2011). Importantly, several studies have described a dysregulation of miRNAs in many diseases, including cancer (Lin and Gregory 2015; Medina and Slack 2008; Croce 2009). In fact, many of the targets of miRNAs are frequently related to tumor development (Ryan et al. 2010). For these reasons, miRNAs have been classified as **oncogenic (oncomiRs), tumor suppressors, or context-dependent miRNAs** (Kasinski and Slack 2011). For example, oncogenic miRNAs, such as miR-155 or miR-21, are frequently overexpressed, and tumor-suppressive miRNAs, such as miR-146a or the cluster miR-15/16, are deleted in cancer (Kasinski and Slack 2011). Moreover, miRNAs can target epigenetic modifiers such as EZH2 (Friedman et al. 2009) and SMARCA4 (Coira et al. 2015), and therefore, this results in further widespread epigenetic alterations. Thus, the **tight control of miRNA biogenesis** is a crucial process to guarantee their physiological levels. This pathway involves several enzymes and protein complexes that regulate the production of mature and functional miRNAs (O'Brien et al. 2018) (***Figure 5***). The canonical miRNA biogenesis starts in the nucleus with RNA polymerase II or III transcribing miRNAs genes to generate primary miRNAs (pri-miRNAs). Those pri-miRNAs are capped, spliced, and polyadenylated. Pri-miRNAs present a secondary structure of a hairpin that is recognized and cleaved by the Microprocessor Complex, consisting of an RNA binding protein DiGeorge Syndrome Critical Region 8 (DGCR8) and a

ribonuclease III enzyme (DROSHA). This results in the formation of hairpin-shaped precursors miRNAs (pre-miRNAs) that are exported to the cytoplasm by the exportin 5 (XPO5)/RanGTP complex. Later, pre-miRNAs are processed by DICER1, another ribonuclease III, with the help of its cofactor TRBP (transactivation-responsive RNA-binding protein). This complex removes the terminal loop of the pre-miRNAs and generates a mature ~22-nucleotide miRNA duplex. Finally, one strand of the mature miRNA duplex, either the 5p or 3p, is loaded into the Argonaute (AGO) family of proteins to form a miRNA-induced silencing complex (miRISC). In most cases, complementary base-pairing of the miRNA guides miRISC to a specific sequence at the 3' UTR (untranslated region) of a target mRNA and induces either its translation repression or its decapping and deadenylation (Huntzinger and Izaurralde 2011). However, other studies have revealed that there are miRNA binding sites at 5' UTRs, coding sequences, and even at promoter regions (Xu et al. 2014). Moreover, although the canonical effect of the miRNA-mRNA interaction is related to gene silencing, the binding of miRNA to a promoter region induces transcription (Dharap et al. 2013).

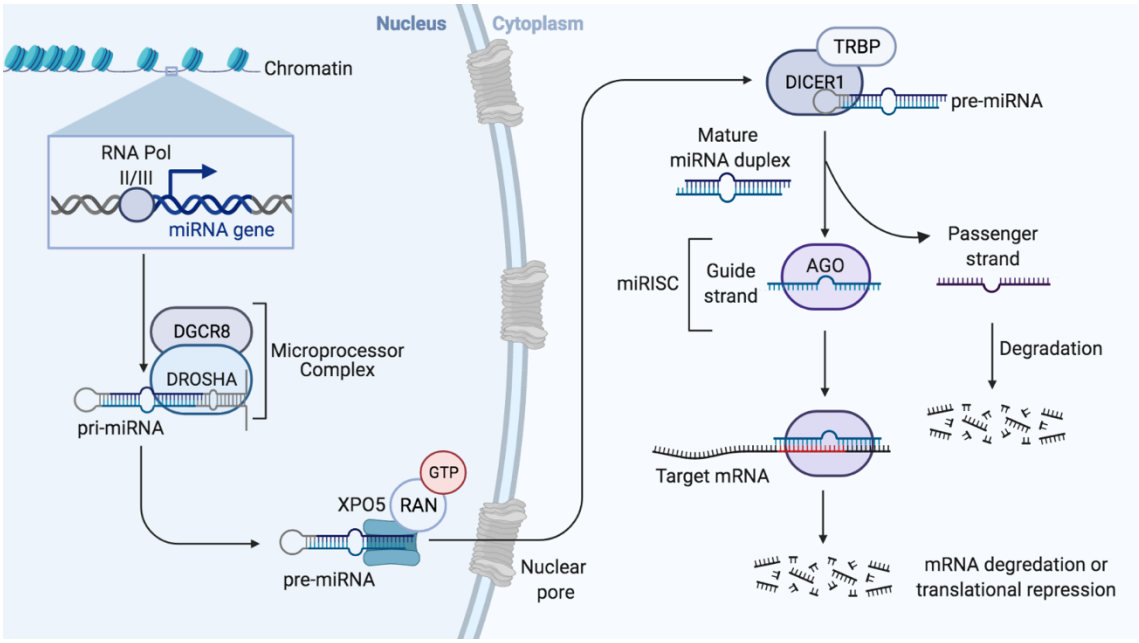


Figure 5: Schematic model of canonical microRNA biogenesis. The miRNA genes are transcribed by RNA polymerase II or III generating the primary miRNAs transcripts (pri-miRNA). They are cleaved by a Microprocessor complex that is composed of an RNAse III, DROSHA, and DGCR8. This reaction produces the precursor miRNAs (pre-miRNAs) that are transported to the cytoplasm by a complex XPO5-RAN-GTP. In the cytoplasm, another RNAse III, DICER1, and its cofactor TRBP process

pre-miRNAs to generate mature miRNAs. Later, the functional strand of the miRNA duplex is loaded into AGO proteins to produce the miRNA-induced silencing complex (miRISC). The unused strand called 'the passenger strand' is degraded. The miRISC complex is guided to the target mRNA by complementary-base pair and it induces mRNA degradation and translational repression. Image made in BioRender.com and adapted from (Winter et al. 2009).

Overall, the effect of miRNAs on gene expression is highly diverse and relevant for an adequate biological balance. In fact, several studies have shown that many proteins of the miRNA biogenesis machinery, such as DROSHA and DICER1, are downregulated (Dedes et al. 2011; Sand et al. 2010; Merritt et al. 2008; Karube et al. 2005) or present inactivating mutations in some tumor types (Heravi-Moussavi et al. 2011; Anglesio et al. 2013; Rakheja et al. 2014). In addition, apart from the proteins involved in miRNA processing, other epigenetic mechanisms, such as promoter hypermethylation and histone modifications, play a crucial role in the expression of miRNAs involved in cancer (Liu et al. 2013; Sato et al. 2011; Moutinho and Esteller 2017). For example, miR-127, which targets the proto-oncogene BCL6, is abnormally methylated and silenced in cancer (Saito et al. 2006).

1.2.3. Relevance of 'epimutations' in cancer

During the last decades, thanks to the performance of several high-throughput sequencing studies, researchers have shown the impact of genetic alterations in many epigenetic regulators. In fact, more than 50% of human cancers harbor mutations in genes involved in chromatin organization and regulation (Valencia and Kadoch 2019; Beck et al. 2012; Weinstein et al. 2013). The discovery of the high mutational rate of some of the proteins implicated in epigenetic control has highlighted the relevance of the interplay between genetics and epigenetics in cancer initiation and progression (You and Jones 2012). Several studies have identified genetic alterations that affect the three types of epigenetic modulators described in section 1.1.1. These 'epimutations' could be considered as the second hit of Knudson's two-hit model postulated for tumor initiation (Jones and Laird 1999). These epimutations could contribute to silencing the remaining active allele

of previously mutated tumor suppressors. Next, we mention some of the most relevant genetic alterations found in the different classes of epigenetic modulators:

A) *DNA methylation regulators*: Different studies have observed that cancer cells present genome-wide hypomethylation that contributes to genomic instability and increases aneuploidy. In addition, site-specific CpG island promoter hypermethylation that leads to the silencing of tumor suppressor genes is a frequent event in tumorigenesis (Esteller 2008). **The source of these DNA methylation alterations is related to an absence or a malfunctioning of the enzymes involved in this process**. Some authors have described in different types of tumors the presence of somatic mutations in many of the DNMTs that maintain and create DNA methylation patterns (Kanai et al. 2003; Yan et al. 2011). For example, Ley and colleagues found that 22% of patients with acute myeloid leukemia (AML) had mutations in DNMT3A, a *de novo* methyltransferase (Ley et al. 2010). In addition, they observed that DNMT3A mutated patients showed worse survival.

B) *Histone modifiers:* A comprehensive analysis of the post-translational modifications of the histone H4 revealed that there was a characteristic global reduction of the trimethylation of H4K20 (H4K20me3) and acetylation of H4K16 (H4K16Ac) in different cancer types (Fraga et al. 2005). This aberrant signature of histone marks highlights the importance of the alterations found in **histone modifiers**, which are **frequently mutated in many types of tumors** (reviewed in (Baylin and Jones 2016)). For instance, EZH2, which catalyzes the trimethylation of histone H3 lysine 27 (H3K27me3), presents both gain-of-function and loss-of-function mutations in distinct tumor types (Kim and Roberts 2016). Specifically, EZH2 mutant lymphomas show an increase in H3K27me3 and a depletion of active

chromatin marks. This results in the arrest of B-cell development in a proliferative state (Béguelin et al. 2013).

C) *Chromatin remodelers*: Among the four families of chromatin remodeling complexes that exist in the cells, the **SWI/SNF complex is the most mutated chromatin remodeler** (Kadoch et al. 2013). Next-generation sequencing studies have revealed that **almost 25% of all human cancers** present a genetic alteration in any of the subunits that form this complex (Shain and Pollack 2013; Kadoch et al. 2013; Mittal and Roberts 2020). Moreover, many of its subunits are considered as key **drivers** of oncogenesis (Bailey et al. 2018).

In this Ph.D. thesis, the focus of attention is on the study of the implications of the SWI/SNF complex in cancer, specifically in lung adenocarcinoma.

1.3. The SWI/SNF complex as a large multi-subunit chromatin remodeler

The SWI/SNF complex was the first remodeler described, and it is also one of the best-studied since it is present in all eukaryotic species. The SWI/SNF complex was initially identified in yeast, where it acted as a transcriptional regulator. Specifically, some studies found a set of genes involved in the activation of crucial pathways in *Saccharomyces cerevisiae,* such as mating-type switching and growth on sucrose media (Neigeborn and Carlson 1984; Stern et al. 1984; Breeden and Nasmyth 1987; Abrams et al. 1986). For this reason, those genes received the name of SWI-SNF (SWItch/Sucrose Non-Fermentable). However, those initial studies only analyzed the effect of individual proteins, and it wasn't until some years later when other researchers proposed that the yeast SWI-SNF proteins could behave like a **large multi-subunit complex** (Peterson and Herskowitz 1992). In addition, Hirschhorn and colleagues noted that, since some of those SWI/SNF proteins had homologs in other species, their function in transcription activation could be **functionally conserved throughout eukaryotes** (Hirschhorn et al. 1992). Finally, in 1994, two independent studies corroborated the existence of a human SWI/SNF complex (Kwon et al. 1994; Imbalzano et al. 1994). However, a few years later, Gerald R.

Crabtree and colleagues showed that there was not a unique human SWI/SNF complex. Instead, they identified different human SWI/SNF complexes depending on the cell line used. With this observation, they introduced a relevant trait of the SWI/SNF complex: its **heterogeneity** and how it adapts to the needs of each cell type (Wang et al. 1996). It was the first time that it was proposed that different SWI/SNF complexes could be involved in the chromatin remodeling of tissue-specific genes.

1.3.1. The combinatorial subunit composition of the SWI/SNF complex

Nowadays, advances in proteomics, biochemistry, molecular biology, and structural biology have identified **29 human genes that encode SWI/SNF subunits.** However, some of these genes are paralogs and only one of the homologous subunits can be found in a complex. In fact, **only 10 to 15 of these subunits constitute a single SWI/SNF complex** (Pulice and Kadoch 2017; Centore et al. 2020). This is the reason why there is a high combinatorial potential that originates **SWI/SNF complexes that are tissue-specific and context-dependent** (Lessard et al. 2007; Ryme et al. 2009). For instance, this combinatorial effect can be observed in the differentiation process of embryonic stem cells where the expression levels of the different SWI/SNF subunits vary depending on the developmental stage (Kaeser et al. 2008; Yan et al. 2008). This combinatory potential of the SWI/SNF complex could explain its functional diversity that we will describe in more detail in the following section.

In general, depending on the subunit composition of the SWI/SNF complex, there are **three main SWI/SNF complexes** currently identified: canonical BAF (BRM/BRG1 Associated Factors), PBAF (polybromo-associated BAF complexes), and the recently discovered non-canonical BAF (ncBAF) (Kaeser et al. 2008; Alpsoy and Dykhuizen 2018; Mashtalir et al. 2018) (***Figure 6***).

All of them are characterized for having an ATPase-helicase subunit that can be either be SMARCA4 (previously known as BRG1) or SMARCA2 (previously known

as BRM) together with a group of structural or core subunits and a group of modulating subunits.

In 1999, a study showed that only four SWI/SNF subunits were required for the chromatin remodeling activity *in vitro*: the catalytic ATPase-helicase subunit and the structural subunits SMARCB1 (BAF47), SMARCC1 (BAF150), and SMARCC2 (BAF170) (Phelan et al. 1999). However, years later, other studies have demonstrated that those structural subunits, such as SMARCB1, are not always present in the SWI/SNF complexes and can also be replaced by alternative subunits (Mashtalir et al. 2018; Doan et al. 2004).

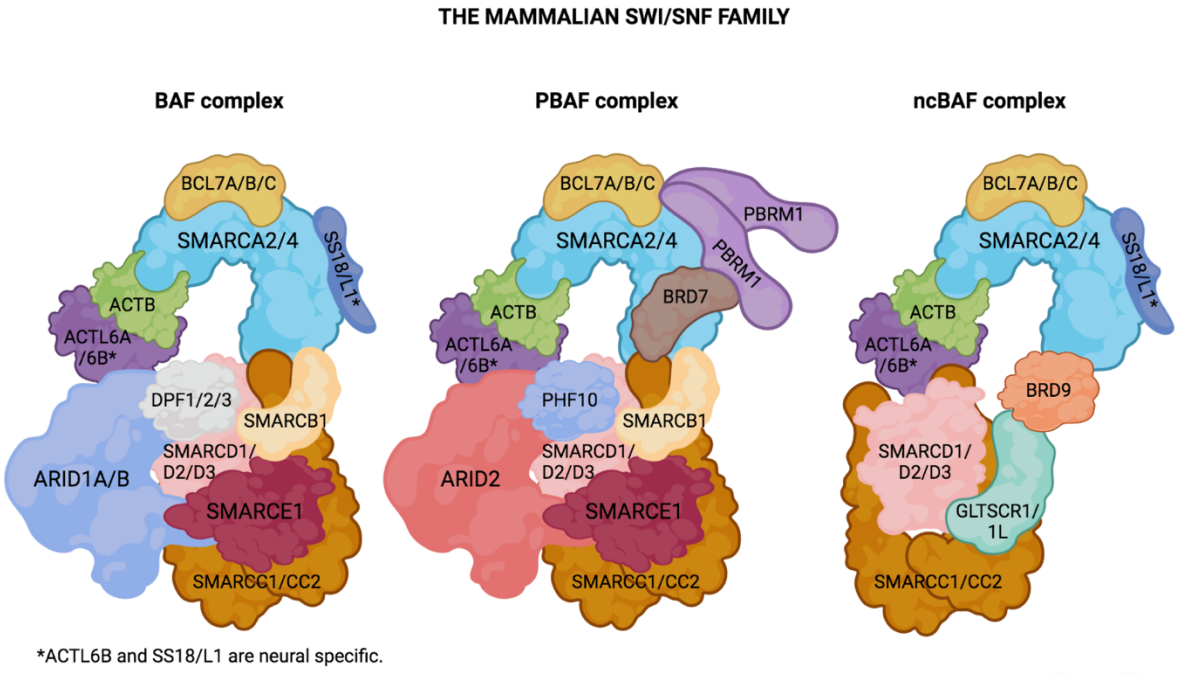


Figure 6: Graphical representation of the three main SWI/SNF complexes in mammalian cells. Different sets of subunits compose the canonical BAF complex (BRM/BRG1 Associated Factors), PBAF (polybromo-associated BAF complexes), and ncBAF complex. Asterisk marks those subunits that are only found in neural-specific SWI/SNF complexes. Image made with BioRender.com and adapted from (Mashtalir et al. 2018).

Annex 1 contains the subunits that can compose the three different configurations of the SWI/SNF complex. The table of ***Annex 1*** includes both the official HUGO (Human Genome Organization) nomenclature that uses the acronym SMARC (**S**WI/SNF-related, **M**atrix-associated, **A**ctin-dependent **R**egulator of **C**hromatin) and the previous nomenclature with the acronym BAF (**B**RG1/BRM-

Associated **F**actor) followed by the molecular weight of each protein. In addition, the table shows a classification of the subunits depending on their functional role in the different SWI/SNF complexes (Mashtalir et al. 2018). Overall, the SWI/SNF remodelers comprise a wide range of complexes that allow the performance of specialized rather than generic functions.

1.3.2. Biochemical roles of the SWI/SNF complex

Although the initial studies of the SWI/SNF complex showed a role in the transcriptional control in yeast, its exact mechanism was unclear. In 1992, two studies elucidated for the first time the mechanism of action of the SWI/SNF complex. Those studies found that some SWI/SNF proteins could promote the expression of specific genes in *Saccharomyces cerevisiae* by modifying chromatin structure (Hirschhorn et al. 1992; Peterson and Herskowitz 1992). A few years later, in 1994, Kwon and colleagues performed the first study that showed that the human SWI/SNF complex mediated nucleosome mobilization through ATP-hydrolysis (Kwon et al. 1994). Other studies have also added that this chromatin remodeling activity could require the **interaction between specific SWI/SNF subunits and different transcription factors** (Valletta et al. 2020; Chang et al. 2018), co-repressors (Zhang et al. 2007; Nagl et al. 2007; Ooi et al. 2006), or nuclear receptors (Belandia et al. 2002; Inoue et al. 2002). Moreover, several studies have shown that the SWI/SNF complex can also modulate gene expression through the interaction with **other chromatin-modifying enzymes**, such as the Polycomb Repressor Complex (PRC) (Bracken et al. 2019; Kadoch et al. 2017; Weber et al. 2021). Thus, one of the principal roles of the SWI/SNF complex is the **regulation of chromatin accessibility leading to activation or repression of gene expression** (Zhang et al. 2007; Nagl et al. 2007). An extensive study that performed chromatin immunoprecipitation sequencing (ChIP-Seq) of different SWI/SNF subunits showed that the **SWI/SNF complex binds to promoters and enhancers**, which are both critical regions that require tight control of the dynamics of nucleosome occupancy (Euskirchen et al. 2011). This study also described that the SWI/SNF complex was located proximal to targets involved in fundamental biological processes. Moreover, many of the functional categories that were significantly overrepresented in this analysis had disease implications, such as cancer. This observation also explains the

relevant role that the SWI/SNF complex has in several **differentiation processes**, which require a tightly regulated transcriptional regulation (Ho et al. 2009; Xu et al. 2012; Lickert et al. 2004; Lessard et al. 2007).

Apart from its direct role in gene expression regulation at a transcriptional level, the **SWI/SNF complex is involved in post-transcriptional pathways, such as alternative splicing**. To date, many studies have described interactions between some subunits of the SWI/SNF and members of the spliceosome, and how these interactions can affect the alternative splicing of crucial genes for the cell, such as *CCND1* (Cyclin D1) (Feng et al. 2021; Underhill et al. 2000; Batsché et al. 2006; Tyagi et al. 2009).

However, besides its function in either transcriptional or post-transcriptional regulation, the SWI/SNF complex also has unique roles in many other cellular processes. One of these **non-transcriptional roles** of the SWI/SNF complex is related to **DNA replication**. Cohen and colleagues showed that SMARCA4, one of the catalytic subunits of this remodeling complex, co-localizes with replication factors at sites of DNA replication (Cohen et al. 2010). Moreover, they observed that SMARCA4 physically associates with TOPBP1, which is a component of the replication fork. They demonstrated that SMARCA4 is involved in replication elongation and how the depletion of this protein significantly decreases the replication elongation rate. Other researchers also found that ARID1A, another SWI/SNF subunit, mediates the physical interaction between the SWI/SNF complex and TOP2 (topoisomerase II) (Dykhuizen et al. 2013). This enzyme is crucial for DNA replication and chromosomal segregation. They showed how depletion of ARID1A caused aberrancies in both DNA replication and chromosomal segregation.

A recent study has also revealed that the SWI/SNF complex plays a **relevant role in preventing replication stress** through its direct interaction with RPA (replication protein A) (Gupta et al. 2020). The authors demonstrated how the loss of SMARCA4-containing SWI/SNF complexes slowed down replication, increased origin firing, and induced replication stress that promoted genomic instability. In addition, Bayona-Feliu and colleagues discovered that the SWI/SNF complex is involved in

the resolution of the so-called transcription-replication conflicts (TRCs) (Bayona-Feliu et al. 2021) through the interaction with factors of the Fanconi's Anemia (FA) pathway. They found that depletion of either SMARCA4 or ARID1A in HeLa cells increased TRCs, and consequently, DNA damage. A similar trend was observed in a different cell model of ovarian clear cell carcinoma, where the authors showed that upon ARID1A loss, there was an increase in TRCs because of the aberrant recruitment of TOP2A (topoisomerase IIA) to R-loops, which led to replication stress (Tsai et al. 2021).

On the other hand, the involvement of the SWI/SNF complex in genomic instability can also be the result of its role in **DNA damage response (DDR)**. Several studies have shown that the SWI/SNF complex is implicated in multiple DDR mechanisms, including nucleotide excision repair (NER) and double-strand break (DSB) repair (extensively reviewed in (Ribeiro-Silva et al. 2019; Harrod et al. 2020)). Regarding NER, Hara and Sancar demonstrated *in vitro* for the first time that DNA damage recognition factors, such as RPA, XPA, and XPC facilitate the remodeling activity of the SWI/SNF complex, which in turn enhances the overall excision activity of the NER machinery (Ryujiro and Aziz 2002). Later, other researchers confirmed the direct interaction between the SWI/SNF subunits and members of the NER pathway (Gong et al. 2006).

Besides the role of the SWI/SNF in NER, other studies have shown the repercussions of the depletion or inactivation of SWI/SNF subunits in **modulating the sensitivity of the cells to DNA damage-inducing agents** by reducing Homologous Repair (HR) and Non-Homologous End-Joining (NHEJ), the two mechanisms to repair DSBs (Watanabe et al. 2014; Agnes et al. 2006; Park et al. 2006; Ogiwara et al. 2011; Shen et al. 2015; de Castro et al. 2017).

Furthemore, other researchers have associated the involvement of the SWI/SNF complex in DDR with transcriptional repression of genes flanking DSBs. This mechanism called **DNA-damage induced transcriptional silencing** is part of an evolutionary process that impedes transcription near DSBs to prevent additional DNA damage, such as translocations (Kakarougkas et al. 2014).

The other non-transcriptional role of the SWI/SNF complex is related to **chromatin architecture**. Euskirchen and colleagues showed that SWI/SNF subunits bind to regions critical for chromosome organization such as CTCF (CCCTC-binding factor), cohesins, lamins, and DNA replication origins (Euskirchen et al. 2011). In addition, other studies have described the interplay between the SWI/SNF complex and centromeric cohesion (Xue et al. 2000; Brownlee et al. 2014). Since centromeric cohesion is crucial for chromosome orientation and proper segregation, the loss of SWI/SNF subunits can cause abnormal anaphase events and aneuploidy.

Figure 7 depicts a summary of the main functional roles of the SWI/SNF complex, excluding its canonical function in nucleosome mobilization.

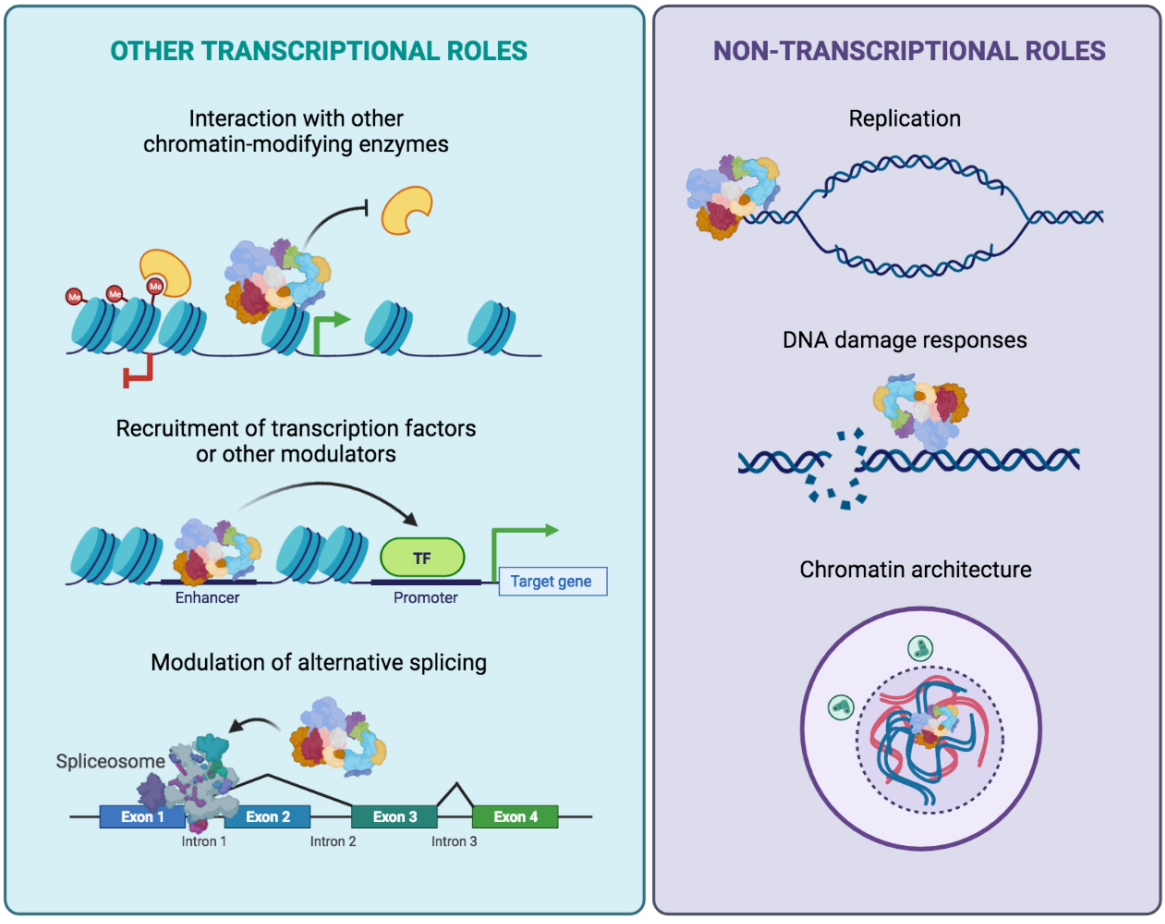


Figure 7: Functional roles of the SWI/SNF complex. Left panel: Transcriptional functions derived from the interactions with other chromatin-modifying enzymes, such as PRC2, transcription factors, or the splicing machinery. Right panel: Non-transcriptional roles of the SWI/SNF complex, such as its involvement in replication, DNA damage responses, and chromatin topology, and 3D organization. Image made with BioRender.com and adapted from (Hodges et al. 2016).

1.3.3. The SWI/SNF complex and disease

Due to the functional diversity of the SWI/SNF complexes, alterations in some of its components are related to the development of many diseases, specifically neurological disorders and cancer.

Several studies have discovered that **SWI/SNF mutations are causative in different neurodevelopmental and intellectual disability syndromes** (Bögershausen and Wollnik 2018). The relevance of the SWI/SNF complex in neurological disorders can be explained by the existence of a tissue-specific complex: neural BAF complexes (nBAFs) (Son and Crabtree 2014). A typical example of a neurodevelopmental disease is Coffin-Siris syndrome (CSS), which is characterized by intellectual disability, frequent infections, and hypoplasia or aplasia of the fifth fingernail and its distal phalange. In CSS patients, 75% of all the mutations affect the *ARID1B* gene (Wieczorek et al. 2013). Other SWI/SNF subunits such as *ARID1A*, *SMARCA4*, *SMARCA2*, *SMARCB1*, and *SMARCE1* were also mutated. Overall, **87% of CSS patients present a mutated SWI/SNF subunit** (Tsurusaki et al. 2012).

On the other hand, the first evidence of a **causative role of SWI/SNF alterations in cancer** was found in 1998 in a rare and aggressive pediatric tumor type: malignant rhabdoid tumors (MRT). Versteege and colleagues demonstrated that MRT patients lost one allele of the *SMARCB1* gene, and the other allele was inactivated by truncating mutations (Versteege et al. 1998). Later, an extensive analysis revealed that **almost 100% of MRTs showed biallelic inactivation of the SWI/SNF gene *SMARCB1*** (Chun et al. 2016). In animal studies, Roberts and colleagues also observed that when they induced somatic homozygous *Smarcb1* inactivation, 100% of mice developed lymphomas and sarcomas at a median of only 11 weeks, which is less than half the time that is required for cancer to form after *Tp53* inactivation (Roberts et al. 2002).

During the last decade, high-throughput studies have revealed that genetic alterations in SWI/SNF subunits occur in a broad spectrum of solid and hematologic

tumors. In general, SWI/SNF mutations can be found in nearly 25% of human cancers (Shain and Pollack 2013; Kadoch et al. 2013; Mittal and Roberts 2020) (***Figure 8***). This breadth of the SWI/SNF involvement in cancer resembles that of relevant tumor suppressors such as TP53, albeit with lower mutation frequencies (Kadoch et al. 2013).

In general, the SWI/SNF subunits are considered as **the most frequently mutated chromatin-related cancer genes** (Kadoch et al. 2013), and both their genetic alterations and expression changes are **prognostic markers for survival across several tumor types** (Savas and Skardasi 2018; Cho et al. 2013; Bai et al. 2013; Endo et al. 2013).

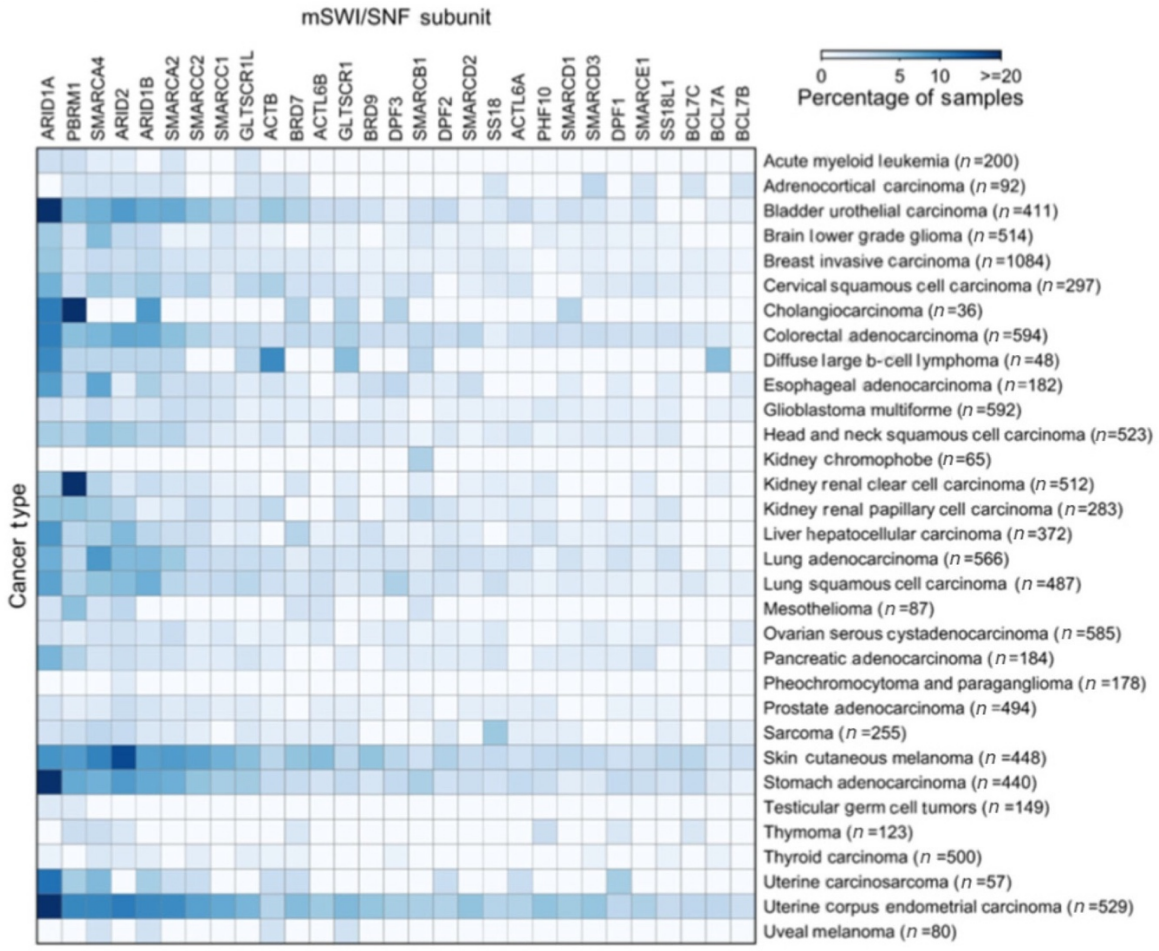


Figure 8: Heatmap of the mutation frequencies of the SWI/SNF subunits in different cancer types. This mutational study analyzed 29 SWI/SNF subunits in 32 cancer types using The Cancer Genome Atlas (TCGA) Pan-Cancer data set (N = 10 967) (Centore et al. 2020).

Although some SWI/SNF subunits alterations are characteristic of specific malignancies, such as the loss of SMARCB1 in MRTs and the SS18-SSX fusion in synovial sarcoma (Clark et al. 1994), other SWI/SNF subunits are altered in several cancers. For instance, ***ARID1A*, *ARID2*, *SMARCA4*, and *PBRM1* are driver genes in three or more tumor types** (Bailey et al. 2018).

To date, many studies have demonstrated that SWI/SNF subunit alterations are tumor-promoting. For example, genetically engineered mouse models have shown that inactivation of *Smarcb1*, *Arid1a*, *Smarca4,* or *Pbrm1* boosts tumorigenesis (Mathur et al. 2017; Gu et al. 2017; Bultman et al. 2008; Roberts et al. 2002). In addition, other researchers have observed that *in vitro* models of re-expression of SMARCB1, ARID1A, SMARCA4, or SMARCA2 in distinct tumor types can lead to growth arrest (Betz et al. 2002; Guan et al. 2011; Karnezis et al. 2016). Thus, all of these articles supported the **tumor-suppressive role** of these SWI/SNF subunits. Nevertheless, the cancer-driving mechanism behind an SWI/SNF-mutant complex is more complicated than just the absence of an SWI/SNF subunit. Several researchers have discovered that SWI/SNF-mutant tumors present **residual SWI/SNF complexes** that are essential for cell survival. This is why the SWI/SNF complex has gained more interest in the clinic. Many articles have described a process called **synthetic lethality** that arises among different SWI/SNF subunits in SWI/SNF-mutant contexts (reviewed in (Centore et al. 2020). For example, some SMARCB1-mutant tumors depend on SMARCA4 for maintaining cell growth. This shows that the oncogenic effect of the SMARCB1 loss does not result from SWI/SNF inactivation but a residual SWI/SNF complex driven by SMARCA4 (Wang et al. 2009). Moreover, other synthetic lethal relationships are found with other proteins of different pathways, such as EZH2, the catalytic subunit of PRC2. Specifically, SMARCB1-deficient tumors depend on EZH2 for their survival, and therefore, EZH2 inhibitors are showing promising effects in clinical trials with MRTs patients (Kim and Roberts 2016).

Furthermore, during the last years, several articles have introduced a **dualistic role of the SWI/SNF complex in cancer** that is related to its high context-dependency. In fact, when analyzing The Cancer Genome Atlas (TCGA) data, researchers have

found amplification, overexpression, and somatic and potentially activating missense mutations in many SWI/SNF subunits, including the well-known tumor suppressors SMARCA4 and SMARCB1 (reviewed in (Orlando et al. 2019)). Moreover, many *in vitro* studies have confirmed this duality of the SWI/SNF in cancer. For example, although SMARCB1 loss is associated with tumor progression in MRTs, it is an essential gene in other tumor types or tissues (Meyers et al. 2017; Agnes et al. 2006). This shows that SMARCB1 loss is only tolerated in unique cell states because of the residual SWI/SNF complexes or their genetic background.

In general, these studies have raised important questions about how the cellular context determines the role of an aberrant SWI/SNF complex in human tumor development.

1.4. Two SWI/SNF subunits: two LUAD drivers

The rapid expansion and development of high-throughput techniques have allowed for more accurate molecular profiling of the different cancer types. In lung adenocarcinoma, four whole-exome or genome sequencing studies have shown the relevance of the SWI/SNF complex in this tumor type (Bailey et al. 2018; Campbell et al. 2016; Imielinski et al. 2012; Network 2014). Specifically, two SWI/SNF subunits, SMARCA4 and ARID1A, are found in the top 15 of LUAD drivers (Bailey et al. 2018). However, the role of both proteins in LUAD initiation and progression remains under study.

1.4.1. SMARCA4

SMARCA4 (BRG1) is one of two mutually exclusive ATPases of the SWI/SNF complex, and it is the most frequently mutated ATPase subunit of all the chromatin remodeling families (Hodges et al. 2016; Kadoch et al. 2013).

SMARCA4 mutations are enriched in diverse cancer types such as Non-Small Cell Lung Cancer (NSCLC) (Dagogo-Jack et al. 2020; Campbell et al. 2016; Rodriguez-Nieto et al. 2010; Medina et al. 2008a; 2004; Wong et al. 2000), Small Cell Carcinoma

of the Ovary Hypercalcemic Type (SCCOHT) (Witkowski et al. 2014; Jelinic et al. 2014), melanoma (Hodis et al. 2012), thoracic sarcomas (Le Loarer et al. 2015), and lymphomas (Love et al. 2012; Lunning and Green 2015). On the contrary, its paralog SMARCA2 (BRM) is rarely mutated (<2%), but rather, it is epigenetically silenced (Karnezis et al. 2016; Kahali et al. 2014; Marquez et al. 2015).

The first time that researchers observed that SMARCA4 was a target of mutations in cancer was in 2000 in a study performed by Wong and colleagues. They also found that re-expression of **SMARCA4 functioned as a tumor suppressor** in multiple types of tumor cell lines (Wong et al. 2000). In the same year, another study showed that mice with monoallelic inactivation of *Smarca4* had a marked predisposition to develop tumors (Bultman et al. 2000). This tumor-suppressive role of SMARCA4 was confirmed by other articles that discovered the importance of this protein in cell cycle control by regulating relevant cell cycle modulators such as p21 and the retinoblastoma protein RB (Dunaief et al. 1994; Kang et al. 2004). Indeed, multiple studies have demonstrated that the deregulation of SMARCA4 alters the cellular transcriptome to increase the expression of genes that promote malignant proliferation (Stanton et al. 2017; Hodges et al. 2018).

Although SMARCA4 mutations are found in a variety of human tumors, SMARCA4 inactivation either by genetic alterations or other epigenetic mechanisms is **especially prevalent in NSCLC** (Dagogo-Jack et al. 2020; Campbell et al. 2016; Marquez et al. 2015; Rodriguez-Nieto et al. 2010; Medina et al. 2008a; 2004; Reisman et al. 2003; Wong et al. 2000). In this tumor type, the loss of both SMARCA4 and SMARCA2 is associated with a worse prognosis (Reisman et al. 2003; Fukuoka et al. 2004). In addition, a recent study has revealed that one-third of the *SMARCA4* mutations in NSCLC correspond to truncating alterations, which 84% of them induce protein loss (Dagogo-Jack et al. 2020). Therefore, this indicates that these genetic alterations often occur in the context of LOH. Other researchers have also shown that *SMARCA4* missense mutations can be deleterious (Hodges et al. 2018).

In vivo studies have found that the deletion of *Smarca4* in a carcinogen-induced lung cancer model enhances lung cancer progression and promotes metastasis (Glaros et

al. 2008). Moreover, *in vitro* analyses have demonstrated that SMARCA4 inactivation promotes NSCLC aggressiveness by changing chromatin organization at regions that include genes implicated in the etiology of NSCLC (Orvis et al. 2014). Additional studies have also confirmed this effect of SMARCA4 on gene expression by its direct interaction with gene regulatory regions (Medina et al. 2005; Banine et al. 2005; Romero et al. 2012; Xue et al. 2019). Importantly, Song and colleagues discovered that re-expression of SMARCA4 activated more genes than other treatments such as the DNMT inhibitor 5dAzaC or an HDAC inhibitor (Song et al. 2014). This result indicated that, although many studies in cancer are focused on the mechanisms of DNA methylation and histone modifications as inducers of gene silencing, SMARCA4 loss has a remarkable impact on epigenetic silencing during NSCLC development.

Specifically, **in LUAD, the main subtype of NSCLC, *SMARCA4* is the SWI/SNF gene most frequently mutated** (Bailey et al. 2018; Imielinski et al. 2012; Network 2014) with a mutation rate of 7% in primary tumors, and it is considered as a LUAD driver gene (Bailey et al. 2018). A recent study performed in a SMARCA4-deficient LUAD cell line showed that re-expression of SMARCA4 increased chromatin accessibility and reactivated genes involved in epithelial cell differentiation, regulation of cell morphogenesis, and development (Lazar et al. 2020). However, despite the evidence that supports the genome-wide regulatory activity of SMARCA4, most studies so far have focused on the regulation of protein-coding genes. Thus, this neglects another layer of regulation that is equally important in tumorigenesis and that is attracting interest in the clinic: the non-protein-coding part of the genome.

1.4.2. ARID1A

ARID1A (the AT-rich interactive domain 1A gene) is an SWI/SNF subunit characterized for having a 100-amino acid AT-rich interaction domain (ARID), which interacts with DNA in a sequence-nonspecific manner (Dallas et al. 2000; Wilsker et al. 2004). Thus, it is responsible for the binding of the SWI/SNF complex to DNA. The ARID1A subunit is present in SWI/SNF complexes in a mutually exclusive fashion with two homologous subunits, ARID1B and ARID2, which also have the DNA-binding ARID domain. Apart from the DNA-binding domain, the C-terminal region of ARID1A also contains a domain that interacts with other

SWI/SNF complex subunits, serving as a hub for the SWI/SNF complex assembly (Mashtalir et al. 2018). Moreover, in that C-terminal region, there are LXXLL motifs that facilitate the interaction with nuclear hormone receptors (Heery et al. 1997; Nie et al. 2000). Overall, ARID1A confers specificity to the SWI/SNF complex by recruiting the complex to its targets through either protein-DNA or protein-protein interactions.

ARID1A is **the most mutated SWI/SNF subunit gene in cancer** (Wu and Roberts 2013). Indeed, fourteen tumor types have *ARID1A* as a driver gene (Bailey et al. 2018). For instance, 50% of ovarian clear cell carcinomas (OCCCs) (Jones et al. 2010), 36% of pancreatic adenocarcinomas (Birnbaum et al. 2011), 35% of breast cancers (Cornen et al. 2012), 30% of endometroid carcinomas (Wiegand et al. 2010), and 29% of gastric cancers (Wang et al. 2011) harbor mutations or chromosomal deletions in *ARID1A*. These *ARID1A* mutations occur across the length of the gene, and most of them are inactivating alterations (nonsense or frameshift) that result in the loss of ARID1A protein (Hodges et al. 2016). Thus, this shows a **loss of function** mechanism of ARID1A in a broad spectrum of human cancers.

In general, tumor suppressor genes can be categorized as "caretakers" or "gatekeepers" (Kinzler and Vogelstein 1997). Gatekeepers control cellular proliferation by regulating the cell cycle or promoting apoptosis. Caretakers are responsible for maintaining the integrity of the genome. In the case of ARID1A, different functional studies have shown a **dual tumor-suppressive role as gatekeeper and caretaker** in various types of cancer (reviewed in (Wu et al. 2014)). However, other researchers have also revealed an additional layer of complexity of ARID1A in tumorigenesis: its role as an **oncogene**. In 2017, Sun and colleagues demonstrated the oncogenic and tumor-suppressive networks of ARID1A in liver cancer (Sun et al. 2017). They observed that ARID1A behaved as an oncogene during tumor initiation, whereas it had a tumor suppressor capacity during tumor progression and metastasis. In addition, a previous study also showed that 83% of hepatocellular carcinoma (HCC) tumors presented ARID1A overexpression in comparison with adjacent tissues (Zhao et al. 2016). In fact, Sun and colleagues found that the survival of HCC patients negatively correlated with

ARID1A expression levels (Sun et al. 2017). This dualistic role of ARID1A was also observed in colorectal cancer (Mathur et al. 2017). On the one hand, Mathur and colleagues found that ARID1A inactivation impaired cellular growth in a particular genetic context: *Apc*-mutant colorectal cancer. On the other hand, they discovered that, in a model of colon cancer of the hypermutated/microsatellite-instable (MSI) type, ARID1A loss drove the formation of invasive colon tumors. In endometrial cancer, Gibson and colleagues observed that primary tumors expressed wild-type *ARID1A,* but metastatic samples from the same patient harbored deleterious mutations in this gene (Gibson et al. 2016). Thus, this result suggested that ARID1A could be required in the early stages of endometrial cancer. Overall, the role of ARID1A in cancer is rather conflictive and **requires a better understanding of its context-specific functions**.

In lung adenocarcinoma, although only 6-8% of patients harbor mutations in *ARID1A* (Imielinski et al. 2012; Network 2014), which is a low mutation frequency in comparison with other malignancies, Bailey and colleagues showed that *ARID1A* is also a driver gene in this tumor type (Bailey et al. 2018). Another study also revealed the presence of *ARID1A* mutations in the plasma DNA of 12% of LUAD patients (Karachaliou et al. 2018). Moreover, a recent analysis of the mutational status of *ARID1A* of an NSCLC cohort, comprised of 80% of lung adenocarcinomas, confirmed a tumor suppressor profile of *ARID1A* with mutations spread throughout the gene and a great proportion of loss-of-function (LOF) alterations (Hung et al. 2020a). In addition, they observed that a complete loss or a reduced ARID1A protein expression in tumor samples was significantly associated with *ARID1A* LOF mutations. They also found that *ARID1A*-mutant tumors showed a higher mutational burden than *ARID1A*-wild-type tumors confirming the previously mentioned role of ARID1A as a caretaker to maintain genome stability (Wu et al. 2014). However, the mutational status of ARID1A did not correlate with differences in survival (Hung et al. 2020a). Instead, among *ARID1A*-mutant NSCLC patients, an aberrant expression of ARID1A was significantly associated with decreased overall survival. Another article also found that the loss of ARID1A expression was an independent prognostic factor of shorter survival in NSCLC (Jang et al. 2020).

Nevertheless, **only 1-2% of NSCLC show loss of ARID1A expression** (Hung et al. 2020a; Naito et al. 2019), indicating that the majority of NSCLC tumors express ARID1A.

Given the context-dependent functions of ARID1A in other cancer types, considering *ARID1A* as a tumor suppressor gene because of its mutational profile is a simplification. Therefore, there is a need for functional studies to elucidate the role of ARID1A in lung adenocarcinoma.

Objectives

2) OBJECTIVES

In this Ph.D. thesis, our aim was to provide a comprehensive perspective of the SWI/SNF complex in LUAD patients and cell models and to focus on some of the open-ended questions that are currently in the field.

For that purpose, we set the following objectives:

1. To study the composition of the SWI/SNF complex in lung epithelial cells.
2. To evaluate the mutational status of the SWI/SNF complex in LUAD.
    a. In a cohort of 70 LUAD patients.
    b. In a panel of 38 LUAD cell lines.
3. To study the clinical impact of the mutational status of the SWI/SNF complex.
4. To characterize the expression profile of the SWI/SNF subunits.
    a. In a cohort of 70 LUAD patients.
    b. In a panel of 38 LUAD cell lines.
5. To assess the impact of the SWI/SNF catalytic subunit SMARCA4 on the regulation of the non-protein-coding genome.
6. To evaluate the role of ARID1A, the most mutated SWI/SNF subunit, in LUAD cell lines.

Materials &Methods

3) MATERIALS AND METHODS

3.1. Cell culture

3.1.1. Cell lines

38 lung adenocarcinoma cell lines (detailed in ***Annex 2***) were grown under the following standard conditions: humidified 5% CO_2 atmosphere at 37 °C in Dulbecco's modified Eagle's medium (DMEM) or Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with glutamine, 10% Fetal Bovine Serum (FBS), and 1% penicillin/streptomycin/amphotericin. Normal bronchial epithelial cells, NL20, and its tumorigenic version NL20-TA were grown in Ham's F12 medium with 4% FBS, 2.0 mM L-glutamine, 1.5 g/L sodium bicarbonate, 2.7 g/L glucose, 0.1 mM nonessential amino acids, 1 µg/mL transferrin, 5 µg/mL insulin, 10 ng/ml Epidermal Growth Factor (EGF), and 500 ng/mL hydrocortisone. All cells were obtained from the ATCC (American Type Culture Collection) where they were authenticated by multiplex PCR (Polymerase Chain Reaction) of minisatellite markers. All cell lines tested negative for mycoplasma contamination.

3.1.2. Transfections

A) *For SMARCA4 restoration*

The A549 LUAD cell line was transiently transfected with a plasmid encoding the isoform E of SMARCA4, which is the most abundant isoform of this protein in lung (Romero et al. 2012). This plasmid was provided by the authors of the previously cited article. The transfections were performed with Lipofectamine 2000 (Thermo Scientific) following the manufacturer's instructions and using Opti-MEM (Fisher Scientific), as a reduced serum medium. Specifically, one million A549 cells were plated in a 100 mm petri dish. After 24 h of the seeding, they were transfected with the empty vector (PCDNA4) or the SMARCA4-plasmid (SMARCA4-PCDNA4) following the protocol of Lipofectamine 2000's manufacturer. 24 h later, complete growth media was added to the cells. At this time point or after 48 h, depending on

the study's aims, cell pellets were collected or used for the functional assays described below.

B) *For microRNA overexpression*

Negative control miRNA mimic (mirVana® miRNA mimic Negative Control # 4464058) and miR-222 mimic (mirVana® miRNA mimic #4464066) were purchased from Thermo Scientific. We plated 250 000 cells of the A549 cell line per well in a 6 well plate and 24 h later, the mimics (at a final concentration of 40 nM) were added to the wells with Lipofectamine RNAimax (Thermo Scientific) and Opti-MEM following the manufacturer's instructions. After 24 h of transfection, complete growth media was added to the cells and 24 h later (48 h post-transfection), cell pellets were collected or used for the functional assays described below.

C) *For ARID1A silencing*

Two different silencing RNAs (siRNAs) were purchased from Thermo Scientific to inhibit the expression of ARID1A: Silencer® Select Pre-designed siRNA #4392420 s15784 (siARID1A#1) and s15785 (siARID1A#2). A nonsense scrambled RNA (#4390843; Thermo Scientific) was used as a negative control (mentioned as "SCR" in the text). Six different cell lines (A549, H2009, H358, H1395, NL20, and NL20-TA) were transfected with a final concentration of 40 nM siRNAs following the Lipofectamine RNAimax's protocol. Specifically, 250 000 A549 cells, 400 000 H2009 cells, 350 000 H358 cells, 400 000 H1395 cells, 700 000 NL20 cells, or 500 000 NL20-TA cells were plated per well of 6 well plates. After 24 h of the seeding, the cells were transfected with Lipofectamine RNAimax and the siRNAs. 24 h later, the growth media was replaced by complete growth media. At this time point or 48 h post-transfection, cell pellets were collected or used for the functional assays described below.

3.1.3. Treatments

After 48 h of the transfection of the A549 cell line with the siRNAs specific for ARID1A silencing or the negative control siRNA, the transfected cells were treated

with doxorubicin (provided by Dr. Rosario M. Sanchez's group) or etoposide (provided by Dr. David Landeira's group) at a final concentration of 0.5 μM and 10 μM, respectively. After different incubation times with these drugs, cell pellets were collected or cell viability was measured as described in the next section.

3.1.4. Functional assays

A) *Cell viability assays*

Cell viability was measured by resazurin assays at different time points. 24 h or 48 h post-transfection, 1000 A549 cells, 2000 H2009 cells, 2000 H358 cells, 4000 H1395 cells, or 5000 cells of either NL20 or NL20-TA, were plated per well in 96 well plates. For all the experiments, a time zero plate was included to normalize the seeding procedure. For each of the time points evaluated, the cells were treated with a final concentration of 0.12 mM resazurin sodium salt (Sigma Aldrich) and incubated for 4 h prior to the addition of 3% sodium dodecyl sulfate (SDS). Fluorescence was measured in a Glomax® Discover Multimode Microplate Reader (Promega) (excitation fluorescence 520 nm and emission fluorescence 580-640 nm).

B) *Colony assays*

After 48 h of transfection, clonogenicity was evaluated by seeding 1000 A549 cells, 2000 H2009 cells, 2000 H358 cells, or 4000 cells of either H1395, NL20, or NL20-TA per well in 6-well plates. Fourteen days later, colonies were stained for 15 minutes with a solution of 0.1% Crystal Violet, 1% Formaldehyde, 1% Methanol, PBS, and H_2O. Then, the plates were rinsed with tap water and dried.

C) *Apoptosis assays*

Six days after the transfection of the A549 cell line with the siRNAs specific for ARID1A silencing or the negative control siRNA, 200 000 cells were collected. The cell pellets were washed twice with cold PBS (Phosphate-Buffered Saline) and processed following the protocol of the PE Annexin-V Apoptosis Detection Kit I (#559763, BD Biosciences). The samples were examined by flow cytometry using

BD FACSCanto II (BD Biosciences). The experiments were performed in triplicate acquiring a minimum of 10 000 events per sample.

3.2. Patients

Lung adenocarcinoma patients were diagnosed from August 2008 to January 2016. The inclusion criteria for our patients required the following information: 1) histological diagnosis of lung adenocarcinoma, 2) availability of demographic and clinical data, 3) availability of DNA and RNA samples for genomic and transcriptomic analyses, and 4) provision of signed informed consent. An independent experienced pathologist confirmed all diagnoses with pathological examinations.

Our patient cohort was homogeneous and no statistically significant differences were found in terms of age, sex, stage, relapse, or survival when comparing the different subgroups. Patients included 50 men and 20 women, whose ages ranged from 47.6 to 83.2 years. This cohort had a median age of 66.1 years at the diagnosis of LUAD, a median time to relapse of 17.4 months, and an overall survival of 20.1 months. The main characteristics of these 70 patients are shown in ***Table 1***.

Table 1: Clinicopathological features of our Spanish LUAD cohort and the LUAD cohort of The Cancer Genome Atlas (TCGA-LUAD project, last updated October 1, 2019).

	Spanish cohort (n=70)	**TCGA (n=522)**
Sex		
Male	20 (28.6%)	242 (46,4%)
Female	50 (71.4%)	280 (53,6%)
Age		
<60 years	30 (42.9%)	137 (26,2%)
>60 years	40 (57.1%)	354 (67,8%)
Stage		
Stage I	18 (25.7%)	279 (53,4%)
Stage II	7 (10%)	124 (23,6%)
Stage III	8 (11.4%)	85 (16.3%)
Stage IV	1 (1.4%)	26 (5%)

3.3. DNA and RNA extraction

3.3.1. From LUAD cell lines

DNA and RNA were extracted using E.Z.N.A.® Tissue DNA kit and TRI Reagent® (Sigma Aldrich), respectively. Cell pellets were collected at 80% confluence.

3.3.2. From LUAD patients

DNA and RNA from 70 LUAD tumors and their paired normal adjacent tissues were obtained from the Basque Biobank (www.biobancovasco.org) and were processed following standard operating procedures. The study was approved by the Ethics Committee (CEI Granada), Department of Health, Government of Andalucía and from the Basque Foundation for Health Innovation and Research, Spain. The signed informed consent, following the procedures of the Declaration of Helsinki and institutional and national guidelines, was obtained from all participants.

3.4. High-throughput techniques

3.4.1. Liquid chromatography with tandem mass spectrometry (LC-MS/MS)

A) *Immunoprecipitation and sample preparation*

Lysates from NL20 cells were prepared in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and 50 mM Tris-HCl pH 7.5) containing protease and phosphatase inhibitors (0.2 mM PMSF, 7 mM OV_4, and 1x complete Mini EDTA-free Protease Inhibitor Cocktail Tablets). 5 mg of protein from each condition were immunoprecipitated overnight at 4 °C using 1 µg antibody (Anti-SMARCA4 (G-7), sc-17796, Santa Cruz Biotechnology) per µg of total protein. In each experiment, one sample with an irrelevant antibody (anti-IgG) was included as a negative control of nonspecific binding. Immune complexes were recovered by adding 200 µL of Dynabeads Protein G (#10004D, Thermo Scientific) and incubating the samples for 3 hours at 4 °C. Beads were washed three times with 1x PBS containing protease inhibitors. The final elution was performed with 8 M urea in 0.1 M Tris-HCl pH 8.

Eluates were processed by the Proteomics Unit of the Spanish National Cancer Research Center (CNIO) following the standard Filter Aided Sample Preparation (FASP) protocol. Proteins were reduced (30 min incubation at Room Temperature RT with 15 mM TCEP (Tris-(2-carboxyethyl)-phosphine)), alkylated (20 min incubation at RT in the dark with 50 mM CAA (chloroacetamide)), and sequentially digested with the lysyl endopeptidase Lys-C (FUJIFILM Wako Chemical Corporation) (protein:enzyme ratio 1:50, o/n at RT) and trypsin (Promega) (protein:enzyme ratio 1:100, 6 h at 37 °C). The resulting peptides were desalted using C_{18} stage-tips.

B) *Mass spectrometry*

Liquid chromatography with tandem mass spectrometry (LC-MS/MS) was performed by the Proteomics Unit of the CNIO. They coupled the UltiMate 3000 HPLC system with a Q Exactive Plus mass spectrometer (Thermo Scientific).

Peptides were loaded into a trap column Acclaim™ PepMap™ 100 C_{18} LC Columns 5 µm, 20 mm length) for 3 min at a flow rate of 10 µl/min in 0.1% Formic Acid (FA). Then peptides were transferred to an analytical column (PepMap RSLC C18 2 µm, 75 µm x 50 cm) and separated using a 90 min linear gradient (buffer A: 4% acetonitrile (ACN), 0.1% FA; buffer B: 100% ACN, 0.1% FA) at a flow rate of 250 nL/min. The gradient used was: 0-3 min 4% B, 5-7.5 min 6% B, 7.5-60 min 17.5% B, 60-72.5 min 21.5% B, 72.5-80 min 25% B, 80-94 min 42.5% B, 94-100 min 98% B, 100-104.5 min 4% B, 105-110 min 0% B. The mass spectrometer was operated in a data-dependent mode, with an automatic switch between MS (350-1400 m/z) and MS/MS scans using a top 15 method (intensity threshold signal ≥ 3.9e4, z ≥2). MS spectra were acquired in the Orbitrap with a resolution of 70,000 FWHM (200 m/z) and MS/MS spectra with a resolution of 17,500 FWHM (200 m/z). An active exclusion of 40 sec was used. Peptides were isolated using a 2 Thompson unit (Th) window and fragmented using higher-energy collisional dissociation (HCD) with normalized collision energy of 27. The ion target values were 3E6 for MS (25 ms max injection time) and 1E5 for MS/MS (90 ms max injection time).

C) *Mass spectrometry data analysis*

Raw files were processed with MaxQuant (v1.6.2.6a) using the standard settings against a human protein database (UniProtKB/Swiss-Prot, 20,373 sequences) supplemented with contaminants. Label-free quantification was done with match between runs (match window of 0.7 min and alignment window of 20 min). Carbamidomethylation of cysteines was set as a fixed modification whereas oxidation of methionines and protein N-term acetylation as variable modifications. Minimal peptide length was set to seven amino acids and a maximum of two tryptic missed-cleavages were allowed. Results were filtered at 0.01 False Discovery Rate (FDR) (peptide and protein level). Then, the "proteinGroup.txt" file was loaded in Perseus (v1.6.0.7) for further statistical analysis. Missing values were imputed from the observed normal distribution of intensities. To define potential interactors, a one-sided T-test was performed requiring at least two LFQ (relative Label-Free Quantification) valid values in the "bait" group, FDR < 0.15 and a $\log_2$ ratio > 2.

3.4.2. Deep sequencing

A) *Gene capture and targeted sequencing*

The baits for the gene capture were designed using the NimbleDesign software (Roche, v4.0). The baits were targeted against 20 SWI/SNF genes and the top 10 LUAD drivers identified by Bailey and colleagues (Bailey et al. 2018) (see ***Annex 3***). We included the known LUAD drivers as positive controls. The design spanned the exons (including UTRs) of all target genes. Each target was padded by 10 nucleotides at 5' and 3' in order to ensure the inclusion of splice regions.

The library preparation and gene capture protocol were performed by the Genomic Unit of GENYO (Centre for Genomics and Oncological Research: Pfizer/University of Granada/Andalusian Regional Government) using the SeqCap EZ Choice Enrichment kit (Roche).

300 ng of genomic DNA were fragmented using a Covaris S2 sonicator yielding 180-220 base pair (bp) fragments. After end repair and adapter ligation, the adapter-

ligated fragments were amplified by PCR (9 cycles). The PCR amplicons were purified and the fragments with the correct size were selected. DNA was denatured and hybridized against biotinylated probes, which were then captured using streptavidin-bound magnetic beads. The DNA bound to the beads was isolated and amplified by PCR (14 cycles). The quality and the concentration of the DNA was evaluated using NanoDrop 2000 (Thermo Scientific) and BioAnalyzer 2100 (Agilent Technologies). The paired-end sequencing was performed on a NextSeq 500 instrument (Illumina) using a NextSeq 500/550 Mid Output Kit (Illumina), 2x150 cycles.

B) *Deep sequencing data analysis*

The quality of the raw FASTQ sequencing files was evaluated using FastQC (v0.11.5, http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Then, the adapter sequences were removed using Cutadapt (Martin 2011) with the following options: -b AGATCGGAAGAGC -B AGATCGGAAGAGC –q 20 –m 50. After trimming the adapters, the reads were aligned to the hg38 human genome (http://hgdownload.cse.ucsc.edu/goldenPath/hg38/bigZips) using BWA-MEM (v0.7.13-r1126) with the -M option. Afterward, we used Picard (v2.1.1) to convert the SAM files to BAM format, to sort the BAM files, and to mark PCR duplicates. Quality metrics were collected using Qualimap (v2.2.1) (García-Alcalde et al. 2012) and MultiQC (v1.7) (Ewels et al. 2016).

For the paired variant calling on the 27 tumor-normal matched samples, we used Mutect2 (GATK version 4.1.4.0). We generated a panel of normals (PoN) using the sequencing data from our 27 sequenced normal samples, including any mutation found in at least one normal sample (`--min-sample-count 1` option in `CreateSomaticPanelOfNormals`). Although using related normal samples for the creation of the PoN is known to introduce slight biases in the results, we concluded that it was a better approach than using an external PoN because our PoN was able to better capture the sequencing artifacts that were specific to our experimental approach. We ran Mutect2 in paired mode using default parameters and providing our PoN and an external germline resource from gnomAD v3 (https://storage.googleapis.com/gnomad-

public/release/3.0/vcf/genomes/gnomad.genomes.r3.0.sites.vcf.bgz). Then, we used `FilterMutectCalls` with default parameters to filter out false positive variants. We merged, normalized and left-aligned the mutations that passed the filters using BCFtools (HTSlib version 1.7) and we annotated the multi-sample VCFs (Variant Call Format) using ANNOVAR (version 2017-07-17) with the following databases: ensGene (v20170912), 1000g2015aug_all, exac03, avsnp150 and dbnsfp33a. Variants with a minor allele frequency ≥ 0.01 in 1000 Genomes or ExAc were excluded. We also excluded synonymous mutations and non-coding mutations.

For the non-paired mutational analysis, we combined two approaches:

I. In our first approach, we used BCFtools applying the following filters: individual variant QUAL ≥ 20 and either total coverage ≥ 8 or ≥ 5 mutant reads. We also flagged as 'LowFreq' the mutations that had a mutant allele frequency below 20%:

```
bcftools mpileup -f hg38.fa -R primary_targets.bed -q 1 -Q 13 -a
'FORMAT/AD' -Ou ${bam} | bcftools call -vmO z | bcftools filter -e
"%QUAL<20 | ((FMT/AD[0:0]+FMT/AD[0:1])<8 & FMT/AD[0:1]<5)" -s
"LowQual" -O u | bcftools filter -e
"FMT/AD[0:1]/(FMT/AD[0:0]+FMT/AD[0:1])<0.2" -s "LowFreq" -m + -O u |
bcftools sort -O z > ${out}.vcf.gz
```

We merged and annotated the resulting VCFs files as described in the paired-analysis. Then, we filtered out the following mutations: those present in any of our 27 normal samples, those with frequencies above 0.01 in ExAc or 1000 Genomes Project, those that overlapped with simple repeats or low complexity regions according to RepeatMasker (downloaded from http://hgdownload.cse.ucsc.edu/goldenpath/hg38/database/rmsk.txt.gz), and those that were synonymous or non-coding mutations.

II. In our second approach, we applied Mutect2 (GATK version 4.1.4.0) in non-paired mode using default parameters, our in-house PoN, and gnomAD v3 as a germline resource.

We compared our two approaches for the non-paired analyses and we individually evaluated the discrepancies between the two pipelines using Integrative Genomics Viewer (v2.3.94) and public databases (see ***Annex 4***). We decided to combine the results from the BCFtools analysis with manually 'rescued' mutations from the Mutect2 approach after a careful inspection of the Mutect2-exclusive mutations. We flagged such mutations as 'Mutect2'.

3.4.3. MicroRNA Sequencing (miRNA-Seq)

A) *Library preparation and sequencing*

After 48 h of the transfection of the A549 cell line with the empty vector or the SMARCA4-plasmid (explained in Section 3.1.2), the cells from three biological replicates were collected. Their total RNA was extracted using the mirVana RNA Isolation Kit (Thermo Scientific). RNA concentration and quality were analyzed by NanoDrop 2000 (Thermo Scientific) and Bioanalyzer 2100 (Agilent Technologies), respectively. The Genomic Unit of GENYO used Illumina TruSeq® Small RNA Library Prep Kit (Illumina) to generate sequencing libraries following the manufacturer's instructions. Specifically, 1 µg of RNA was the input for the library preparation. RNA samples were barcoded to allow pooling of samples and they were size-selected using acrylamide gel electrophoresis. Sequencing was carried out on a NextSeq 500 System using a NextSeq 500/550 High-Output Kit v2.5 (75 cycles), obtaining an average of 3 million reads per sample.

B) *miRNA-Seq analysis*

miRNA-Seq data analysis was performed using the QuickMIRSeq suite (Zhao et al. 2017). Briefly, raw sequences were trimmed to remove adapter sequences at the 3' ends. Reads shorter than 15 nucleotides (nt) and larger than 28 nt after trimming were discarded to select those sequences more likely to map against miRNAs. Reads were aligned to the human reference genome GRCh38. We normalized the miRNA

raw counts with the trimmed mean of M values (TMM) method (Robinson and Oshlack 2010) implemented in the edgeR R package (Robinson et al. 2010). Differential expression analysis was performed with the DESeq2 R package (Love et al. 2014). We considered as significantly differentially expressed miRNAs those with an FDR corrected *p*-value < 0.05 and an absolute $\log_2$ Fold-Change > 1. We used pheatmap and EnhancedVolcano R packages for generating the heatmap and volcano plot respectively.

C) *Pathway analysis*

The functional relevance of the differentially expressed microRNAs upon SMARCA4 restoration was assessed with DIANA-miRPath v3.0 (Vlachos et al. 2015). This online software combines the experimentally validated miRNA targets of the DIANA-TarBase platform (Karagkouni et al. 2018) with enrichment analyses in Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG).

3.4.4. Messenger RNA Sequencing (mRNA-Seq)

A) *Library preparation and sequencing*

Total RNA from three biological replicates of transfected A549 cells (after 6 days of silencing ARDI1A) was extracted using Direct-zol RNA Kit (Zymo) following manufacturer's instructions. Concentration and quality of extracted RNA were measured using the Qubit 4 Fluorometer (Thermo Scientific) and the 2100 Bioanalyzer Instrument (Agilent Technologies). The Genomic Unit of GENYO prepared the libraries with 1 µg of RNA and the TruSeq Stranded mRNA Library Prep Kit (Illumina) following the manufacturer's protocol. Adapters and samples codes (index-barcodes) were also added to the libraries to allow simultaneous sequencing. mRNA libraries were sequenced on the NextSeq 500 system (Illumina) using the highest output mode and paired-end 75 bp read lengths with a depth of 15 million reads for each sample.

B) *mRNA-Seq analysis*

Quality of the raw FASTQ files was assessed using FastQC (v0.11.8). If present, the 76th nucleotide of the reads was trimmed using bbduk (BBMap version 35.85, "ftr=74" option). The sequencing reads were aligned to the reference human genome (GRCh38.d1.vd1, downloaded from https://api.gdc.cancer.gov/data/254f697d-310d-4d7d-a27b-27fbf767a834) using STAR (v2.6.1b) in "--twopassMode Basic" mode. The "--sjdbOverhang" parameter was set to 75 and all other parameters were set according to the GDC mRNA Analysis Pipeline (https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/Expression_mRNA_Pipeline). The GENCODE v29 GTF file was used as an annotation reference (downloaded from ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_human/release_29/gencode.v29.annotation.gtf.gz). Coordinate-sorted BAM files were obtained. QualiMap (v2.2.1) was used to estimate the quality of the aligned files. Read counts per gene were obtained using htseq-count (v0.11.2) with the following options: "–m intersection-nonempty –f bam –t exon –i gene_id –r pos –s reverse".

The differential expression analysis was performed using the R package 'edgeR' (Bioconductor version 3.8; R version 3.5.3). Genes that did not have more than 5 reads in at least 2 samples, assuming a library size of ~30x2 million mapped reads, were filtered out. Then, data were normalized using the TMM method. Differential expression analyses between siARID1A-treated and SCR-control conditions were performed using edgeR's estimateDisp(robust = TRUE), followed by glmQLFit(robust = TRUE) and glmQLFTest(). The thresholds for statistical and biological significance were set at FDR < 0.05 and absolute (fold change) > 1.5.

C) *Pathway analysis*

Gene set enrichment analyses (GSEA) were performed using the GSEA Java desktop application (https://www.gsea-msigdb.org/gsea/index.jsp) with default settings. Genes were pre-ranked using a "π-score", which combines the information from the effect size and the FDR (Xiao et al. 2014). All MSigDB gene set collections (H, C1-C7) were tested.

3.5. Bioinformatic analyses

3.5.1. *In silico* study of the SWI/SNF complex

A) *In LUAD cell lines*

We downloaded the merged mutation calls from the Cancer Cell Line Encyclopedia (CCLE, https://portals.broadinstitute.org/ccle/data, last updated July 18, 2018). We converted the genomic coordinates of the mutations from hg19 to hg38 using liftOver in R (version 3.5.2, Bioconductor version 3.8). We analyzed the presence of those mutations in the sequencing files of our cell lines using the Integrative Genomics Viewer (IGV) (Robinson et al. 2011).

B) *In LUAD patients*

Mutation, gene expression, and clinical data of LUAD patients were obtained from The Cancer Genome Atlas (TCGA-LUAD project, last updated October 1, 2019). We used the R packages 'TCGAbiolinks' (v2.12.3, R version 3.6.1) and 'cgdsr' (v1.3.0, R version 3.6.1). For the mutation data, we chose the variant calls from the Mutect2 pipeline (N = 567 patients) and we restricted the analysis to the following mutation types: missense_variant, stop_gained, frameshift_variant, splice_acceptor_variant, splice_donor_variant, inframe_insertion, inframe_deletion, start_lost, and stop_lost.

Tumor mutation burden (TMB) estimates for the TCGA-LUAD cohort were predicted by Hoadley et al (Hoadley et al. 2018) and they were downloaded from https://gdc.cancer.gov/about-data/publications/PanCan-CellOfOrigin, file "mutation-load-updated.txt". The rate of non-silent mutations per Mb was used as the TMB value. Values of TMB = 0 were assumed to be errors, and therefore, they were excluded from the analysis.

C) *Evaluation of the functional impact of missense mutations*

To predict the functional impact of missense mutations, the SIFT algorithm (Sim et al. 2012) was used (integrated in Ensembl VEP 95) for both our mutational data (of LUAD cell lines and primary tumors) and TCGA-LUAD mutations. Only one

consequence per variant allele was kept. We considered 'possibly damaging' mutations as 'damaging', and 'possibly tolerated' mutations as 'tolerated'. If a variant originally annotated as 'missense' affected an isoform that was missing in the newer Ensembl 95, we assumed that it was 'tolerated'.

3.5.2. *Three-dimensional organization and interaction study*

The data of high-throughput chromosome conformation capture (Hi-C), chromatin interaction analysis by paired-end tag sequencing (ChIA-PET), and topologically associating domains (TAD) were obtained from the Encyclopedia of DNA Elements: ENCODE (https://www.encodeproject.org, v111). The IDs of the TADs files were ENCFF336WPU (for the A549 cell line) and ENCFF307RGV (for the normal lung fibroblast cell line IMR-90). The IDs of CTCF ChIA-PET files were ENCFF269OED (A549, rep1), ENCFF561KUX (A549, rep2), ENCFF371UFY (IMR-90, rep1), and ENCFF746ZQV (IMR-90, rep2). All genomic coordinates were represented in hg38. TAD coordinates were converted from hg19 to hg38 using UCSC liftOver.

3.5.3. *In silico* study of ARID1A in LUAD patients

ARID1A mRNA expression data from LUAD patients was obtained from GEPIA2 (Gene Expression Profiling Interactive Analysis) (Tang et al. 2019). Tumor values corresponded to the TCGA-LUAD cohort and normal values were a combination of the normal samples of TCGA and GTEx (Genotype-Tissue Expression) databases.

Protein expression data of ARID1A were retrieved from the analysis of LUAD patients performed by Gillette and colleagues (Gillette et al. 2020) (see Table S3D of the cited publication). Paired t-tests were performed with these data.

3.6. Real-time quantitative polymerase chain reaction (RT-qPCR)

3.6.1. For detection of protein-coding-RNAs

1-2 µg of total RNA were used to prepare cDNA (complementary DNA) with the RevertAid RT Kit (Thermo Scientific). Real-time quantitative PCR (RT-qPCR)

reactions followed KAPA SYBR® FAST qPCR Master Mix recommendations. The RT-qPCRs performed to profile the expression of the SWI/SNF complex in LUAD primary tumors and cell lines were optimized using the 7900HT Real-Time PCR System with its adaptor for 384 well-plates (Applied Biosystems). For the functional assays of SMARCA4 and ARID1A, the RT-qPCRs were performed with QuantStudio™ 3 (Thermo Scientific). Relative expression was calculated using *GAPDH* as reference gene and applying the ΔΔCt method. Primers for each gene are shown in ***Supplementary Table 5*** and ***Supplementary Table 7*** of ***Annex 5***. All experiments were carried out in triplicate.

3.6.2. For detection of miRNAs

2 µg of RNA were polyadenylated with Poly(A) Polymerase Tailing Kit (Lucigen/Epicentre) following the manufacturer's protocol. The cDNA was obtained with the RevertAid RT Kit (Thermo Scientific) and a polyT-adapter (see ***Supplementary Table 6*** of ***Annex 5***). RT-qPCR reactions followed KAPA SYBR® FAST qPCR Master Mix (Merck) recommendations including a universal-reverse primer in the mix (***Supplementary Table 6*** of ***Annex 5***). RT-qPCR was optimized using QuantStudio™ 3 (Thermo Scientific). Relative expression was calculated using *U6* as a reference gene and applying the ΔΔCt method. The experiments were carried out in triplicate.

3.7. Protein techniques

3.7.1. Western Blot

Total protein was extracted from cell pellets lysed with RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, and 50 mM Tris-HCl pH 7.5) supplemented with the following protease and phosphatase inhibitors: 0.2 mM PMSF, 7 mM OV_4 and 1x complete Mini EDTA-free Protease Inhibitor Cocktail Tablets (Thermo Scientific). For cellular fractionation, one million of cells were subjected to the protocol described by Beringer et al (Beringer et al. 2016). Protein concentration was measured using the Bradford method (PanReac AppliChem). 20-50 µg of protein were separated by SDS-PAGE (sodium dodecyl sulfate

polyacrylamide gel electrophoresis) and transferred to PVDF (polyvinylidene difluoride) membranes. After an incubation of an hour with blocking solution (PBS containing 5% nonfat dry milk and 0.1% Tween), the membranes were incubated overnight with the primary antibodies detailed in ***Table 2***. Next, the membranes were incubated with the following secondary antibodies purchased from Dako: anti-rabbit HRP (#P0448, 1:2000), anti-mouse HRP (#P0447, 1:1000); anti-goat HRP (#P0449, 1:1000). The targeted protein bands were visualized using the reagents Clarity™ Western ECL Substrate (BioRad) or SuperSignal™ West Femto (Thermo Scientific) and measured by ImageQuant LAS4000 (GE Healthcare). Band quantification was carried out with ImageJ.

Table 2: List of primary antibodies used for the different objectives of this Ph.D. Thesis.

Protein	Supplier	Reference	Dilution
ARID1A	Cell Signaling Technology	#12354	1:500
ARID1B	Abcam	ab57461	1:500
ARID2	Santa Cruz Biotechnology	sc-166117	1:500
β-ACTIN (ACTB)	Sigma Aldrich	A5441	1:10 000
CASPASE 3	Cell Signaling Technology	#9662	1:500
Cleaved CASPASE 3	Cell Signaling Technology	#9661	1:500
γ-H2AX	Cell Signaling Technology	#9718	1:500
LAMIN B	Santa Cruz Biotechnology	sc-6216	1:1000
c-MYC	Santa Cruz Biotechnology	sc-40	1:500
PARP-1	Santa Cruz Biotechnology	sc-8007	1:100
SMARCA4	Santa Cruz Biotechnology	sc-17796	1:500
SMARCA2	Cell Signaling Technology	#11966	1:1000
SMARCC1	Cell Signaling Technology	#11956	1:1000
α-TUBULIN	Santa Cruz Biotechnology	sc-23948	1:3000

3.7.2. Immunoprecipitation (IP)

Total protein from transfected A549 cells with the empty vector or the SMARCA4-plasmid was extracted as detailed in the previous section. 500 µg of total protein were immunoprecipitated overnight at 4 °C using 2 µg of Anti-SMARCA4 (sc-17796, Santa Cruz Biotechnology) or a 1:50 dilution of either Anti-SMARCA2 (#11966, Cell Signaling) or Anti-SMARCC1 (#11956, Cell Signaling). In each experiment, an irrelevant antibody (anti-IgG-rabbit, #12-370, Merck; anti-IgG-mouse, #12-371,

Merck) was included as a negative control for nonspecific binding, and an input of 5% of the total protein was saved. Immune complexes were recovered by adding Dynabeads Protein G (#10004D, Thermo Scientific) and incubating the samples for 3 hours at 4°C. Beads were washed three times with 1x PBS containing protease inhibitors. Final elution was performed in 50 µL of 2x Loading Buffer (0.1 M Tris-HCl pH 6.8, 4% SDS, 0.2% bromophenol blue, 200 mM DTT, and 20% glycerol). Samples were denatured by heating at 95°C for 10 minutes before performing the SDS-PAGE.

3.8. Chromatin Immunoprecipitation (ChIP)

3.8.1. ChIP-qPCR

ChIP assays for SMARCA4 and SMARCA2 were performed following the protocol of Asenjo H. et al (Asenjo et al. 2020) with minor modifications. After the transfection with the empty vector or the SMARCA4-plasmid, two million of transfected A549 cells at different time points were used for this protocol. Bioruptor® Sonicator (Diagenode) was set to sonicate the cells for 35 cycles 45 sec ON, 15 sec OFF. 1 µg of the antibody of SMARCA4 (ab11064, Abcam) or a 1:100 dilution of the antibody of SMARCA2 (#11966, Cell Signaling Technology) were used per million of cells for immunoprecipitation. For each condition, a control of nonspecific binding (anti-IgG-mouse, #12-371, Merck; or anti-IgG-rabbit, #12-370, Merck) was included at the same amounts as the SMARCA4 or SMARCA2 antibodies. Eluted DNA was quantified with a Qubit™ dsDNA HS (High Sensitivity) Assay Kit (Thermo Scientific). 0.1 ng of DNA were used for measuring absolute levels of binding to the enhancer of miR-222 with the following primers: miR-222-enhancer-FW: 5'-GAGGCAACTCACTTGCC-3', miR-222-enhancer-RV: 5'-CCTGCTTCACCTTGTAAATTC-3'. The reactions were performed using a KAPA SYBR® FAST qPCR Master Mix (Merck) in a QuantStudio™ 3 system (Thermo Scientific) following manufacturer's instructions. Fold enrichment was calculated relative to IgG controls.

3.8.2. Public ChIP-Sequencing (ChIP-Seq) data

The following NCBI Gene Expression Omnibus (GEO) accession numbers correspond to the data used for analyzing the histone mark profile in the A549 cell

line: H3K4me3 (GSE91218), H3K27ac (GSE91337), and H3K4me1 (GSE91306). These data were generated by the ENCODE project (Davis et al. 2018).

3.9. Statistical analyses

3.9.1. Of the study of the SWI/SNF complex in LUAD patients

All statistical analyses of the study of the SWI/SNF complex in LUAD patients were performed using R (version 3.6.1). Normality of the data was assessed using quantile-quantile plots and data transformations and statistical tests were chosen accordingly.

Co-occurrence or mutual exclusion of mutations in gene pairs was analyzed by Fisher's exact tests and *p* values were adjusted for multiple testing using the Benjamini-Hochberg method.

Univariate and stepwise multivariate Cox Proportional-Hazards regressions were performed with the R packages 'survival' (v2.44-1-1) and 'My.stepwise' (v0.1.0). These R packages used the TCGA-LUAD mutation data of 20 SWI/SNF genes and the top 10 LUAD driver genes, as well as clinical covariates (age at diagnosis, gender, and tumor stage). We considered mutations on SWI/SNF genes as one binary variable that classifies patients in those with at least one mutated SWI/SNF subunit and those with wild type SWI/SNF. Only variables significant at $p < 0.2$ were selected for the stepwise multivariate analysis. Kaplan-Meier curves were drawn with the R package 'survminer' (v0.4.6) and compared with the log-rank test.

3.9.2. Of *in vitro* studies

Normality of the data was addressed by the Shapiro-Wilk test. Two-tailed Student's t-tests were applied for normally distributed data. Differences were considered significant at p-value < 0.05. Results were expressed as mean ± SD of at least three different biological replicates.

To test differences in cell growth rate of the combined treatment of silencing *ARID1A* and genotoxic drugs, we first fit the following linear model to the data:

$$log_2(y) = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot trx + \beta_3 \cdot t \cdot trx + \beta_4 \cdot rep + \varepsilon_i$$

Where y is the signal from the cell viability assay, t is time (in hours; t = 0 is the start of the experiment), trx is the treatment, and rep is the replicate. To account for the heteroscedasticity of the data, variance was modeled as:

$$Var(\varepsilon_i) = \sigma^2 \cdot (\delta_1 + t^{\delta_2})^2$$

The model was fit using the `gls(method = "REML")` function from the nlme R package. Pairwise differences in the cell growth rates over time (β_3 in the model) between the conditions of interest were tested using the `emmeans()` function from the `emmeans` R package, followed by `pairs()` and correction of the p values by the Holm method.

3.10. Data availability

3.10.1. Mass spectrometry data

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD017397.

3.10.2. DNA sequencing data of LUAD cell lines

Cell line DNA sequencing data has been uploaded to the European Nucleotide Archive (ENA) under the accession PRJEB40655.

3.10.3. DNA sequencing data of LUAD patients

Human DNA sequencing data has been uploaded to the European Genome-phenome Archive (EGA) under the accession EGAD00001005930.

3.10.4. miRNA-Seq data of the cellular model of SMARCA4 restoration

Raw and processed miRNA-Seq data is publicly available at the GEO repository (GSE167140).

RESULTS

4) RESULTS

4.1. Chapter I: Characterization of the SWI/SNF complex in LUAD

4.1.1. Composition of the SWI/SNF complex in lung epithelial cells

The first objective of this Ph.D. thesis was to identify which subunits constitute the SWI/SNF complex in lung epithelial cells. For this purpose, we performed an endogenous immunoprecipitation (IP) of SMARCA4, one of the catalytic subunits of the SWI/SNF complex that is present in all the SWI/SNF complex subtypes. We used the NL20, a non-tumorigenic bronchial epithelial cell line, as our model of study. The IP was followed by an analysis of liquid chromatography-tandem mass spectrometry/mass spectrometry (LC-MS/MS). In this proteomic analysis, we identified proteins that belong to one of the three human SWI/SNF complexes: BAF, PBAF, and ncBAF. In total, we detected twenty SWI/SNF subunits that were pulled down along with SMARCA4 (***Fig. 9A***, see ***Annex 6***). From now on, we will refer to the SWI/SNF subunits found in this immunoprecipitation, plus SMARCA4 and SMARCA2, which are the two catalytic subunits of this complex, as "lung SWI/SNF subunits". We restricted the rest of the analyses to these subunits. With this study, we noted that the following nine known SWI/SNF subunits did not immunoprecipitate in our lung cellular model: ACTB, ACTL6B, BCL7B, BCL11A, BCL11B, BICRAL (GLTSCR1L), DPF1, SS18, and SS18L1. We were particularly intrigued by the absence of ACTB, which has always been depicted as a stable member of all SWI/SNF complexes described so far (reviewed in (Kadoch and Crabtree 2015)). We validated the absence of ACTB by an immunoprecipitation followed by a Western blot analysis (***Fig. 9B***). Moreover, we observed that other actins (ACTA1, ACTA2, ACTC1, and ACTC2) interacted with SMARCA4 in our normal lung epithelial cell model, according to our LC-MS/MS data (***Annex 6***). These results support a change of the SWI/SNF interactome in lung tissue in comparison with other cellular models previously analyzed.

4.1.2. More than 40% of LUAD patients harbor mutations in lung SWI/SNF subunits

To examine the mutational status of the lung SWI/SNF subunits, we performed targeted DNA sequencing in 70 LUAD primary tumor samples and 27 paired normal adjacent tissue samples. We restricted this analysis to 20 lung SWI/SNF subunits with good quality sequencing results (see ***Annex 3***). Using a gene-targeting protocol, we achieved a median coverage of 142X (Interquartile Range (IQR) = 121X-160X) with a median of ~400,000 reads per sample spanning ~269 kb. A median sample had ≥50X coverage in ~81% of the target nucleotides. For the analysis of somatic point mutations and short indels, we combined several paired and non-paired pipelines to minimize false positives while retaining true somatic mutations (see Materials and Methods and ***Annex 4***).

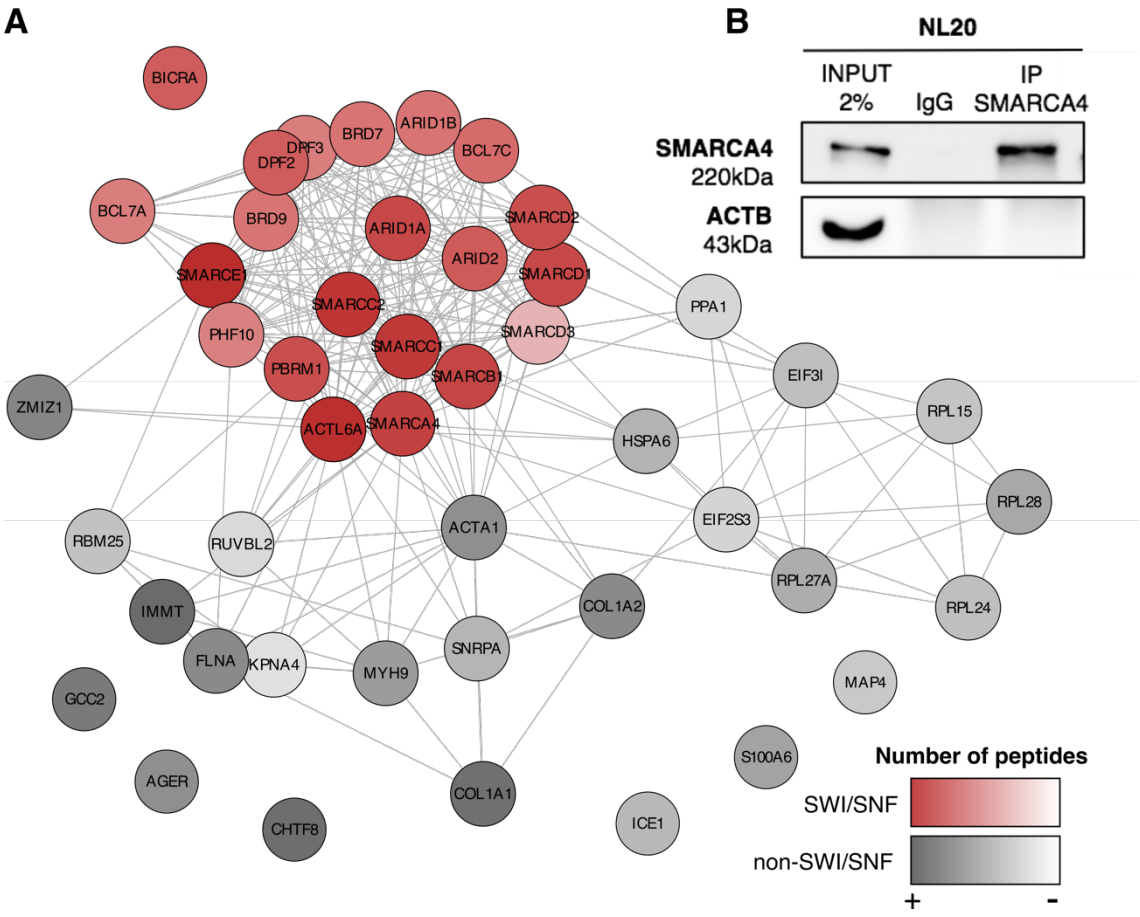


Figure 9: SWI/SNF interactome in lung epithelial cells. (**A**) Protein-protein interactions in NL20 after SMARCA4 IP. SWI/SNF subunits are depicted in red. Non-SWI/SNF subunits are shown in gray. Color intensity is correlated with the number of peptides found. (**B**) Western blot of the IP of SMARCA4 in NL20 that confirms that there was no interaction with ACTB.

We found 38 point mutations and small indels in our LUAD patient cohort (N= 70). Twenty-nine (41.4%) of the primary tumors harbored at least one mutation in a lung SWI/SNF subunit, underlining the importance of SWI/SNF mutations in lung cancer (***Fig. 10A***). The most frequent alterations were missense mutations (65.8%) followed by stop gain (13.2%), frameshift indels (10.5%), splice site alterations (5.3%), and inframe indels (5.3%). In addition, we did not find any recurrent mutations, an observation that agrees with the mutation pattern frequently associated with tumor suppressor genes. *SMARCA4* was the most commonly mutated SWI/SNF gene (11.4% of samples), followed by *ARID1A* (8.6%), *ARID2* (7.1%), *ARID1B* (4.3%), and *PBRM1* (4.3%).

Next, to investigate potential differences in mutation frequencies between our cohort of Spanish LUAD patients and other LUAD cohorts, we examined publicly available data from the TCGA-LUAD project (last updated on October 1, 2019, N = 567). The distributions of the clinical parameters such as age, sex, stage, relapse, and survival statuses were comparable between the two cohorts (***Table 1***). Our patients and the TCGA cohort showed similar mutation frequencies across the 20 lung SWI/SNF genes. However, our cohort presented slightly higher mutation rates (***Fig. 10B***). Specifically, the total mutation frequency of the lung SWI/SNF complex was 41.4% in our cohort and 30.0% in TCGA-LUAD. This difference could be explained due to a greater coverage in our protocol or differences in data analysis protocols. Moreover, we observed that *SMARCA2*, *SMARCD1*, *ACTL6A*, *SMARCE1*, *BCL7A*, and *SMARCD3* were mutated in TCGA-LUAD but not in our cohort, possibly because of our limited sample size, whereas *PHF10* was mutated in our cohort but not in TCGA-LUAD. Nevertheless, regardless of the study group, *SMARCA4*, *ARID1A*, and *ARID2* were the SWI/SNF subunits that accumulated the highest number of truncating mutations.

Since a considerable proportion of the genetic alterations of the lung SWI/SNF genes corresponded to missense mutations, we decided to predict their functional impact using the SIFT algorithm (see Materials and Methods). Based on SIFT predictions, more than half of the missense mutations in our cohort (64%, 16/25) and the external data (65%, 103/159) were “deleterious” (***Fig. 10C***). Overall, considering

the truncating mutations and the predicted deleterious missense mutations, more than 70% of the SWI/SNF mutations may have a functional impact.

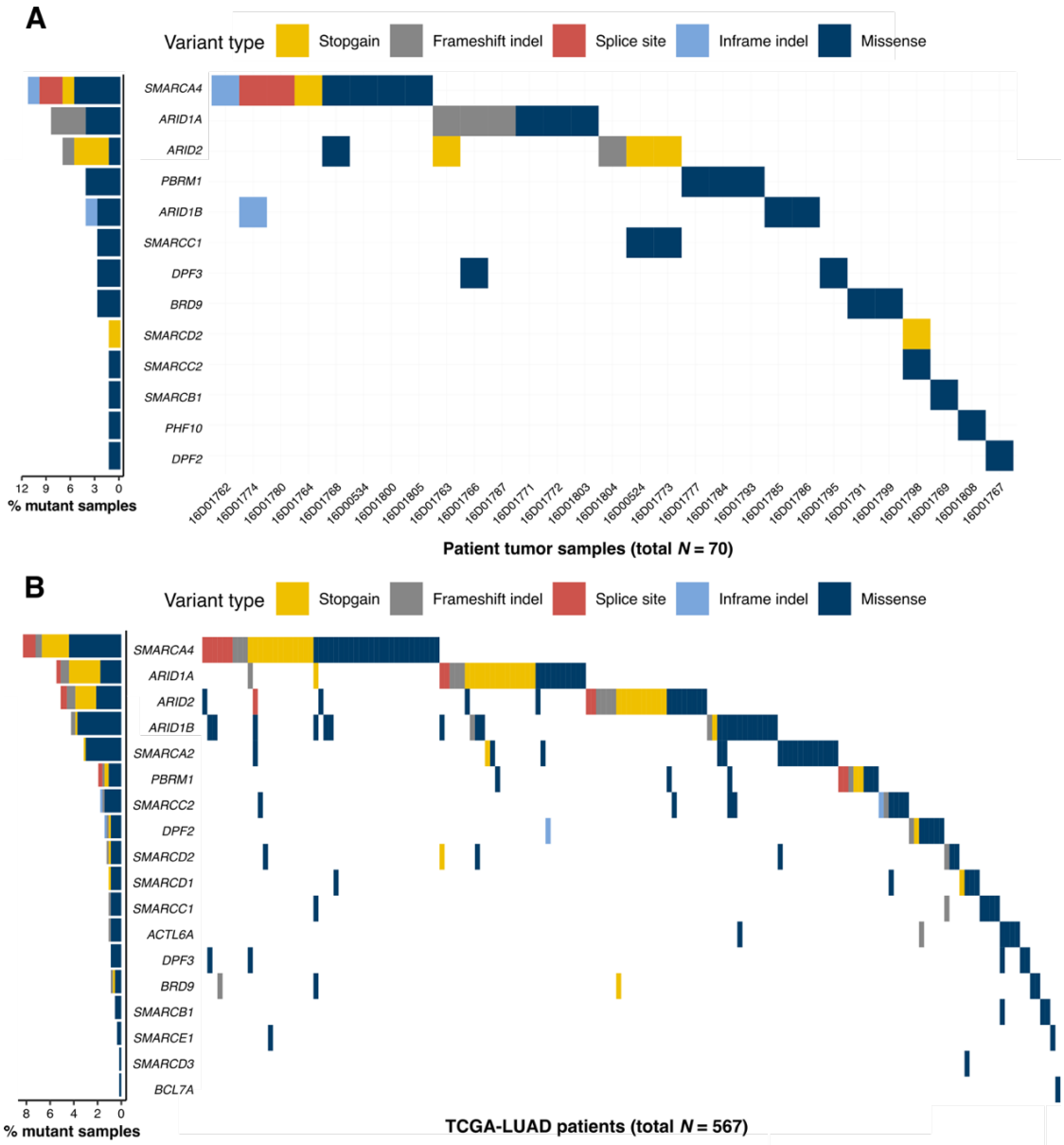


Figure 10: Mutational study of SWI/SNF in LUAD primary tumors. (**A**) Mutation profile of the 20 lung SWI/SNF complex subunits in our LUAD cohort (**B**) or the TCGA-LUAD cohort. Y-axis represents all the subunits that had at least one genetic alteration in at least one LUAD patient. X-axis gathers all LUAD patients with a mutant SWI/SNF complex. On the left, mutation frequencies of these lung SWI/SNF subunits in our LUAD patients. (**C**) Functional prediction of the mutations found in SWI/SNF subunits in our 70 LUAD patients, or in TCGA-LUAD data using SIFT.

4.1.3. More than 75% of LUAD cell lines harbor mutations in lung SWI/SNF subunits

We selected a panel of representative LUAD cell lines that are commonly used (quantified by the number of Pubmed citations, see ***Supplementary Fig. 1A*** of ***Annex 7***) and that combine different genetic and clinical backgrounds (see ***Supplementary Fig. 1B*** and ***Supplementary Table 9*** of ***Annex 7***). We analyzed the mutational status of 20 SWI/SNF subunits and the top five LUAD driver genes (see ***Annex 3***). We also included homozygous deletions in this study by searching for genes or exons that had no reads in a sample but that were properly sequenced in the rest of the samples.

Twenty-nine out of the 38 LUAD cell lines (76.3%) harbored at least one genetic alteration in SWI/SNF genes (***Fig. 11A***). The top five most cited LUAD cell lines presented a mutation affecting a lung SWI/SNF subunit. Specifically, 12 out of the 20 (60.0%) analyzed SWI/SNF subunits were mutated or had homozygous deletions in our panel of LUAD cell lines, accumulating a total of 49 genetic alterations. *SMARCA4* was the top mutated SWI/SNF gene (mutation rate = 42.1%) followed by *ARID2* (15.8%), *ARID1A* (10.5%), *ARID1B* (7.9%), and *SMARCA2* (7.9%). To corroborate our observations with external data, we compared our results with those reported by the Cancer Cell Line Encyclopedia (CCLE). In general, our mutational data highly agreed with the CCLE although we found some discrepancies that are explained in further detail in ***Annex 8*** and ***Supplementary Fig. 2–5***.

Similar to what we observed in LUAD patients, among all genetic alterations that affected SWI/SNF subunits, almost 60% were missense mutations (***Fig. 11B***). For this reason, we decided to predict their functional impact using the SIFT algorithm.

Based on SIFT predictions, we estimated that 41% of missense mutations could be "deleterious" and impact the functionality of the protein (***Fig. 11C***).

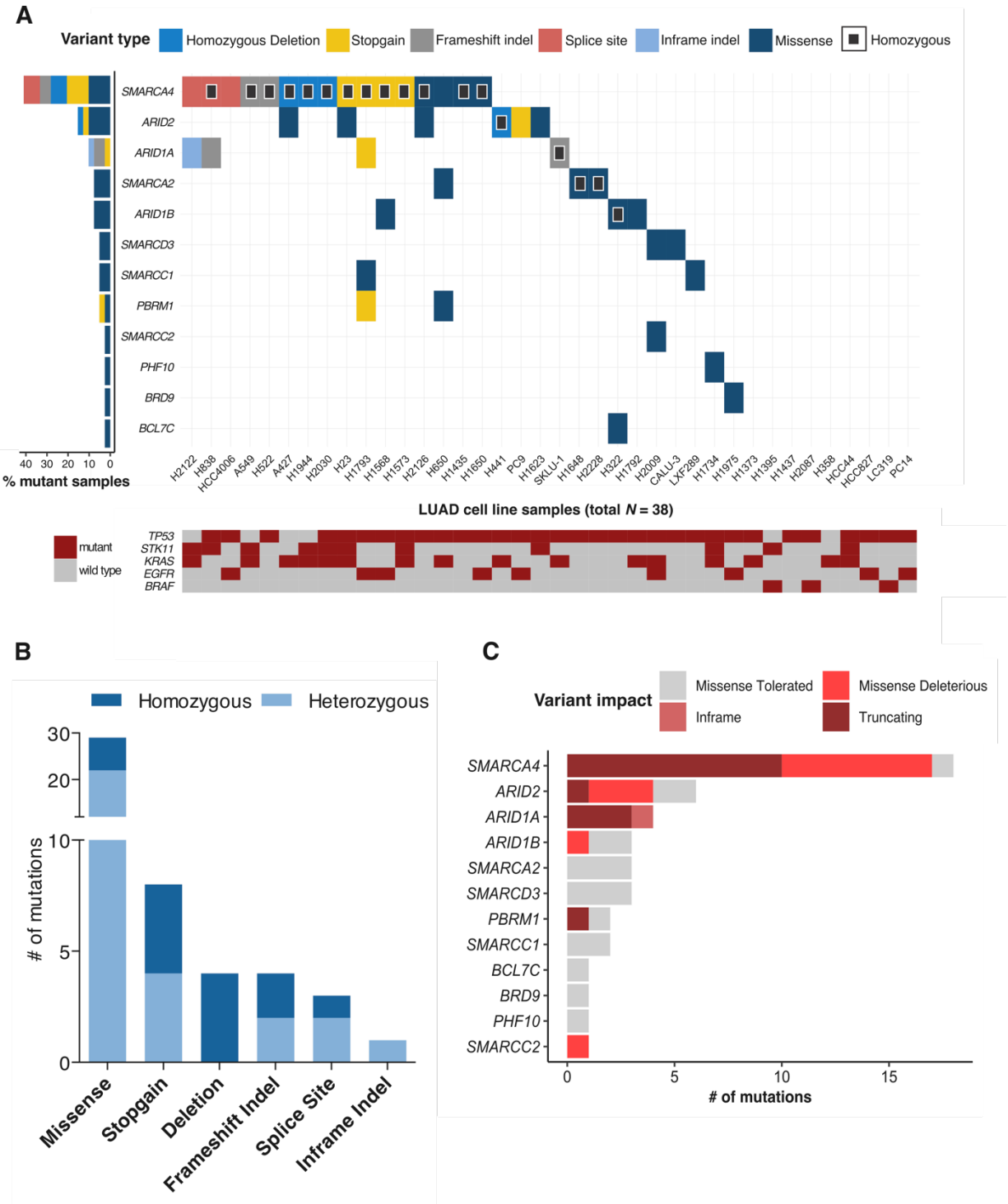


Figure 11: Mutational study of the SWI/SNF complex in LUAD cell lines. (**A**) Mutation profile of 20 SWI/SNF complex subunits in 38 LUAD cell lines. The Y-axis represents all the subunits that had at least one genetic alteration in one LUAD cell line. Homozygous mutations are depicted with a black square. The X-axis gathers all LUAD cell lines included in this study. On the left, mutation frequencies of these SWI/SNF subunits in LUAD cell lines. (**B**) Distribution of the different variant types found in SWI/SNF subunits in our 38 LUAD cell lines. Light blue shows those genetic alterations that were

heterozygous. Dark blue depicts homozygous mutations. (**C**) Functional prediction of the mutations found in SWI/SNF subunits in our panel of 38 LUAD cell lines using SIFT.

As we have previously mentioned, we also analyzed the mutational status of the top five LUAD driver genes in the same panel of cell lines (***Fig. 11A***). On the one hand, we observed that, among these driver genes, *BRAF* was the only one that did not harbor any mutations when there was a mutated SWI/SNF subunit. On the other hand, due to the high mutation rate of *TP53* in our LUAD cell lines (84.2%), most SWI/SNF mutant cell lines were also *TP53*-mutant. Only 25% of *SMARCA4* and *ARID1A* mutant samples and 16% of *ARID2* mutant samples had a wild-type *TP53*. To evaluate whether there was any statistically significant co-occurrence or mutual exclusion of mutations of the SWI/SNF complex and the top 5 LUAD driver genes, we analyzed a larger cohort from TCGA-LUAD (N = 574). We considered all possible pairs between SWI/SNF subunits and the five tested LUAD drivers (see ***Supplementary Fig. 6*** of ***Annex 9***). SWI/SNF mutations significantly overlapped with *TP53* mutations in TCGA data ($p = 0.0018$). However, this overlap could not be corroborated in our cell line data ($p = 0.61$). Nevertheless, the analysis of co-occurrence and mutual exclusion of mutations should be interpreted with caution because they are affected by multiple external variables, and some significant results might be statistical artifacts (van de Haar et al. 2019).

4.1.4. LUAD patients with mutations in the SWI/SNF complex have higher tumor mutation burden and worse prognosis

Another remarkable result that we observed in our mutational analysis of the TCGA-LUAD cohort was that SWI/SNF-mutant tumors showed a significantly higher Tumor Mutation Burden (TMB) than SWI/SNF-wild type tumors ($p < 0.05$) (***Fig. 12A***). This means that a mutated SWI/SNF complex could be related to an increase in genome instability that could potentiate tumor heterogeneity and hinder their treatment. For this reason, we decided to evaluate whether the mutational status of the lung SWI/SNF subunits was associated with LUAD overall survival (OS) in the TCGA-LUAD cohort. To select variables for a multivariate Cox analysis, we first performed univariate Cox analyses on each of the variables under study and we selected those with $p < 0.2$. In the univariate analysis, none of the individual

SWI/SNF subunits were significantly associated with OS, but SWI/SNF mutations altogether were significantly associated with poorer OS (HR = 1.42; 95% CI: 1.04 - 1.93; $p = 2.5 \cdot 10^{-2}$) (***Fig. 12B*** and ***C***). These observations led us to consider the SWI/SNF complex as a single functional unit. On the contrary, mutations in none of the top 10 LUAD driver genes from Bailey et al (Bailey et al. 2018) were significantly associated with OS (see ***Supplementary Fig. 7A-J*** of ***Annex 10***). Next, all variables with $p < 0.2$ in the univariate analysis were used for a stepwise multivariate analysis. According to this analysis, the SWI/SNF mutational status is an independent prognostic factor associated to shorter OS in LUAD patients (HR = 1.37; 95% CI: 1.01 - 1.88; $p = 4.39 \cdot 10^{-2}$) (***Fig. 12D***). Therefore, the lung SWI/SNF mutational status distinguishes between two clinically different subgroups.

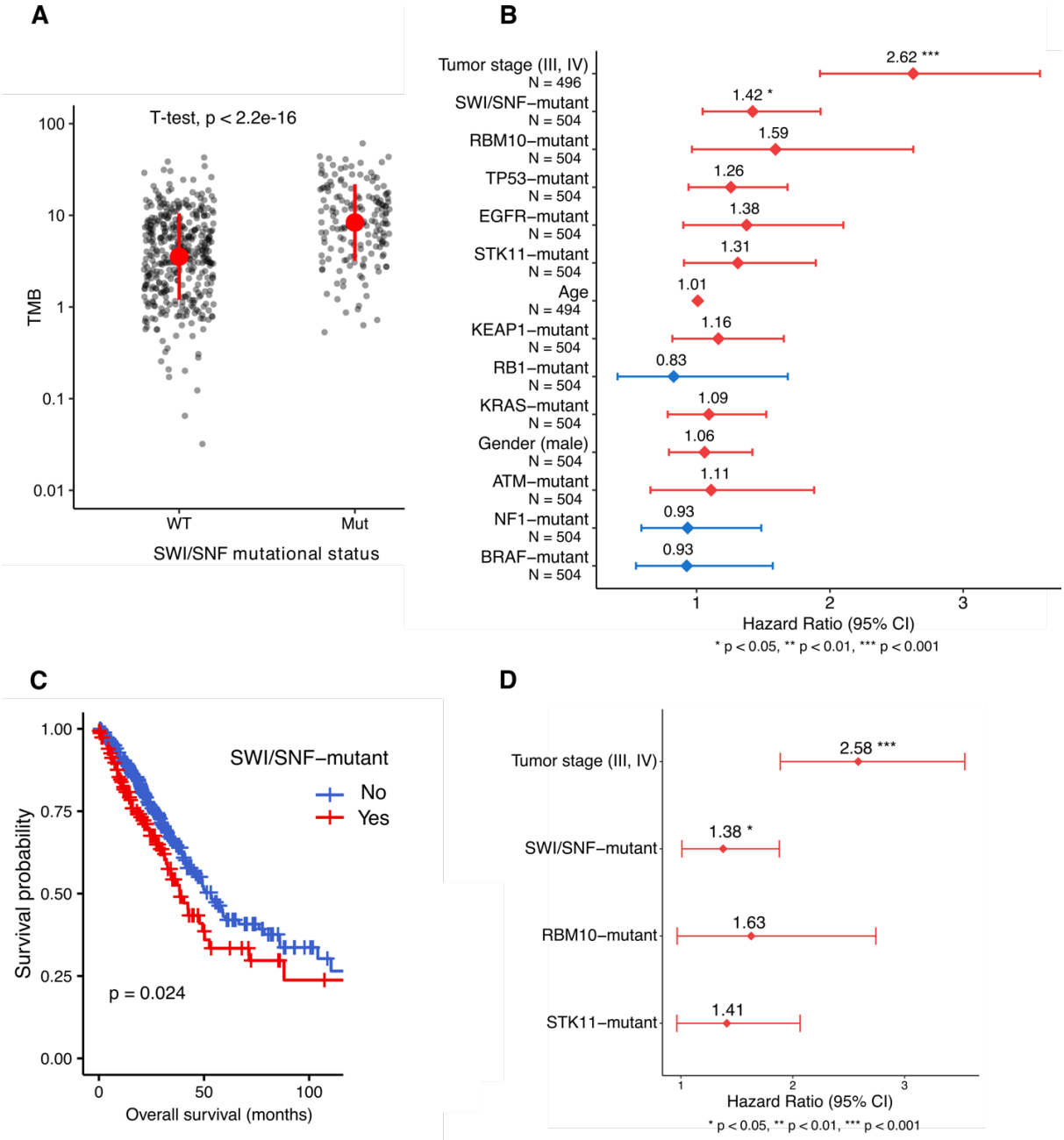


Figure 12: Clinical analyses with the mutational status of the lung SWI/SNF complex. (**A**) Tumor mutation burden (TMB), defined as the number of non-silent mutations per Mb as estimated by Hoadley et al (Hoadley et al. 2018), in SWI/SNF-wild type vs SWI/SNF-mutant patients from TCGA-LUAD. The red dot and the lines represent the mean and standard deviation of the $\log_{10}$(TMB) values, respectively. A two-tailed Student's t test was performed on the $\log_{10}$(TMB) values. (**B**) Univariate Cox Proportional-Hazards regression on mutation and clinical TCGA-LUAD covariates. All variables included in the model are sorted by statistical significance (*p*-value). (**C**) Kaplan-Meier curves grouping the TCGA-LUAD cohort by the mutational status of SWI/SNF complex (Logrank test). (**D**) Stepwise Multivariate Cox Proportional-Hazards regression on mutation and clinical TCGA-LUAD covariates. All variables included are sorted by statistical significance (*p*-value).

4.1.5. The SWI/SNF complex is frequently downregulated in LUAD

The high mutation rate is not the only cause that can affect the function of a protein. The SWI/SNF complex is not an exception. Other mechanisms, such as epigenetic inactivation through methylation or post-transcriptional regulation, can lead to an impairment of SWI/SNF expression (Marquez et al. 2015; Coira et al. 2015). For this reason, we went beyond the mutational study of the lung SWI/SNF subunits and we also analyzed their mRNA levels in our 70 LUAD primary tumors and their paired normal adjacent tissues using RT-qPCR for a better resolution. With this study, we found that all lung SWI/SNF subunits were significantly downregulated in LUAD primary tumors compared to their matched normal adjacent samples (FDR-adjusted $p < 0.05$, ***Fig. 13A***). Next, we decided to set a fold change threshold of +2/-2 between the tumor and the paired normal sample to consider a subunit to be up- or downregulated, respectively. We found 42 tumors (60%) that had more than 10 downregulated subunits (***Fig. 13B***). On average, each lung SWI/SNF subunit was downregulated in ~57% of LUAD patients. The top downregulated SWI/SNF subunit was *SMARCA2* (82% of the cases). Similar results have been observed in other tumors where *SMARCA2* was found to be epigenetically repressed (Glaros et al. 2007; Mizutani et al. 2002; Yamamichi et al. 2005). Moreover, none of the top 5 downregulated subunits (*SMARCA2*, *DPF2*, *SMARCD3*, *PHF10*, and *SMARCD1*) were among the top 5 most frequently mutated subunits. More generally, only 5/11 (45.5%) truncating mutations and 13/23 (56.5%) missense mutations were associated with more than a 2-fold decrease in expression. Overall, these findings suggest a profound silencing in the expression of the whole SWI/SNF machinery in

LUAD and show that genetic alterations are not the only cause of SWI/SNF inactivation.

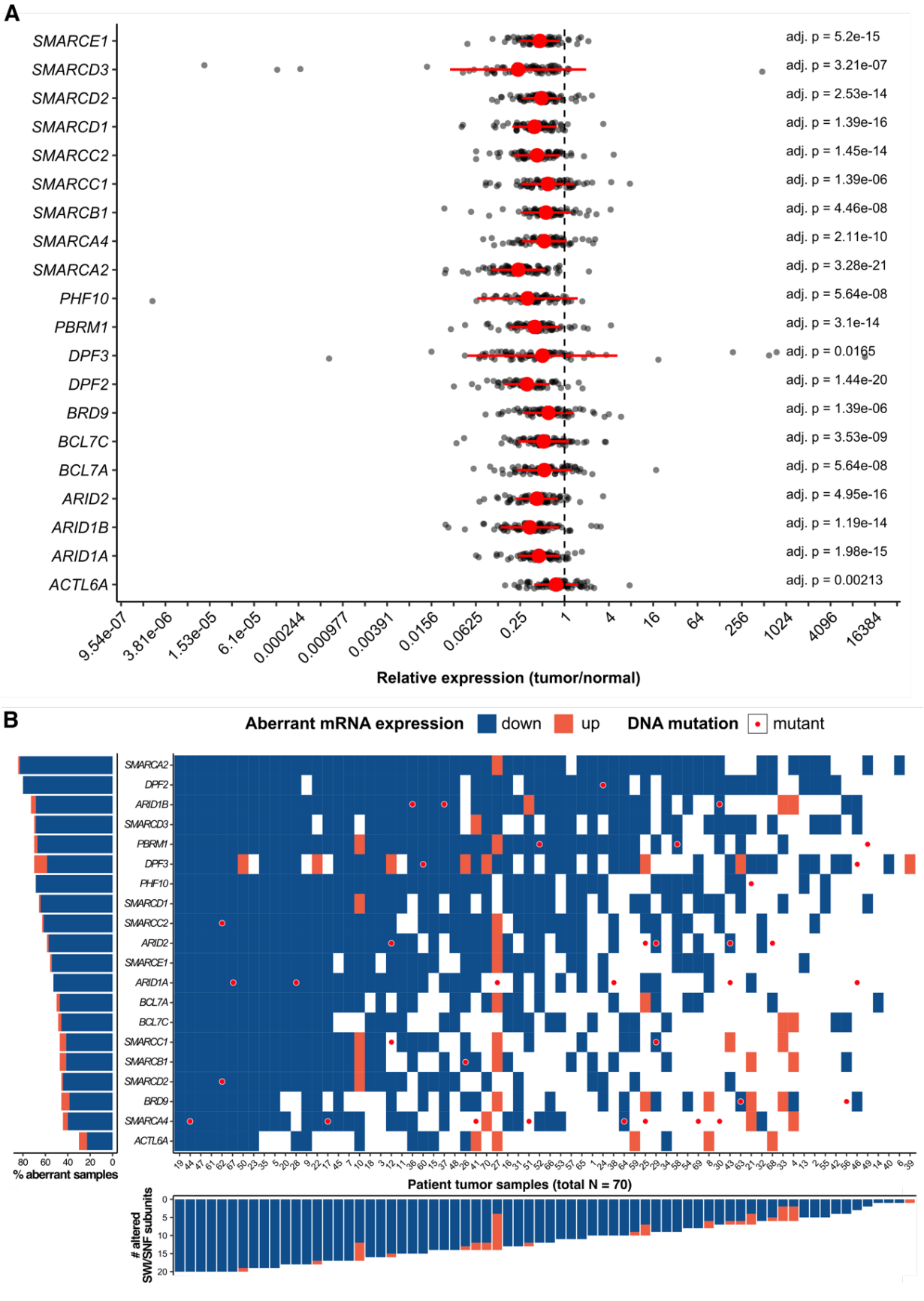


Figure 13: Downregulation of lung SWI/SNF complex subunits in our LUAD cohort. (**A**) The log_2-relative expression between each tumor and its matched normal sample was estimated as ΔCt(normal) - ΔCt(tumor). The red dots and lines represent the mean and standard deviation of the log_2-relative expression values. The FDR-corrected *p* values from one-sample t tests under the null hypothesis that the log_2-relative expression values are equal to 0 are shown. (**B**) Tile plot of the lung SWI/SNF subunits mRNA expression in our LUAD cohort. Blue colors correspond to those genes that showed ≤-2x expression in the tumor sample than in the matched normal sample. Orange colors are displayed when a gene was expressed ≥2x in the tumor. White colors correspond to those expression values that did not reach the thresholds that we defined for upregulation or downregulation. Red circles are present when a certain gene was mutated in a specific patient. On the left side, lung SWI/SNF genes are arranged based on downregulation percentage in our LUAD patients. At the bottom of the tile plot, our 70 LUAD patients are arranged based on the number of lung SWI/SNF subunits that were downregulated in their tumors.

Next, we analyzed the transcriptional levels of the same 20 SWI/SNF subunits in our 38 LUAD cell lines by RT-qPCR. We evaluated the intrinsic variability of mRNA expression among cell lines using the median ΔCt for each gene as the normalization value to calculate a −ΔΔCt (***Fig. 14***).

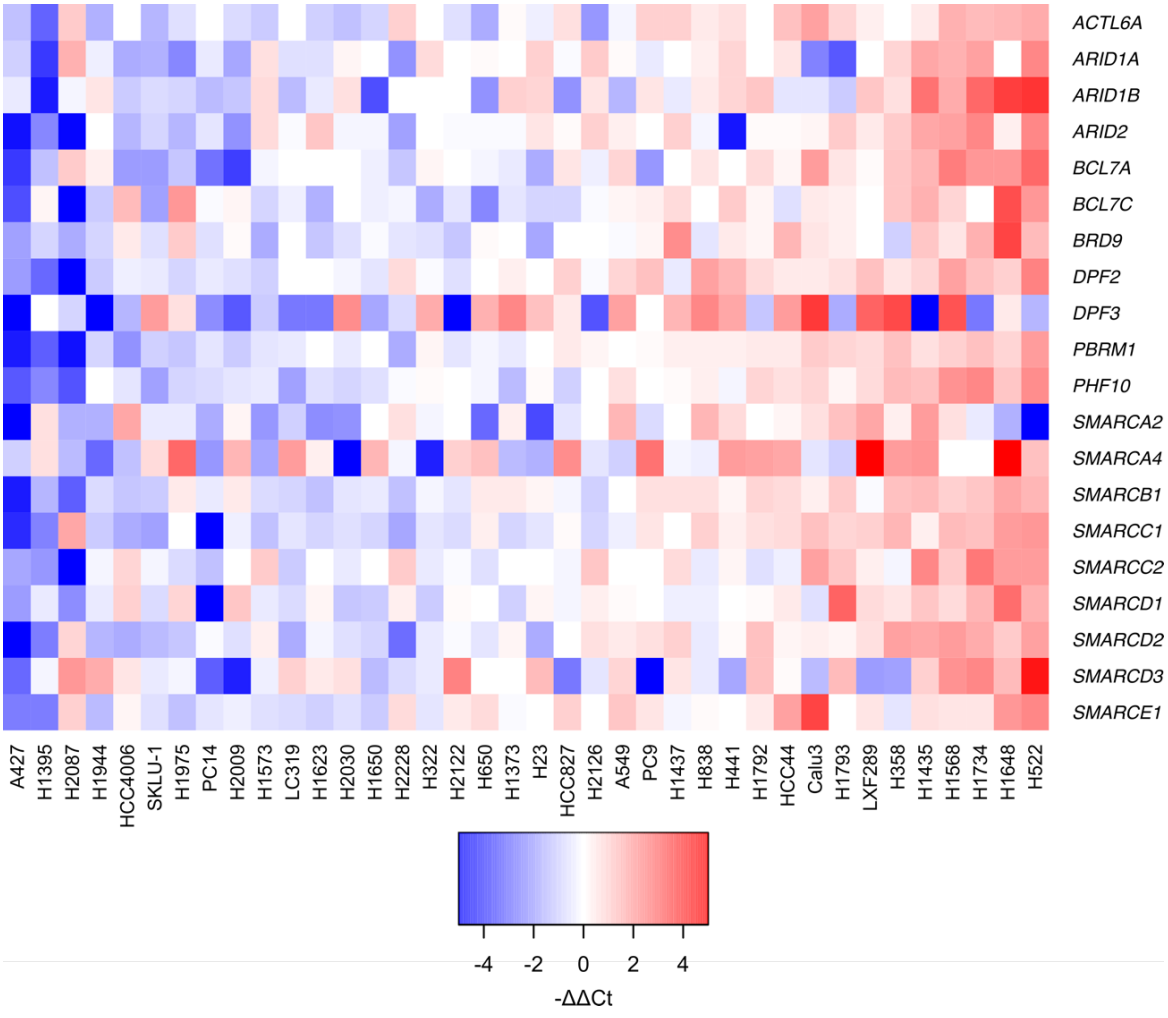


Figure 14: Transcriptional analysis of the SWI/SNF complex in LUAD cell lines. Heatmap of mRNA expression changes within the LUAD cell lines in 20 SWI/SNF subunits (−ΔΔCt was calculated using the median ΔCt for each of the measured genes). The Y-axis represents all measured SWI/SNF subunits. The X-axis contains the 38 LUAD cell lines of our study.

With this analysis, our panel of LUAD cell lines displayed two tendencies. On the one hand, there were cell lines with low relative expression of most of the SWI/SNF complex (e.g., A427 and H1395). On the other hand, our panel also contained cell lines with high relative mRNA levels of most SWI/SNF subunits (e.g., H522 and H1648).

Finally, to provide a complete resource of SWI/SNF alterations in our panel of 38 LUAD cell lines, we also combined data of DNA and mRNA alterations (***Fig. 15***). Intriguingly, some cell lines, such as H1395, had most SWI/SNF subunits with less expression than other LUAD cell lines, despite lacking DNA alterations.

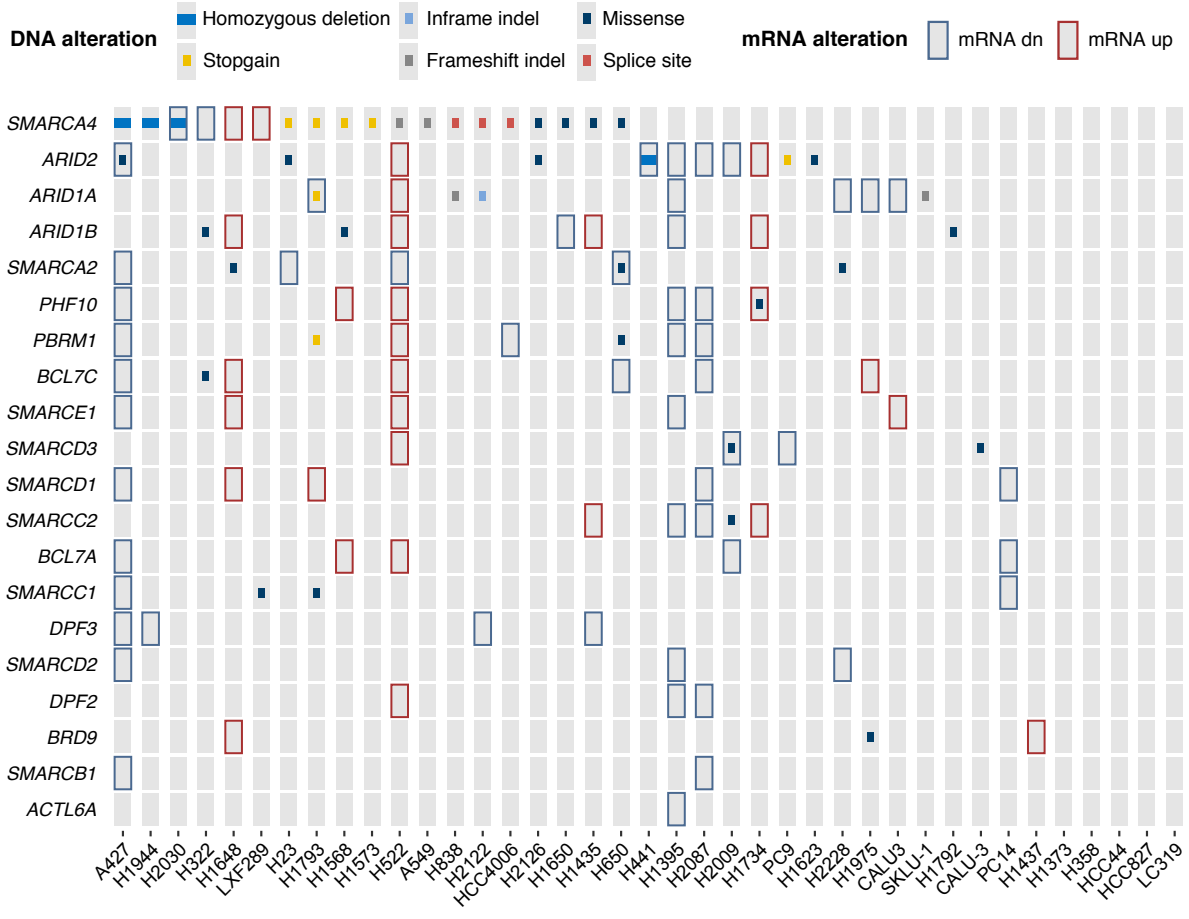


Figure 15: Summary of DNA and RNA alterations of the analyzed SWI/SNF subunits in our collection of LUAD cell lines. Copy number alterations, point mutations, short indels, and alterations of mRNA levels are represented. To find alterations in mRNA expression, robust Z scores were calculated by subtracting the median ΔCt for each gene across all cell lines from the ΔCt for each gene in each cell line, and then dividing by the median absolute deviation for each gene across all cell lines. The cutoffs for down- or up-regulation were robust Z < −2 or robust Z > 2, respectively.

4.1.6. Genetic and Epigenetic Factors Contribute to the Protein Loss of the ATPases and ARID Subunits

Since we developed an extensive profile of the SWI/SNF complex status in our panel of LUAD cell lines, we decided to include an analysis of the protein levels of the determinant SWI/SNF subunits that define the main SWI/SNF subtypes. Generally, the ATPases (SMARCA4 and SMARCA2) and ARID subunits (ARID1A, ARID1B, and ARID2) are used to classify SWI/SNF complexes as BAF, PBAF, or ncBAF (Mashtalir et al. 2018). For this reason, we performed western blots of these subunits to evaluate which SWI/SNF complexes could be found in the 38 LUAD cell lines and in NL20, our control non-tumor lung cell line (***Fig. 16A***). In contrast with NL20 (***Supplementary Fig. 8*** of ***Annex 11***), only 23.7% of the LUAD cell lines had detectable protein levels of the two ATPases and the three ARIDs. However, no LUAD cell line lacked all of the five analyzed proteins, supporting the idea that there may always be residual SWI/SNF complexes controlling gene expression (Helming et al. 2014a).

According to these western blot analyses, ARID1A was the most commonly lost subunit in LUAD cell lines, as we observed it in 19 out of 38 cell lines (50.0%). This was followed by the ATPase subunits SMARCA4 and SMARCA2, which were lost in 12 (31.6%) and nine (23.7%) LUAD cell lines, respectively. Among these three proteins, we observed different explanations for the protein loss (***Fig. 16B***). On the one hand, 84% of the ARID1A loss and 89% of the SMARCA2 loss could not be explained by any genetic alterations. On the other hand, 83% of SMARCA4 losses were directly related to truncating mutations. Remarkably, most SMARCA4 mutations in cell lines were homozygous (***Fig. 11A***), which could partially explain the strong correlation between its mutation and lack of protein. In general, these

results also support our previous observation that there is a combination of genetic, epigenetic, and post-translational regulations that influence the expression and, therefore, the functionality of the SWI/SNF complex.

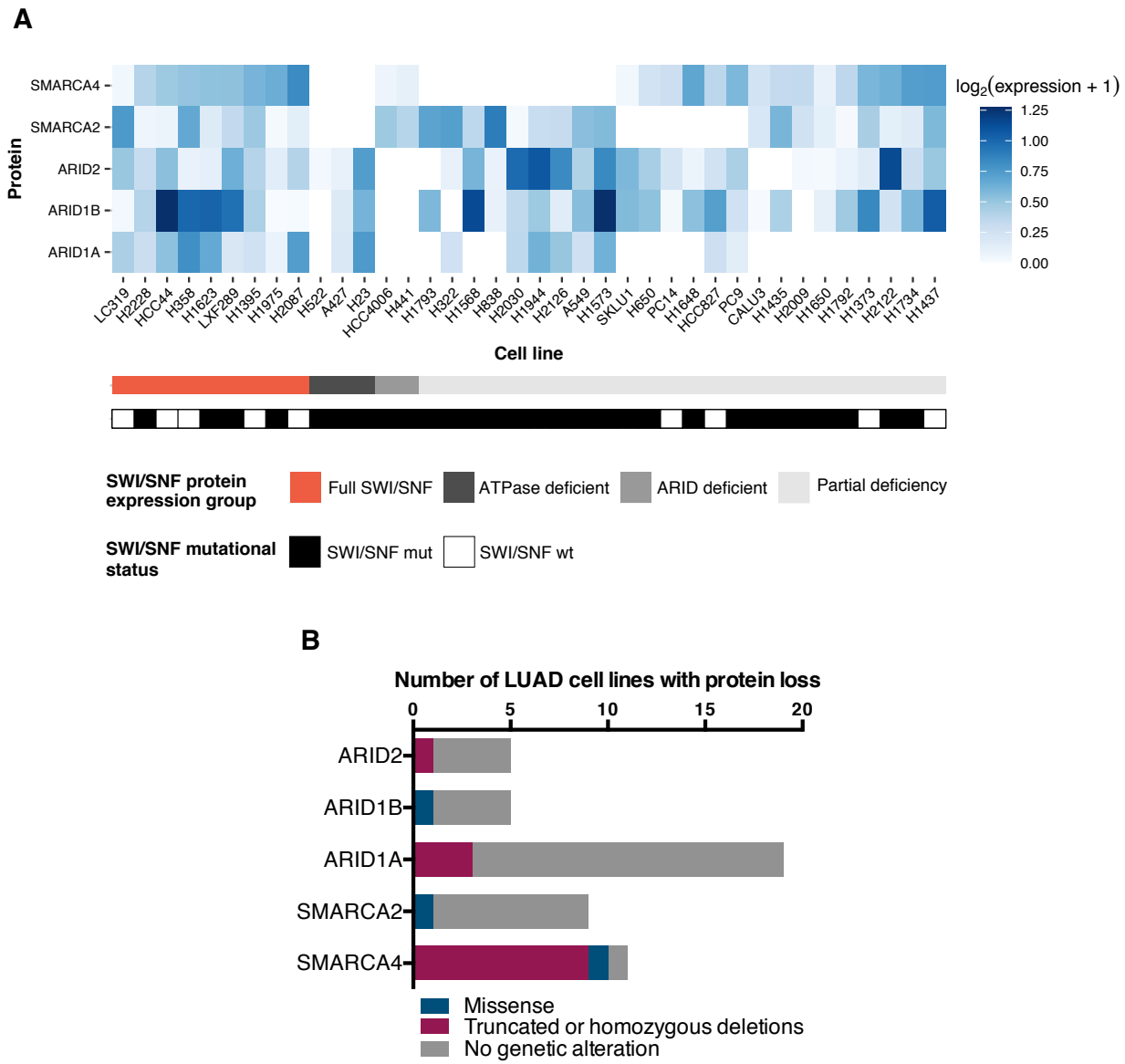


Figure 16: Protein expression profile of ATPases and ARID subunits of the SWI/SNF complex in LUAD cell lines. (**A**) Heatmap with normalized protein expression values of the ATPases and ARIDs subunits of the SWI/SNF in 38 LUAD cell lines. Zero values correspond to the absence of a band in the Western blot (see ***Supplementary Fig. 8*** of ***Annex 11***). Below the heatmap, two lines show a classification of the panel of LUAD cell lines. The first line depicts the classification of LUAD cell lines based upon protein expression of the ATPases and ARIDs subunits. The second line shows the mutational status of the SWI/SNF complex considering all 20 subunits analyzed in this study (**B**) Causative analysis of the protein loss of the ATPases and ARID subunits in LUAD cell lines.

4.1.7. ATPases and ARID Protein Expression Profiles Define Four LUAD Cell Line Subgroups

The protein expression profiles of the five analyzed SWI/SNF subunits also allowed us to distinguish four subgroups of LUAD cell lines (***Fig. 16A***). The first subgroup gathered all cell lines that showed detectable levels of all SWI/SNF proteins and it comprised nine cell lines (23.7%). Inside this subgroup, we could find five cell lines (LC319, HCC44, H358, H1395, and H2087) that were wild-type for all the 20 SWI/SNF subunits analyzed in this study. These cell lines could be a good reference of wild-type lung SWI/SNF contexts. Second, there was an ATPase deficient subgroup where we found three cell lines (7.9%). This subgroup has previously been observed in other studies (Reisman et al. 2003; Fukuoka et al. 2004; Marquez-Vilendrer et al. 2016) and these data restrict the widespread idea of SMARCA2/4 synthetic lethality proposed by Oike and colleagues to only certain genetic contexts (Oike et al. 2013). Third, we defined a subgroup of ARID-deficient cell lines comprising H441 and HCC4006. This observation suggests that ARID1A/B synthetic lethality (Helming et al. 2014b) is also limited to specific contexts. Moreover, this ARID-deficient subgroup could support the existence of the recently described ncBAF that lacks any ARID subunit (Alpsoy and Dykhuizen 2018; Mashtalir et al. 2018). To corroborate this observation, we analyzed the protein expression of BRD9, a specific ncBAF subunit (***Fig. 17***). Both ARID-deficient cell lines expressed BRD9, although their protein levels were lower than that observed in HCC44, one of the cell lines of the Full-SWI/SNF subgroup. Finally, we defined a fourth subgroup of LUAD cell lines bearing partial SWI/SNF loss in various combinations that reflect the diversity of assembly even within the same histological type of tumor. Overall, our observations emphasize the heterogeneity of the SWI/SNF status in LUAD cell lines, which must be considered when using these cell lines as *in vitro* models for studying the functionality of the SWI/SNF complex.

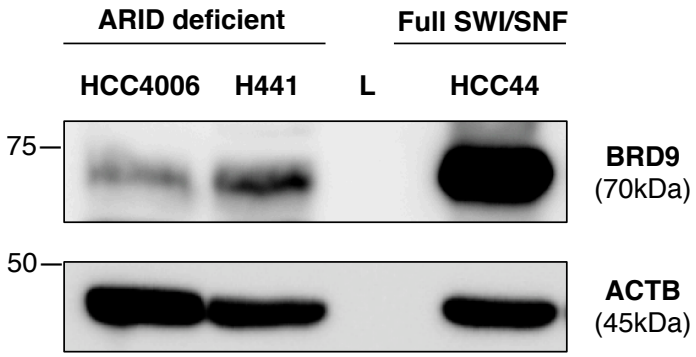


Figure 17: Western blot of BRD9 in our ARID deficient LUAD cell lines and a FULL-SWI/SNF LUAD cell line. The sizes depicted in the image correspond to the observed molecular weight of BRD9.

4.2. Chapter II: Regulation of microRNA expression by SMARCA4, the catalytic subunit of the SWI/SNF complex

4.2.1. SMARCA4 restoration induces expression changes in cancer-related miRNAs

To analyze the specific involvement of SMARCA4 in regulating miRNA expression in lung adenocarcinoma, we chose an *in vitro* model that was SMARCA4-deficient so that we could track the changes derived from its re-expression. We used the A549 cell line because it has a homozygous frameshift deletion in *SMARCA4* that generates a premature stop codon (Medina et al. 2008a). In addition, our previous study of the protein expression of some SWI/SNF subunits in our panel of LUAD cell lines showed that A549 had no residual SMARCA4. Therefore, we concluded that, in this cell line, the ATPase activity of the SWI/SNF complex relied on SMARCA2.

We restored SMARCA4 in this cellular model by transiently transfecting a DNA construct containing the most abundant isoform of this gene in lung tissue (see Materials and Methods). We corroborated that SMARCA4 was re-expressed, reaching its peak expression at 24 h (***Fig. 18A***) and that it was located in the chromatin fraction of the transfected cells (***Fig. 18B***).

To determine whether the exogenous SMARCA4 was incorporated in the endogenous SWI/SNF complexes of the A549 cell line, we performed two complementary IP analyses (***Fig. 18C***). When we pulled down SMARCC1, one of the core SWI/SNF subunits (Mashtalir et al. 2018), we observed that our exogenous SMARCA4 was bound to it. The same interaction was obtained when we performed the opposite pull-down. This result showed that the exogenous SMARCA4 was successfully incorporated into endogenous SWI/SNF complexes, yielding complexes where SMARCA4 was the catalytic subunit. However, we also confirmed that SMARCA2-containing SWI/SNF complexes coexisted with the new SMARCA4-SWI/SNF complexes in the transfected A549 cells (***Fig. 18D***).

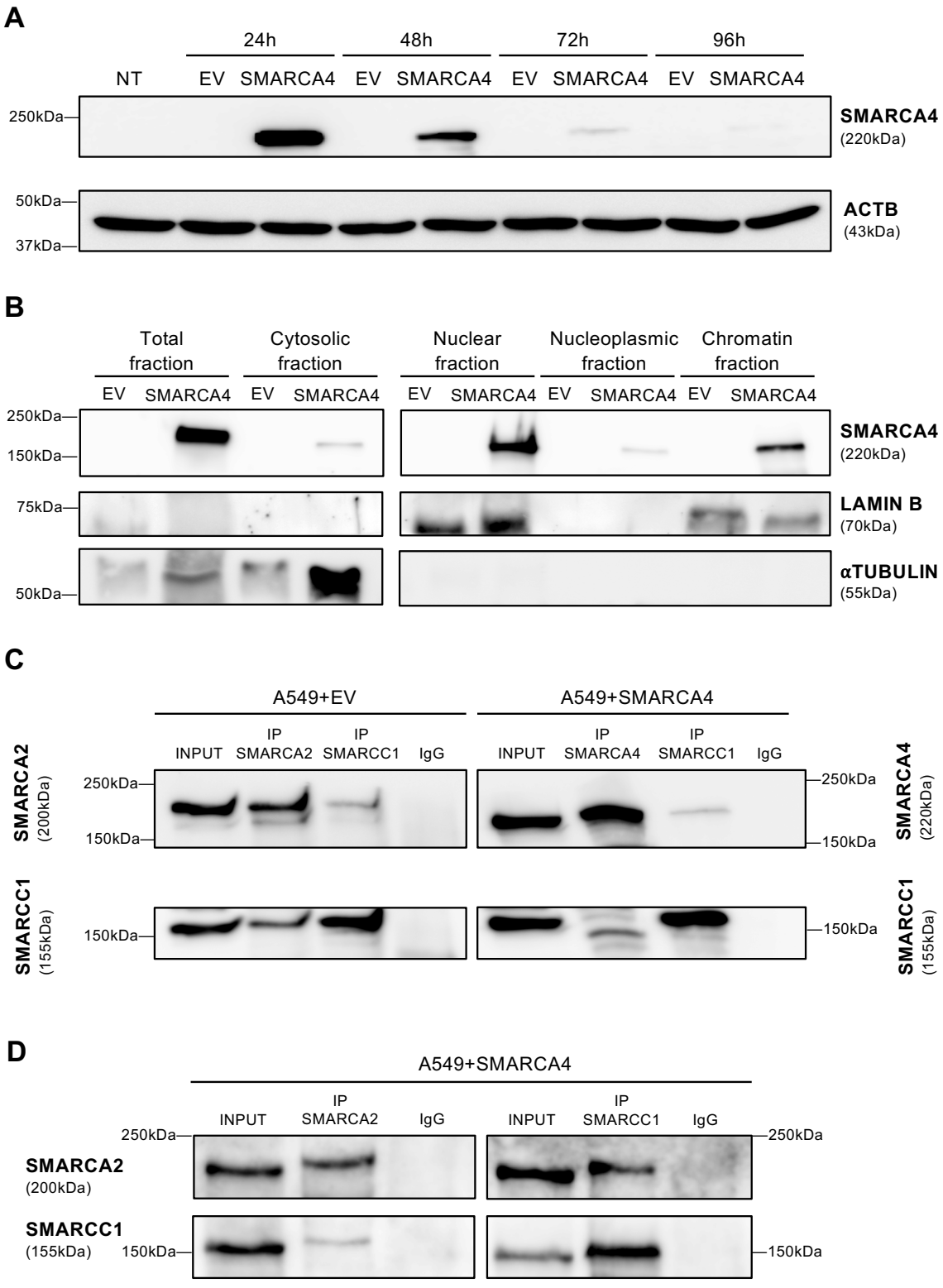


Figure 18: SMARCA4 re-expression and localization. (**A**) Western blot analysis of SMARCA4 protein levels after SMARCA4 restoration in A549 cell line at different time points. ACTB was used as a loading control. (**B**) Subcellular fractionation analysis to determine the localization of the exogenous SMARCA4 in transfected A549. Lamin B was used as nuclear and chromatin control. α-Tubulin was used as a cytoplasmic control. (**C**) Western blot of the IP of SMARCA2 and SMARCC1 in

A549 transfected with empty vector (EV) and the IP of SMARCA4 and SMARCC1 in A549 transfected with the SMARCA4-plasmid. IgG was used as a negative control for nonspecific binding. 5% of INPUT was included for the analysis. (**D**) Western blot of the IP of SMARCA2 and SMARCC1 in A549 transfected with the SMARCA4-plasmid. IgG was used as a negative control for nonspecific binding. 5% of INPUT was included for the analysis.

The tumor suppressor role of SMARCA4 was validated by measuring cell viability and colony formation after its restoration. Ectopic expression of SMARCA4 impaired cell viability up to 48% after four days of transfection (***Fig. 19A***) and reduced the clonogenic capacity by 49% (***Fig. 19B***). Both results agreed with the tumor suppressor function of this protein and showed that transient restoration of SMARCA4 decreased the viability of LUAD cells.

Following the demonstration that SMARCA4 re-expression showed tumor suppressor activity, we determined whether changes in the miRNome could contribute to its tumor suppressor role. For this purpose, we performed miRNA sequencing after 48 h of SMARCA4 re-expression. We found 57 miRNAs that were significantly dysregulated upon SMARCA4 restoration (FDR < 0.05, absolute $\log_2$ Fold Change (FC) > 1). Specifically, 29 miRNAs showed increased levels and 28 miRNAs were downregulated in A549 transfected with the SMARCA4-plasmid compared to A549 transfected with an empty vector (***Fig. 19C***, see **Annex 12**).

We used the online tool DIANA-miRPath (see Materials and Methods) to analyze the pathways that could be altered by the differentially expressed miRNAs. We observed that for both upregulated and downregulated miRNAs, the most significant KEGG pathway was 'MicroRNAs and cancer' followed by 'Transcriptional misregulation in cancer' in the case of the downregulated miRNAs (***Fig. 19D***). These findings suggested that re-expression of SMARCA4 not only regulates protein-coding genes, as it has been previously described (Medina et al. 2005; Romero et al. 2012; Orvis et al. 2014; Lazar et al. 2020), but also changes the expression of key miRNAs involved in tumorigenic pathways.

4.2.2. SMARCA4 binds to an enhancer of miR-222 and controls its expression

Among the cancer-related miRNAs that were differentially expressed upon SMARCA4 restoration, we selected the most significant dysregulated miRNA, miR-222, to perform a detailed study. First, we confirmed by RT-qPCR that miR-222 levels increased 24 h after SMARCA4 restoration, and at 48 h they decreased along with SMARCA4 levels (***Fig. 19E***).

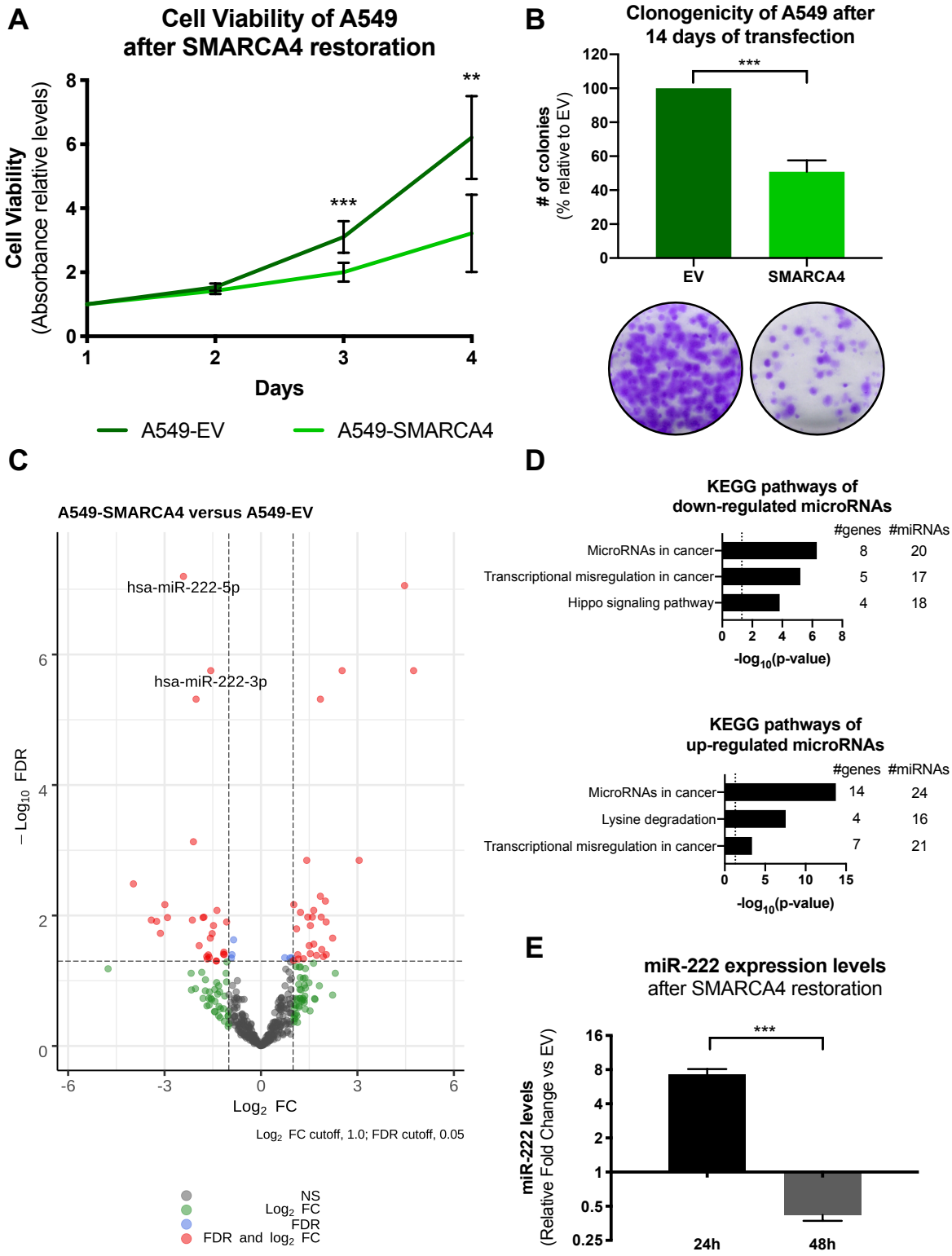


Figure 19: SMARCA4 restoration in A549 cells reduces cell viability and changes the miRNome. (**A**) Resazurin assay at different time points after transfection of A549 with empty vector (A549-EV) or the SMARCA4-plasmid (A549-SMARCA4). (**B**) Clonogenic assay after 14 days of SMARCA4 restoration in A549. Upper panel: Colony number quantification. Lower panel: Representative images of each condition after the staining with crystal violet. (**C**) Volcano plot with the differentially expressed miRNAs upon SMARCA4 restoration. miRNAs with an FDR < 0.05 and an absolute $\log_2$ FC > 1 are displayed in red color. (NS: Non-significant; FDR: False Discovery Rate; FC: Fold Change) (**D**) Top 3 significantly enriched pathways of the dysregulated miRNAs upon SMARCA4 restoration. For each pathway, we represented the number of genes that are targets of at least seven miRNAs of our dysregulated miRNAs. (**E**) miR-222 relative expression levels after 24 h and 48 h of restoration of SMARCA4 in A549. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p*-value < 0.05; ** $p < 0.01$; *** $p < 0.001$.

The decrease in miR-222 expression after 48 h of SMARCA4 restoration agreed with the miRNA-Seq data, which showed that it was the top down-regulated miRNA at that time point (***Fig. 19C***).

We noticed that the *MIR222* locus is contained within an enhancer according to the GeneHancer database (GeneHancer ID: GH0XJ045746) (Fishilevich et al. 2017). Interestingly, enhancers are genomic regions frequently targeted by the SWI/SNF complex (Lazar et al. 2020; Shi et al. 2013; Hodges et al. 2018; Wang et al. 2017; Alver et al. 2017; Mathur et al. 2017; Bossen et al. 2015; Vierbuchen et al. 2017; Hu et al. 2011). Therefore, we hypothesized that miR-222 may be regulated by the enhancer under the control of the SWI/SNF complex. Using public ChIP-Seq data from A549 (Davis et al. 2018), we confirmed that the region upstream of miR-222 contains peaks of H3K4me1 and H3K27ac marks, which are characteristic histone marks of enhancers (***Fig. 20A***).

To determine whether SMARCA4- or SMARCA2-containing SWI/SNF complexes bind to the miR-222 enhancer, we performed ChIP-qPCR analyses. After 24 h of restoration of SMARCA4, we obtained a significant enrichment only of SMARCA4 signal at the miR-222 enhancer (***Fig. 20B***). However, after 48 h, as SMARCA4 transient expression declined (***Fig. 20C***), it was replaced by SMARCA2. This turnaround also matched the expression changes that we found in miR-222 (***Fig. 19E***). These data showed that the expression of miR-222 was directly influenced by SMARCA4, which reached its peak expression 24 h after transfection (***Fig. 20C***).

These results suggest a model where miR-222 expression is activated preferentially by SMARCA4-SWI/SNF complexes and not by SMARCA2-SWI/SNF complexes (***Fig. 21***).

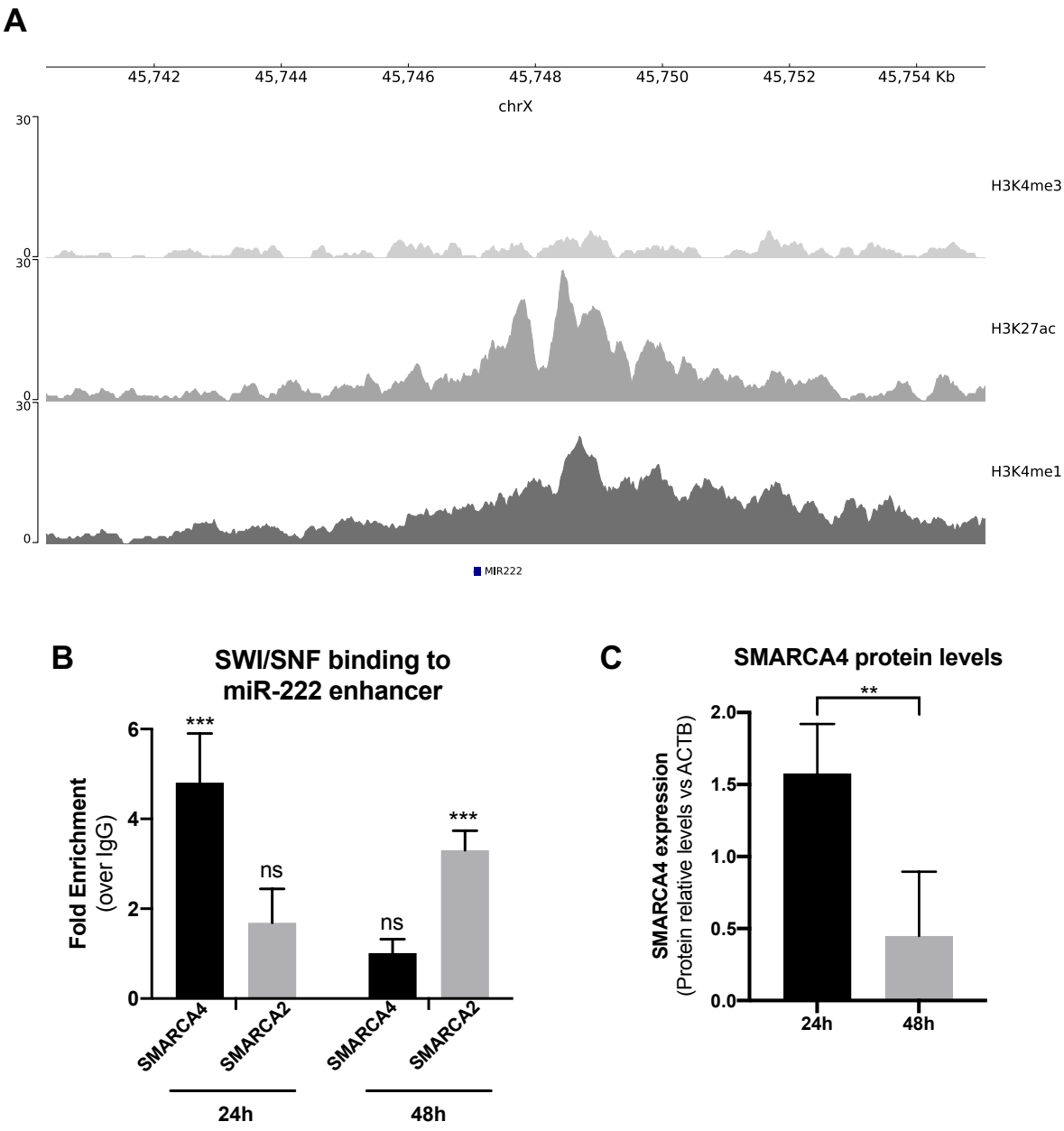


Figure 20: SMARCA4 binds to the enhancer of miR-222 in a time-dependent manner. (**A**) Overview of the miR-222 genomic context including ChIP-seq profiles of H3K4me3, H3K27ac, and H3K4me1 in the A549 cell line. High levels of H3K27ac and H3K4me1 and low levels of H3K4me3 are characteristic of enhancer regions. (**B**) Analysis by ChIP-qPCR of SMARCA4 or SMARCA2 binding to the enhancer of miR-222 after 24 h and 48 h of SMARCA4 restoration in A549. (**C**) Quantification of SMARCA4 protein levels relative to ACTB expression in transfected A549 cells after 24 h and 48 h of SMARCA4 restoration. Values represent mean ± SD ($n \geq 3$). *Two-tailed t-test p-value < 0.05; **p < 0.01.

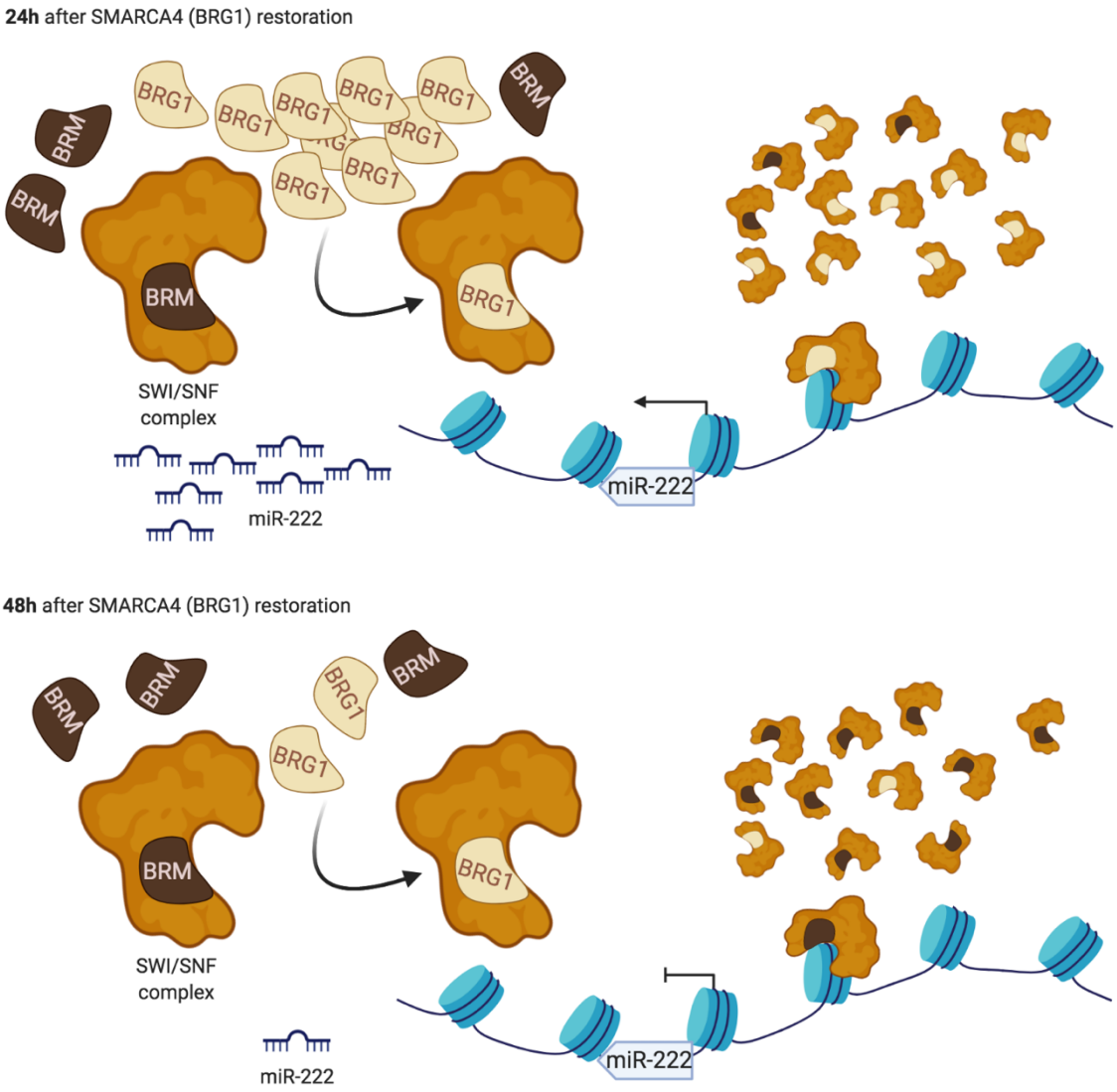


Figure 21: Schematic overview of the regulation of miR-222 by the SWI/SNF complex. Upper panel: After 24 h of SMARCA4 (BRG1) re-expression in A549 cells, there are high levels of SMARCA4 and it is included in the SWI/SNF complex. Thus, there is a predominance of SWI/SNF complexes with SMARCA4 and this type of complexes induces miR-222 expression. Lower panel: After longer times of transfection, SMARCA4 levels in A549 decrease and the SWI/SNF complexes with SMARCA2 (BRM) are predominant. This complex impairs miR-222 expression. Both figures were made using BioRender (https://app.biorender.com/).

4.2.3. miR-222 impairs cell viability phenocopying SMARCA4 restoration in the A549 cell line

Since the levels of miR-222 strongly increased upon SMARCA4 restoration, we aimed to elucidate the phenotypic effects of this miRNA in the A549 cell line. We transfected A549 cells with either a miR-222 mimic or a negative control mimic.

We found that increasing the levels of miR-222 (***Fig. 22A***) decreased cell viability by more than half (50.2%) after seven days of transfection (***Fig. 22B***). Moreover, miR-222 overexpression also impaired cell clonogenicity by 44%, pointing to a tumor suppressor role in this LUAD cell line (***Fig. 22C***). Interestingly, we observed that increasing miR-222 levels resulted in the same phenotype as restoring SMARCA4 in A549. These data highlight that SMARCA4-mutant contexts not only change crucial protein-coding genes needed by the cancer cell but also modify the expression of miRNAs that have relevant functions in tumor progression.

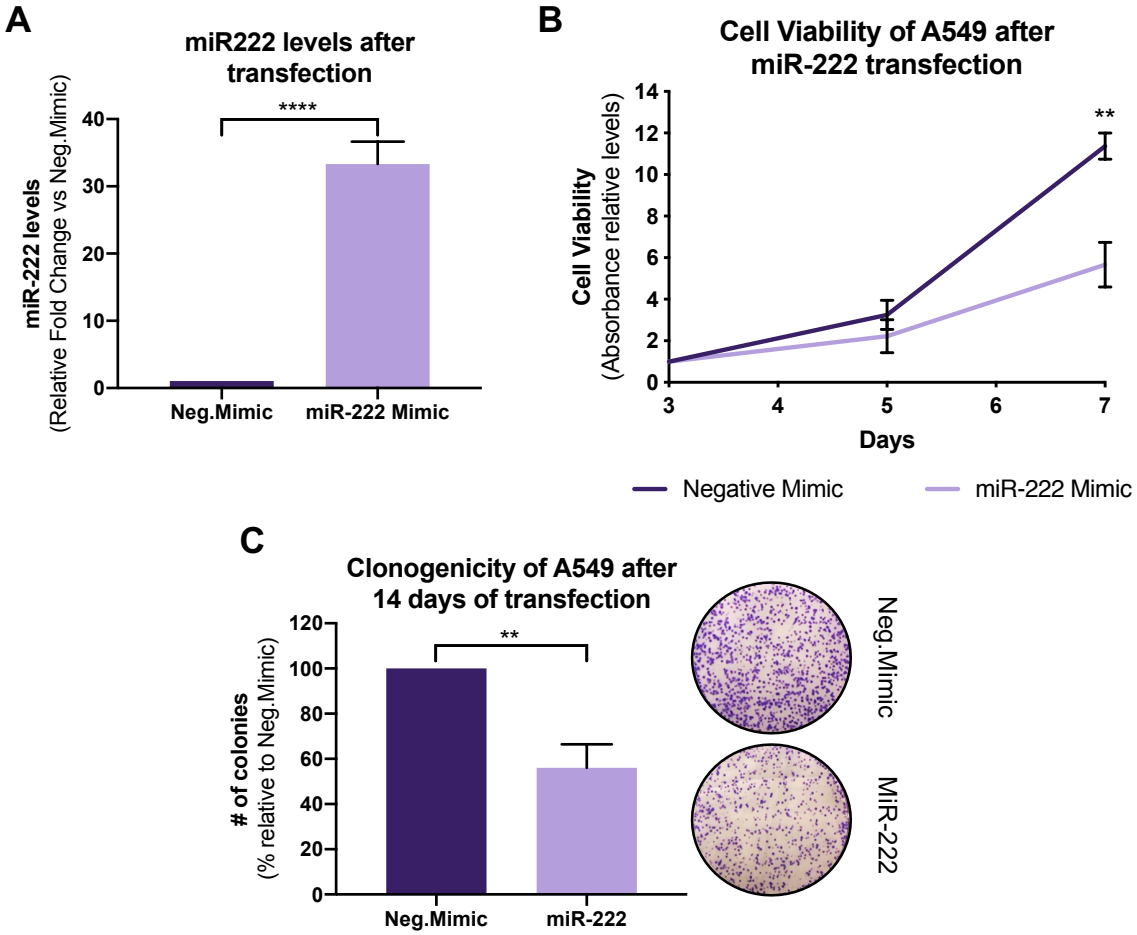


Figure 22: miR-222 behaves as a tumor suppressor miRNA by decreasing cell viability and colony formation. (**A**) Relative miR-222 expression levels after transfection of A549 cells with either a negative control mimic or a miR-222 mimic. (**B**) Resazurin assay at different time points after transfection of A549 with negative control mimic or miR-222 mimic. (**C**) Clonogenic assay after 14 days of transfection of A549 with negative control mimic or miR-222 mimic. Left: Colony number quantification. Right: Representative images of each condition after the staining with crystal violet. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p-value < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

4.2.4. The miR-222 enhancer belongs to a topologically associating domain that does not contain cancer-related protein-coding genes

We observed that SMARCA4-SWI/SNF complex binds to the miR-222 enhancer, increasing the expression of miR-222 and that miR-222 overexpression phenocopies SMARCA4 restoration in A549. However, we wondered whether the miR-222 enhancer modulates cancer-related protein-coding genes that might contribute to the phenotype. To find other potential protein-coding targets of the miR-222 enhancer, we studied the three-dimensional organization and interactions of the enhancer region in lung cells. We queried high-throughput chromosome conformation capture (Hi-C) and chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) data from ENCODE (https://www.encodeproject.org, v111). We found Hi-C and CTCF ChIA-PET data for A549 and for the normal lung fibroblast cell line IMR-90 (see Materials and Methods). In both cell lines, the miR-222 enhancer belonged to a topologically associating domain (TAD) that was delimited by CTCF (CCCTC-Binding Factor) binding sites (see ***Annex 13***). The TADs were more well-defined in IMR-90 than in A549 due to a better resolution (5 kb for IMR-90, 40 kb for A549). In the IMR-90 cell line, the smaller TAD that contained the miR-222 enhancer did not encompass any protein-coding genes. In the A549 cell line, the TAD was larger due to a poorer resolution of the data, and it only contained the protein-coding gene *KRBOX4.* However, to our knowledge, *KRBOX4* has not been linked to cancer. The nearest cancer genes to the miR-222 enhancer were *KDM6A* and *RBM10*, but they belonged to different TADs. This observation suggests that the miR-222 enhancer does not physically interact with these genes despite their proximity in the genomic sequence. Overall, we found no evidence that the miR-222 enhancer modulates any well-known cancer-related protein-coding genes, supporting that its tumor suppressor function may rely on miR-222.

4.3. Chapter III: Functional study of the SWI/SNF subunit ARID1A in LUAD

4.3.1. LUAD cell lines rely on ARID1A for their survival

The last objective of this Ph.D. thesis was to study the biological role of ARID1A in lung adenocarcinoma. First, we decided to analyze ARID1A expression in larger cohorts of LUAD patients such as the TCGA and the study group of Gillette and colleagues (Gillette et al. 2020) with 483 and 110 LUAD samples, respectively. Contrary to the expected for a tumor suppressor gene with high frequency of loss of function (LOF) alterations, which corresponded to 60% of *ARID1A* mutations in LUAD tumors from TCGA, we observed that there was not a significant decrease in *ARID1A* mRNA levels in LUAD tumors (***Fig. 23A***). On top of that, ARID1A protein levels were significantly higher in tumors versus normal adjacent tissue of LUAD patients (***Fig. 23B***).

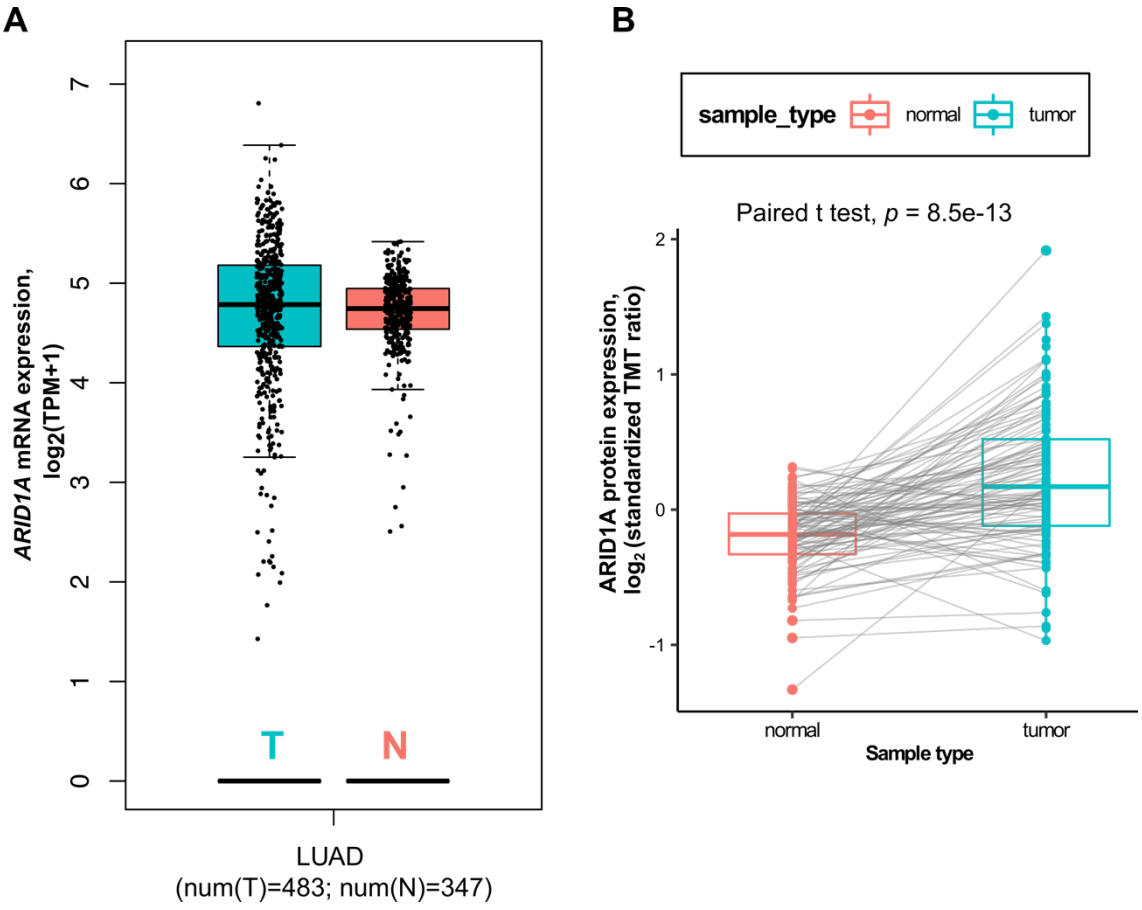


Figure 23: ARID1A expression levels in LUAD patients. (**A**) Box plot of *ARID1A* mRNA levels in tumor samples (T) and normal adjacent tissue (N) of LUAD patients. The tumor values of this analysis

correspond to the TCGA-LUAD cohort and the normal values were obtained by a combination of the normal samples from TCGA-normal and GTEx (Genotype-Tissue Expression) databases. The plot was obtained from GEPIA2 (Gene Expression Profiling Interactive Analysis) (Tang et al. 2019). Y-axis represents the logarithmic scale of the number of Transcripts Per Million (TPM) (**B**) Box plot of ARID1A protein levels in LUAD tumors and normal adjacent samples of the study of Gillette et al (Gillette et al. 2020). Y-axis depicts the Tandem Mass Tag (TMT) ratios provided by Gillette et al. Paired t-tests were performed with these data.

To determine the relevance of ARID1A in LUAD progression, we studied the effect of silencing *ARID1A* expression in a panel of LUAD cell lines with different genetic backgrounds to interrogate various cellular contexts (***Fig. 24*** and ***Fig. 25***). For that purpose, we used two different small interference RNAs (siRNAs) that effectively decreased *ARID1A* mRNA levels, although the silencing activity of siARID1A#1 was higher than siARID1A#2 (***Fig. 24A, D***).

First, we knocked down *ARID1A* in A549 and H2009, two LUAD cell lines that have mutations in other SWI/SNF subunits such as *SMARCA4* in the case of A549 and *SMARCD3* and *SMARCC2* in H2009. In both cell lines, we found that *ARID1A* knockdown (KD) significantly impaired cell viability and colony formation (***Fig. 24B, C, E, F***). Given that KRAS was a relevant LUAD driver gene that was also mutated in these two cell lines (see ***Annex 2***), we decided to study the effect of *ARID1A* silencing in H358, another LUAD cell line mutant for KRAS but wild-type for the SWI/SNF complex. We observed a similar effect on cell viability and colony formation after *ARID1A* silencing (***Fig. 25A-C***). Interestingly, similar results were also found when using H1395, a LUAD cell line wild type for both the SWI/SNF complex and *KRAS* (***Fig. 25D-F***). Overall, we observed that, regardless of the genetic background, these LUAD cell lines required ARID1A for their survival.

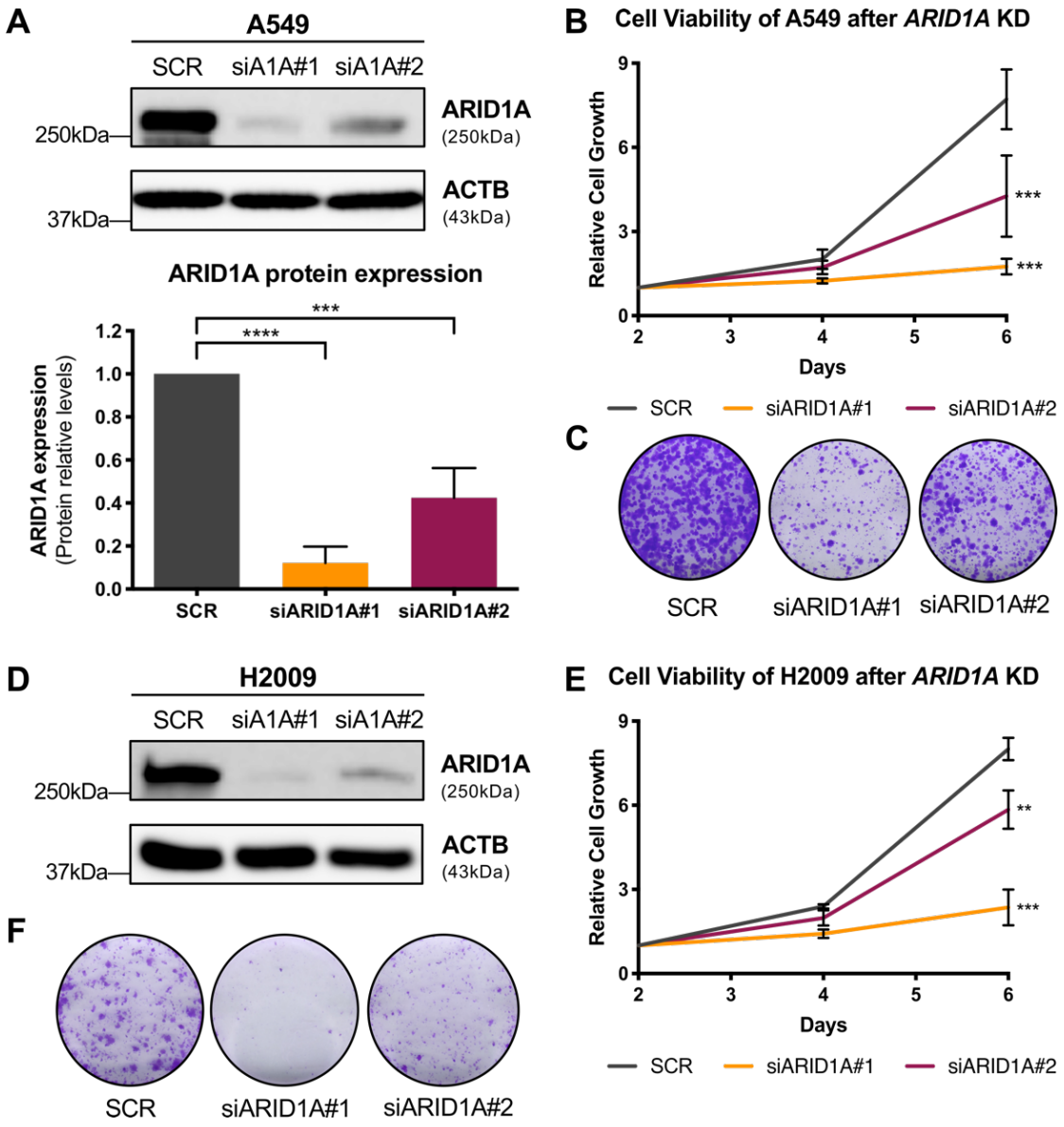


Figure 24: ARID1A dependency of LUAD cell lines with mutant SWI/SNF complexes. (**A**) Western Blot analysis of ARID1A protein levels in the A549 cell line after transfection with either nonsense scrambled siRNAs (SCR) or siRNAs against *ARID1A* mRNA (siA1A#1 or siA1A#2). The graph below represents the band quantification analysis performed with ImageJ using ACTB as the normalization protein. (**B**) Resazurin assays to measure cell viability at different time points after transfection of A549 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. (**C**) Clonogenic assay after 14 days of transfection of A549 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. (**D**) Western Blot analysis of ARID1A protein levels in the H2009 cell line after transfection with either scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. ACTB was used as a loading control. (**E**) Resazurin assays to measure cell viability at different time points after transfection of H2009 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. (**F**) Clonogenic assay after 14 days of transfection of H2009 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. Values represent mean ± SD (n ≥ 3). Two-tailed Mann–Whitney U-test was chosen for

A549 cell viability analysis and two-tailed t-test for the rest of experiments. *p-value < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

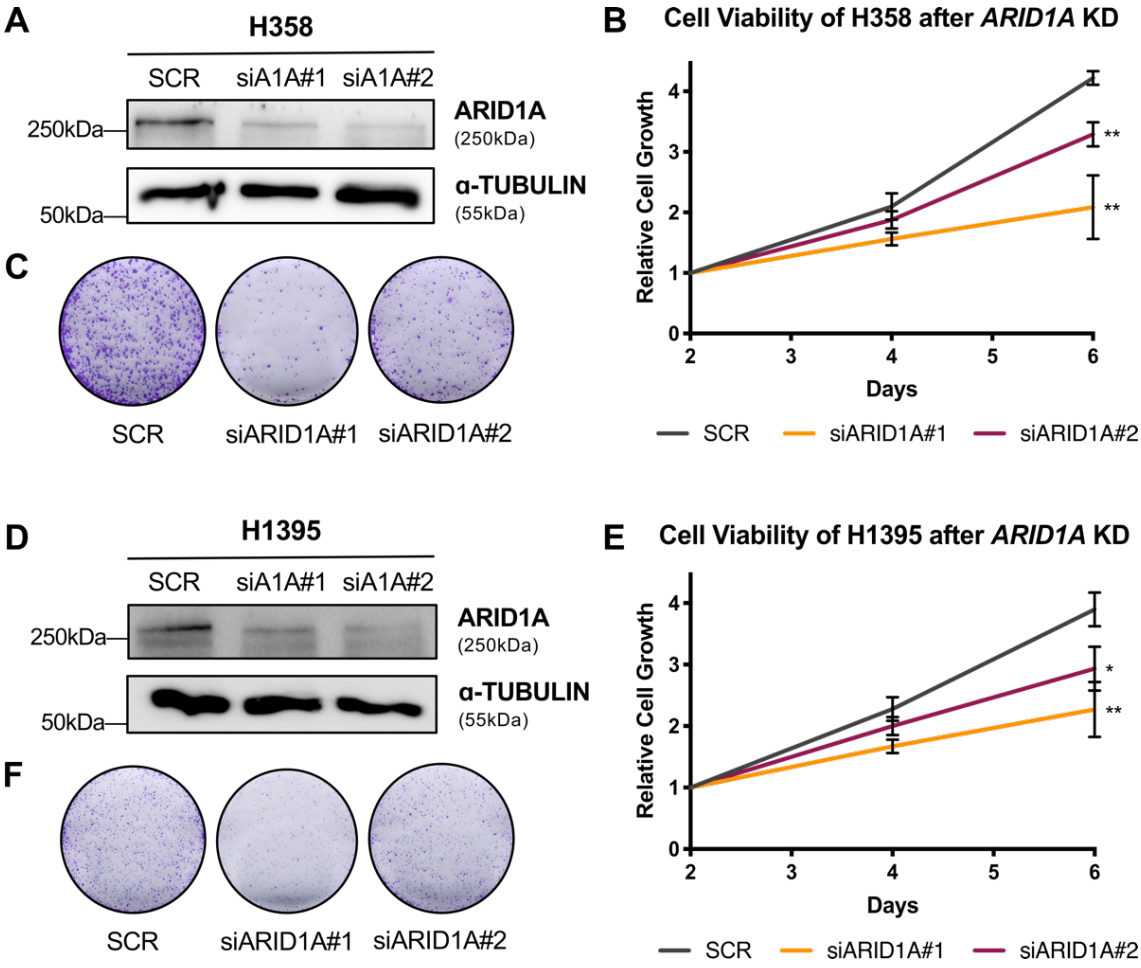


Figure 25: ARID1A dependency of LUAD cell lines with wild type SWI/SNF complexes. (**A**) Western Blot analysis of ARID1A protein levels in the H358 cell line after transfection with either nonsense scrambled siRNAs (SCR) or siRNAs against *ARID1A* mRNA (siA1A#1 or siA1A#2). α-Tubulin was used as a loading control. (**B**) Resazurin assays to measure cell viability at different time points after transfection of H358 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. (**C**) Clonogenic assay after 14 days of transfection of H358 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. (**D**) Western Blot analysis of ARID1A protein levels in the H1395 cell line after transfection with either scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. α-Tubulin is used as a loading control. (**E**) Resazurin assays to measure cell viability at different time points after transfection of H1395 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. (**F**) Clonogenic assay after 14 days of transfection of H1395 with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p-value < 0.05; **p < 0.01.

4.3.2. ARID1A dependency is only observed in tumor contexts

Then, we studied the effect of knocking down *ARID1A* in a dual model of normal epithelial lung cells and their tumor counterpart derived from the previous one: NL20 vs NL20-TA. Interestingly, we only observed a decrease in cell viability and colony formation in the tumor sample, not in the normal epithelial lung cells (***Fig. 26A-D***). Thus, this finding suggested that the ARID1A dependency that we found in different LUAD cell lines could be developed once the tumor is formed.

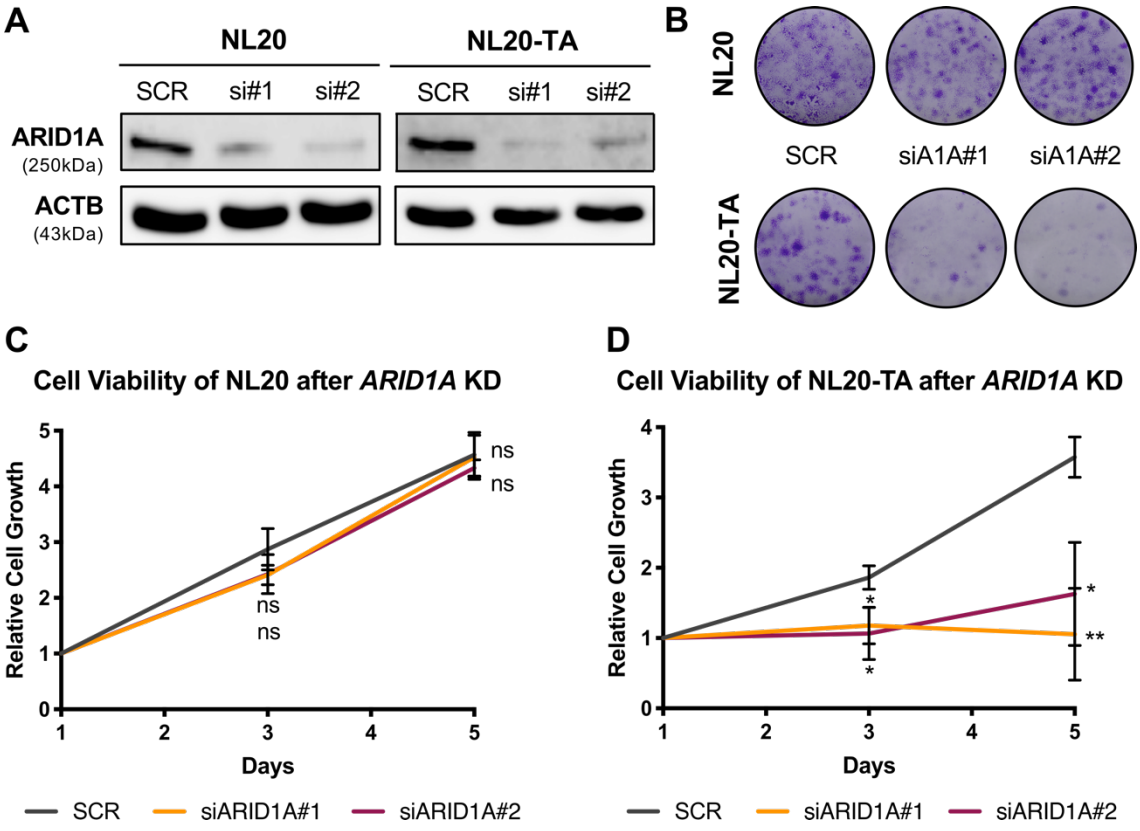


Figure 26: ARID1A dependency only affects tumor contexts. (**A**) Western Blot analysis of ARID1A protein levels in NL20 and NL20-TA cell lines after transfection with either nonsense scrambled siRNAs (SCR) or siRNAs against *ARID1A* mRNA (siA1A#1 or siA1A#2). ACTB was used as a loading control. (**B**) Clonogenic assay after 14 days of transfection of NL20 or NL20-TA with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. (**C**) Resazurin assays to measure cell viability at different time points after transfection of NL20 or NL20-TA. (**D**) with scrambled siRNAs (SCR), siARID1A#1, or siARID1A#2. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p-value < 0.05; **p < 0.01. ns: non-significant.

4.3.3. ARID1A loss triggers DNA damage-induced apoptosis

To uncover the networks behind the impairment of cell viability after silencing *ARID1A* in LUAD cells, we performed an RNA-seq of the *ARID1A* KD model in the A549 cell line. We compared the transcriptome profiles of the A549 after the transfection with either a nonsense scrambled siRNA or an *ARID1A* siRNA and we found that 3090 genes were differentially expressed (FDR < 0.05, absolute FC > 1.5). Specifically, 1770 genes were upregulated and 1320 downregulated after *ARID1A* KD (***Fig. 27A***).

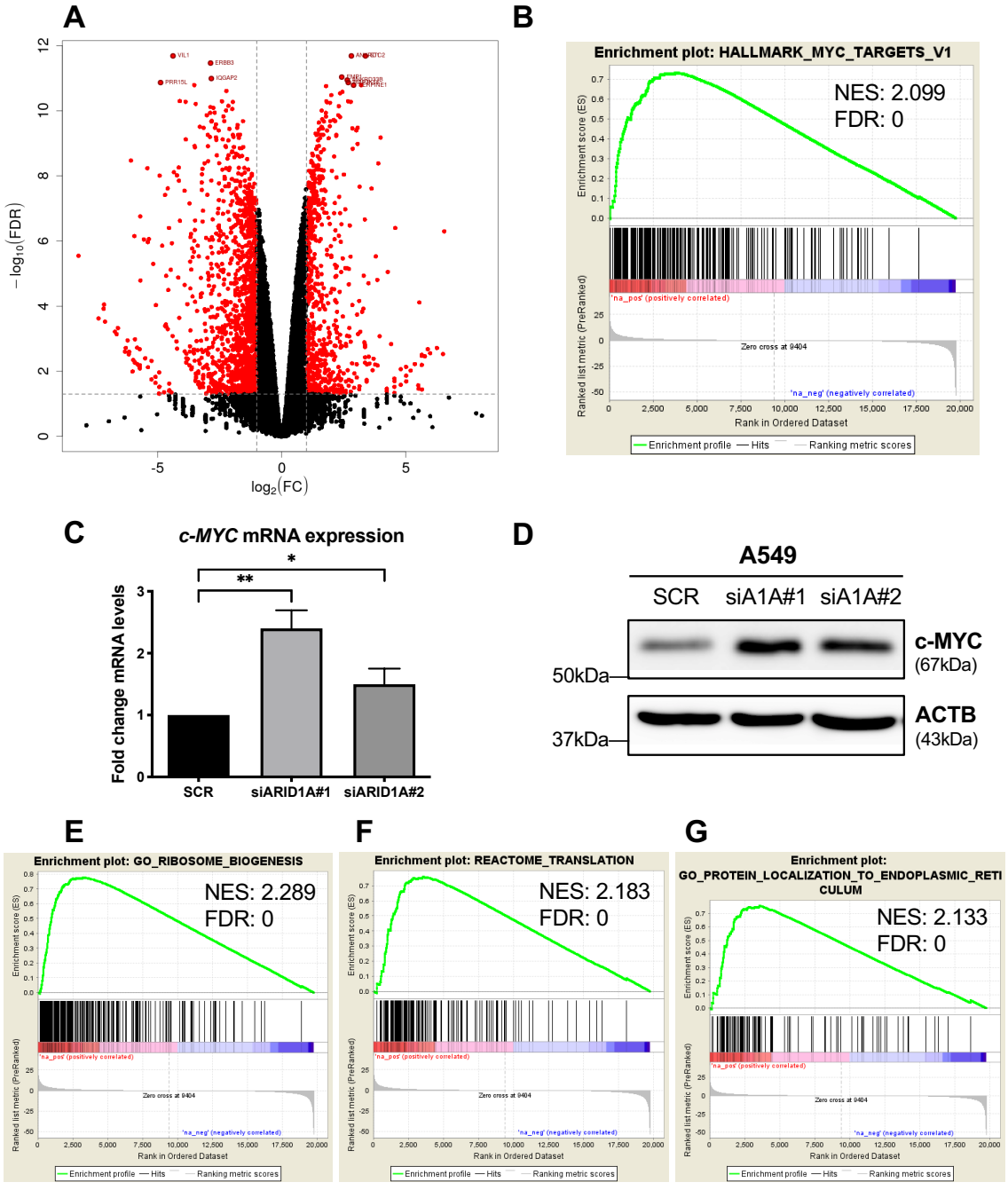


Figure 27: Differentially expressed genes upon *ARID1A* knockdown in the A549 cell line and enriched pathways related to an increase in protein synthesis. (**A**) Volcano plot of the RNA-seq analysis after 6 days of silencing *ARID1A* in A549 cells. Red points represent the genes that are both significant and highly dysregulated. FDR: False Discovery Rate; FC: Fold change. (**B**) Enrichment of the "Hallmark_MYC_targets_v1" in our pathways analysis. NES: Normalized Enrichment Score. (**C**) RT-qPCR analysis of *c-MYC* mRNA levels after *ARID1A* KD in A549. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p-value < 0.05; **p < 0.01. (**D**) Western blot analysis of c-MYC after *ARID1A* KD in A549. ACTB is the loading control. (**E**) Other gene sets significantly enriched in our transcriptome analysis after *ARID1A* KD in A549: GO_Ribosome_biogenesis; (**F**) Reactome_Translation; and (**G**) GO_Protein_localization_to_endoplasmic_reticulum.

Next, we performed a Gene Set Enrichment Analysis (GSEA) to identify enriched biological pathways among the differentially expressed genes using the gene sets derived from the Molecular Signatures Database (MSigDB). Intriguingly, we found some pathways related to an oncogenic stimulus, such as the induction of MYC-targets, ribosome biogenesis, translation, and protein localization to the endoplasmic reticulum (***Fig. 27B, E-G***). However, those pathways coexisted with others that pointed to an elevation of intracellular stress, such as the upregulation of the unfolded protein response (UPR) and reactive oxygen species (ROS) or the decrease of peroxisomal activity (***Fig. 28A***). Interestingly, that signature of intracellular stress derived from an exacerbated oncogenic dysregulation could be responsible for the activation of other gene sets involved in cell death, such as apoptosis and the P53 pathway, which were also enriched in our data (***Fig. 28B***).

We validated by RT-qPCR and western blot the increase in MYC expression that could explain the upregulation of its pathway and the concomitant dysregulation of other oncogenic networks (***Fig. 27C, D***). However, as other authors have described, the upregulation of MYC expression may constitute an intracellular stress that places further weight on protein synthesis and folding (Zhang et al. 2020). Since the UPR and the apoptosis pathways were also enriched in our RNA-seq data, we suspected that MYC overexpression was causing a proteostasis imbalance. The elevation of the protein load of the cells can exceed the folding capacity of the endoplasmic reticulum (ER), leading to severe ER stress, which also increases ROS production and enhances intracellular stress. Thus, the transcriptomic profile of the *ARID1A* KD model could be showing a proteostasis imbalance that induced severe ER stress and a pro-apoptotic UPR signaling. To validate this observation, we

analyzed the expression levels of two regulators of the pro-apoptotic UPR pathway. We found that both, *DDIT3* (also known as *CHOP*) and *ATF5*, were significantly upregulated upon *ARID1A* KD (***Fig. 28D***). Both proteins are relevant transcription factors that are responsible for inducing the expression of pro-apoptotic genes (Tabas and Ron 2011; Teske et al. 2013), such as *BAX*, *BAK1*, *BID*, and *HRK*, which were also upregulated after silencing *ARID1A* (***Fig. 28D***). Importantly, these results coincided with a significant increase of apoptotic cells in our model of *ARID1A* KD after measuring Annexin-V positive cells by flow cytometry (***Fig. 28C***).

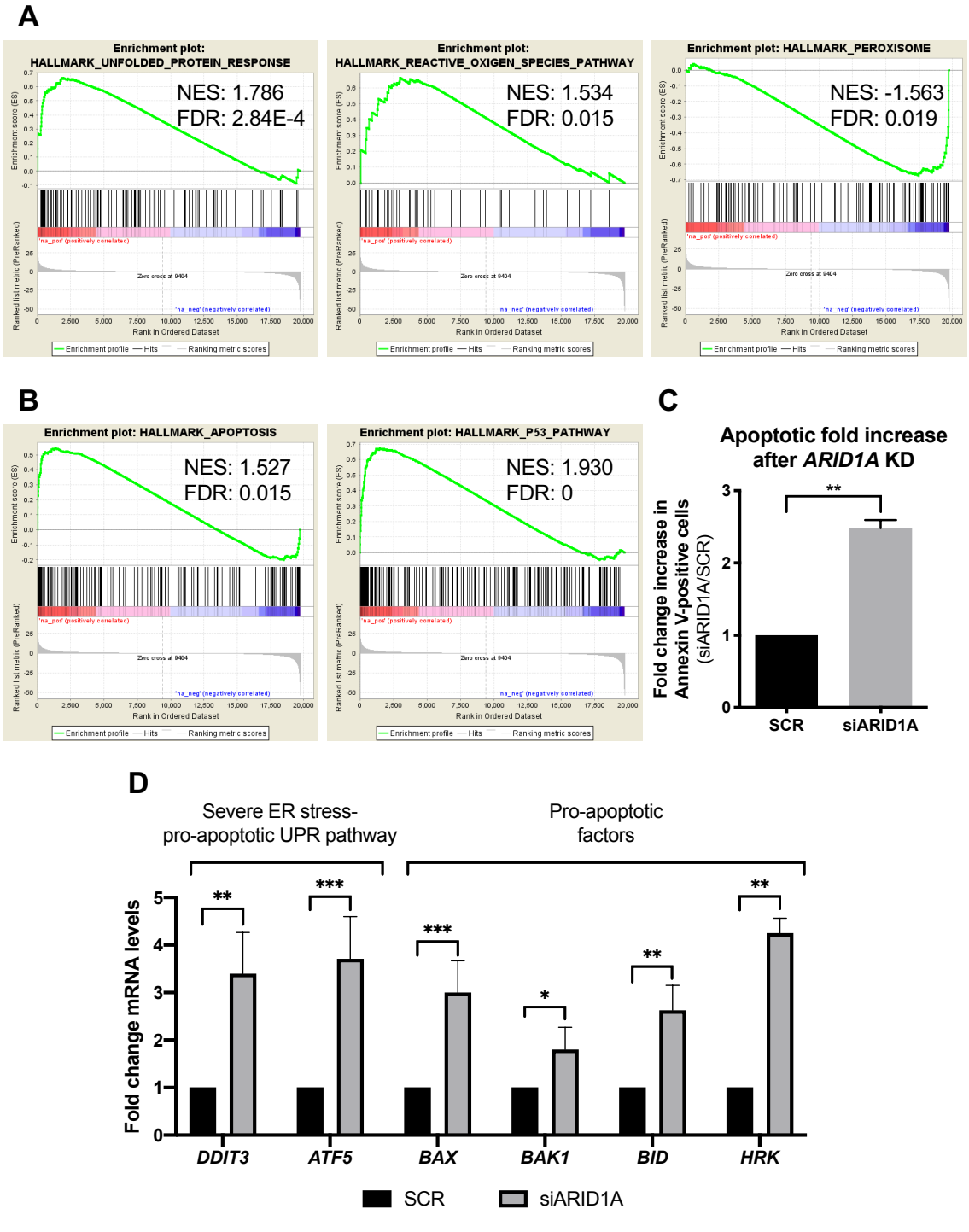


Figure 28: ARID1A loss induces a pro-apoptotic unfolded protein response (UPR). Enriched pathways upon *ARID1A* KD in the A549 cell line related to severe ER stress (**A**) and apoptosis (**B**). NES: Normalized Enrichment Score; FDR: False Discovery Rate. (**C**) Apoptosis analysis by flow cytometry in A549 after *ARID1A* silencing. (**D**) RT-qPCR analysis of a panel of markers of the pro-apoptotic UPR pathway after *ARID1A* KD in A549. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p-value < 0.05; **p < 0.01; ***p < 0.001.

Apart from the profile of cellular stress that we observed in our transcriptomic data, we also found other enriched gene sets related to the UV-response pathway (***Fig. 29A***). This result pointed to a signature of DNA damage and therefore, we wondered whether the depletion of *ARID1A* could also be associated with an increase in DNA damage. It is known that ER stress is interconnected with ROS production, since the protein folding process generates ROS as a byproduct (Malhotra and Kaufman 2007; Santos et al. 2009). In fact, the transcriptome of our *ARID1A* KD model also showed the ROS pathway significantly enriched (***Fig. 28A***). Thus, we suspected that this increase in ROS derived from ER stress could be genotoxic for the cells and potentiate DNA damage (Dizdaroglu et al. 2002). For that reason, we analyzed DNA damage levels by measuring the protein levels of γ-H2AX, a histone mark characteristic of DNA damage (Sharma et al. 2012). We found an increase of this DNA damage-mark after silencing *ARID1A* in LUAD cell lines confirming the genotoxic effect of depleting this gene in LUAD (***Fig. 29B***). In addition, we observed an upregulation of the transcript levels of *E2F1*, a relevant transcription factor that is induced in response to various DNA-damaging agents (Blattner et al. 1999; O'Connor and Lu 2000) (***Fig. 29C***). Importantly, E2F1 is crucial for triggering DNA damage-induced apoptosis (Biswas and Johnson 2012). ATF3 and JUN are also other transcription factors upregulated upon DNA-damage stress with important roles in regulating DNA damage-induced apoptosis (Lu et al. 2006; Turchi et al. 2008; Fan et al. 2002; Lei and Davis 2003; Devary et al. 1991). We also found that these transcription factors presented increased expression levels after silencing *ARID1A* (***Fig. 29C***). Overall, our transcriptomic analysis of the *ARID1A* KD model confirmed that ARID1A is a relevant regulator of cell viability in LUAD. Specifically, we observed that ARID1A fine-tune the expression of important pathways whose dysregulation unbalance protein synthesis, causing an increase of cellular stress and DNA damage that cannot be overcome by the tumor cells and leads to cell death.

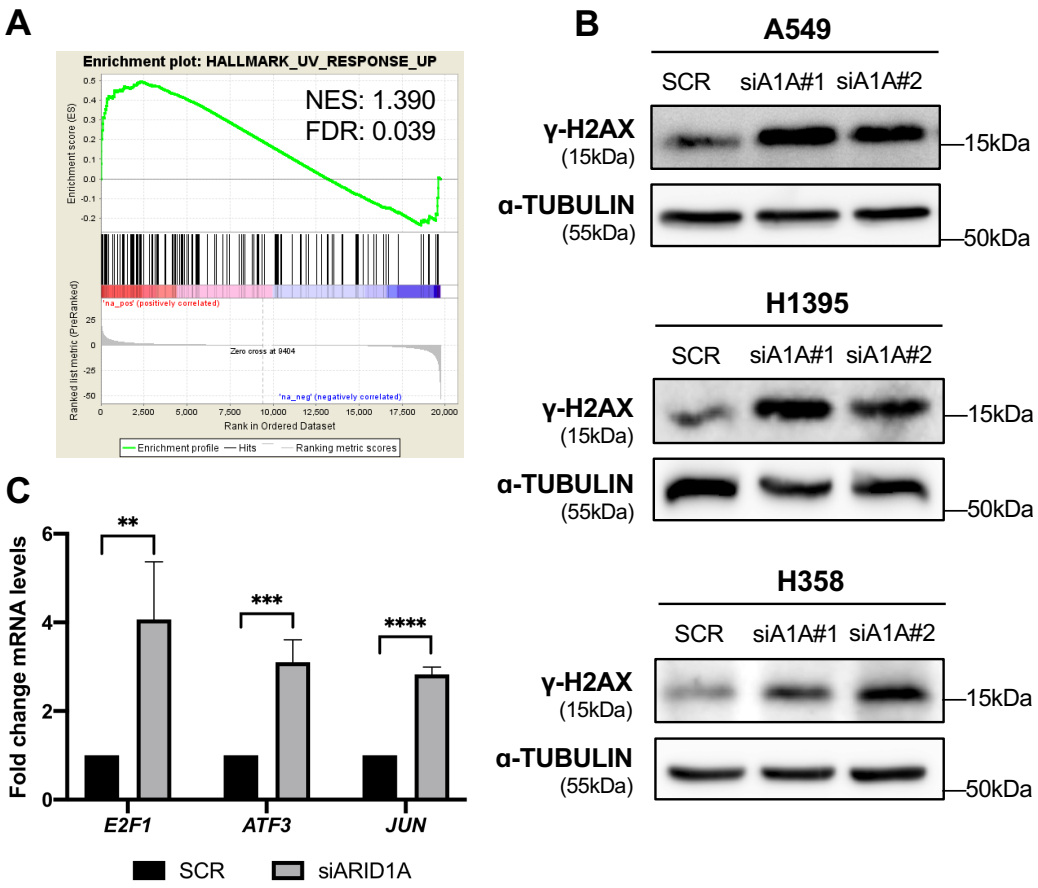


Figure 29: ARID1A loss increases DNA damage in LUAD cell lines. (**A**) Gene set "Hallmark_UV_response_UP" enriched upon *ARID1A* KD in the A549 cell line. NES: Normalized Enrichment Score; FDR: False Discovery Rate. (**B**) Western blot analysis of the DNA-damage mark γ-H2AX in A549, H1395, and H358 cell lines after *ARID1A* silencing. (**C**) RT-qPCR analysis of a panel of markers of DNA damage-induced transcriptions factors that regulate DNA damage-apoptosis after *ARID1A* KD in A549. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p-value < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

4.3.4. Depleting ARID1A behaves as a genotoxic treatment in LUAD

Since we found that knocking down *ARID1A* promoted DNA damage in LUAD cells, we hypothesized that the depletion of ARID1A could behave as a genotoxic agent. Thus, we compared the effect of silencing *ARID1A* with other genotoxic treatments currently widespread in cancer research, such as etoposide and doxorubicin. For that purpose, after transfecting the A549 cell line with the nonsense scrambled siRNA (SCR) or the siRNA against ARID1A (siARID1A), we administered an

additional treatment with either etoposide (10 μM), doxorubicin (0.5 μM), or vehicle/Not Treated.

On the one hand, we observed that only by silencing *ARID1A*, the reduction of cell viability was comparable to that of 0.5 μM doxorubicin treatment. Importantly, this reduction was even significantly higher than only administering etoposide (10 μM) (***Fig. 30, Table 3***).

On the other hand, we found that knocking down *ARID1A* showed an additive effect on cell viability when combined with doxorubicin or etoposide treatments.

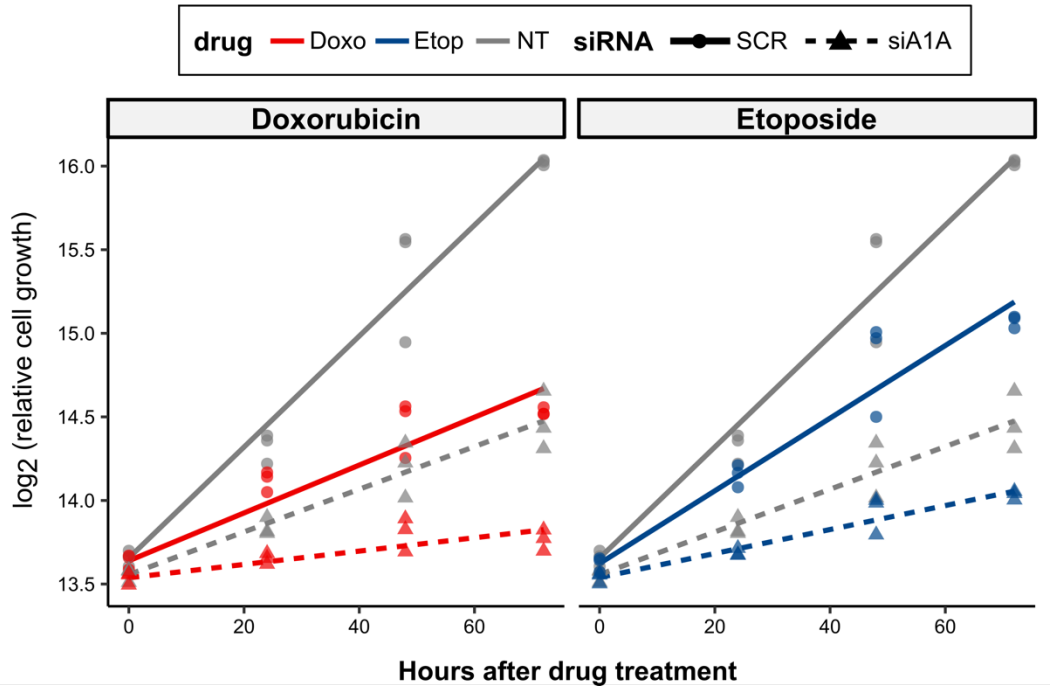


Figure 30: *ARID1A* silencing potentiates the damaging effect of genotoxic drugs. Cell viability assays over time in A549 cells after treatment with a siRNA against ARID1A (siA1A) or a scrambledd siRNA (SCR) plus doxorubicin (Doxo), etoposide (Etop), or vehicle (NT, Not Treated). The experiment was performed in triplicate. The lines summarize the fit of a linear model (see Materials and Methods), averaged across the three replicates. X-axis depicts the $\log_2$ of relative growth values normalized with t=0. Y-axis represents time (hours) after drug treatment. Time 0 corresponds to 48 h after transfection with SCR or siA1A.

***Table 3*: Pairwise comparisons of cell growth rates over time between different treatments of interest.** A549 cells were treated with a siRNA against ARID1A (siA1A) or a scrambledd siRNA (SCR) plus doxorubicin (Doxo), etoposide (Etop), or vehicle (NT, Not Treated). Cell growth rates were estimated as the slopes of the linear fits of $\log_2$(relative cell growth) vs time (in hours; see Materials

and Methods). Then, pairwise comparisons between the slopes were performed for the contrasts of interest. "Estimate": estimated difference in the cell growth rate between the two conditions; SE: standard error; DF: degrees of freedom. *P* values were corrected (*p*.adj) using the Holm method.

CONTRAST	ESTIMATE	SE	DF	T.RATIO	*P*.VALUE	*P*.ADJ
(SCR-NT) - (siA1A-NT)	0,0204	0,0013	55	15,9952	2,43E-22	2,19E-21
(SCR-NT) - (SCR-DOXO)	0,0189	0,0013	55	14,7844	8,26E-21	6,61E-20
(SCR-NT) - (SCR-ETOP)	0,0115	0,0013	55	8,9743	2,32E-12	1,39E-11
(siA1A-NT) - (SCR-DOXO)	-0,0015	0,0013	55	-1,2107	0,2312	0,2312
(siA1A-NT) - (siA1A-DOXO)	0,0088	0,0013	55	6,8738	6,07E-09	1,82E-08
(siA1A-NT) - (SCR-ETOP)	-0,0090	0,0013	55	-7,0209	3,49E-09	1,39E-08
(siA1A-NT) - (siA1A-ETOP)	0,0056	0,0013	55	4,3842	5,30E-05	1,06E-04
(SCR-DOXO) - (siA1A-DOXO)	0,0103	0,0013	55	8,0846	6,35E-11	3,18E-10
(SCR-ETOP) - (siA1A-ETOP)	0,0146	0,0013	55	11,4051	4,08E-16	2,86E-15

To analyze the DNA damage levels produced after all these different treatment conditions, we performed a time-course sample collection to measure γ-H2AX expression by western blot, which is indicative of DNA damage. Interestingly, we observed that the double condition of silencing *ARID1A* and administering etoposide or doxorubicin increased almost twice the levels of DNA damage in comparison with only treating with these genotoxic agents (***Fig. 31A, B***). This observation agrees with our previous results of cell viability, showing that the highest damaging effect of the double treatment siARID1A+Doxo/Etop correlated with the highest levels of γ-H2AX. In addition, we found that in all tested conditions, the DNA damage abruptly decreased over time regardless of the levels of *ARID1A* of the cells, suggesting that all of these treatments triggered an acute cytotoxic response.

Next, we checked the cleaving rate of PARP1 (poly-(ADP-ribose) polymerase 1), a downstream marker of the DNA damage response. In particular, PARP1 is the first responder to DNA damage and its activity mediates DNA repair, but importantly, its activation at high levels also induces apoptosis (Chiarugi and Moskowitz 2002; Pascal 2018). We found that the cleaving rate of PARP1, measured by the ratio Cleaved PARP1/Total PARP1, was higher when *ARID1A* was knocked down with either short-term or long-term exposures to etoposide (***Fig. 31C***). Importantly, we observed that only silencing *ARID1A* showed a cleaving rate higher or similar to that of the treatment with etoposide alone after 6 or 48 h post-administration,

respectively. Thus, this result highlights the potential therapeutic role of ARID1A in LUAD to boost the efficacy of the current genotoxic treatments used in the clinic (reviewed in (Swift and Golsteyn 2014)).

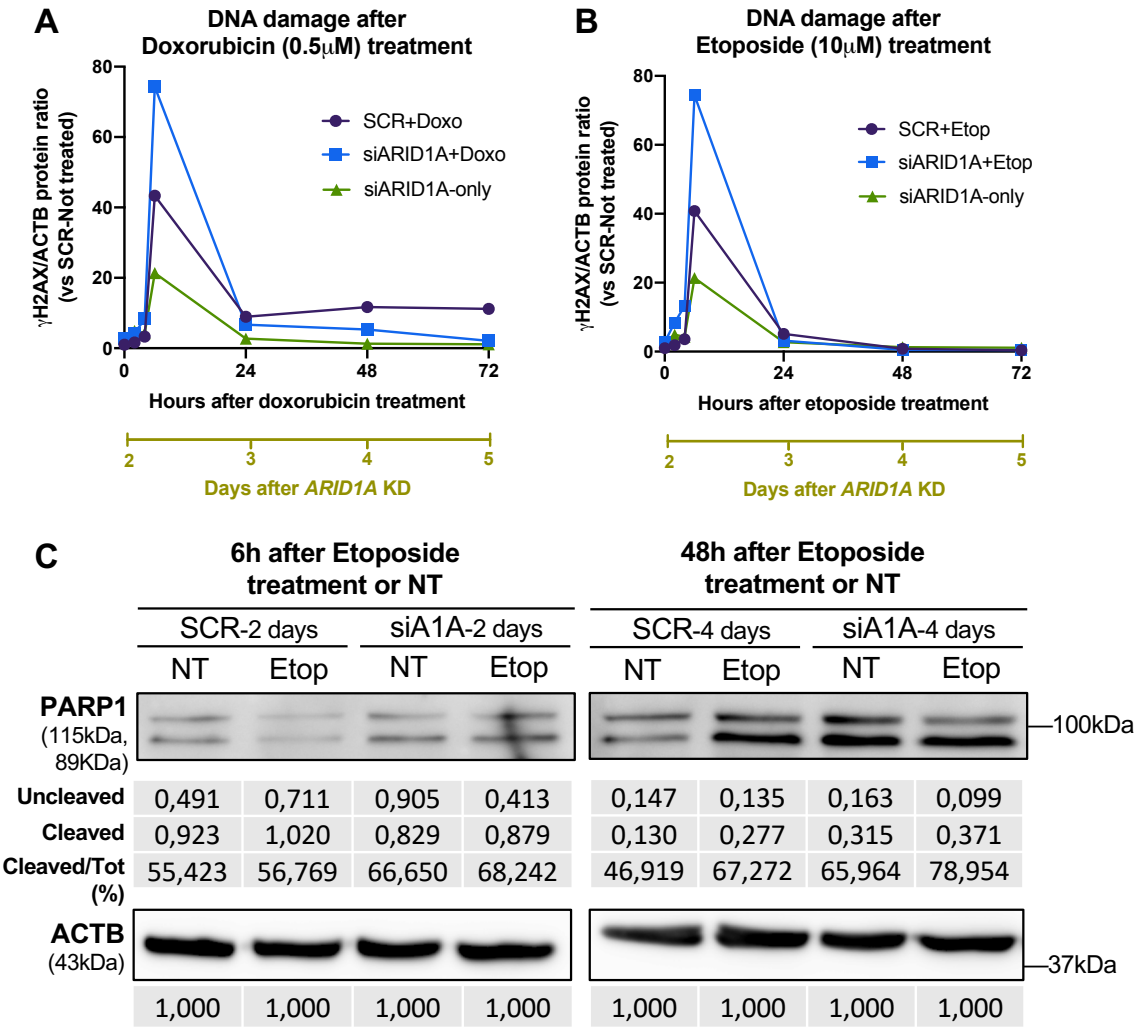


Figure 31: Silencing *ARID1A* potentiates DNA damage in the presence of other genotoxic agents in the A549 cell line. Quantification of the western blot analysis of γ-H2AX levels (see ***Annex 14***) in A549 cells. After 48 h of the transfection with a nonsense scrambled siRNA (SCR) or a siRNA specific against ARID1A (siARID1A), the cells were treated with either doxorubicin 0.5μM (**A**) or etoposide 10μM (**B**). X-axis represents the expression levels of γ-H2AX normalized with the loading control ACTB. Those resulting values were normalized with the condition SCR without any additional treatment. Y-axis shows the different time points that were measured after the treatment with the genotoxic agents. The bar below each of the graphs depicts the corresponding time scale after the transfection with either SCR or siARID1A. (**C**) Western blot analysis of PARP1 levels in the A549 cell line after the transfection of SCR or siARID1A and the subsequent treatment with etoposide 10μM (Etop). NT: Not-treated. The upper band of PARP1 corresponds to the uncleaved protein (115kDa).

The lower band corresponds to cleaved-PARP1 (89kDa). The table below represents the normalized values of each of the bands and the ratio Cleaved/Total PARP1. ACTB was used as the loading control.

4.3.5. Decreasing ARID1A expression alters the protein levels of other BAF and PBAF subcomplexes

Our previous results showed that knocking down *ARID1A* generated a profound dysregulation of relevant pathways that imbalanced the homeostasis of LUAD cells. However, ARID1A is a subunit of only one of the three different SWI/SNF complexes that can be found inside lung cells. Furthermore, although ARID1A is part of the BAF complex, ARID1B can also replace it as the ARID subunit of that type of SWI/SNF complex (Helming et al. 2014b). For all of these reasons, we wanted to evaluate the effect of *ARID1A* silencing on the expression of other components of the remaining SWI/SNF complexes.

Interestingly, we found that SMARCC1, one of the subunits that make up the core module of the three types of the SWI/SNF complex (Mashtalir et al. 2018), showed significantly lower expression levels after *ARID1A* KD (***Fig. 32A, B***). This result suggested that silencing *ARID1A* could have an impact on the expression of other SWI/SNF complexes. This hypothesis was strengthened when we analyzed the protein levels of ARID1B, the alternative ARID subunit of the BAF complexes. In this case, we observed that at shorter times after *ARID1A* KD, there was not a significant change in ARID1B protein levels. However, after longer times of silencing *ARID1A*, the expression of ARID1B significantly decreased (***Fig. 32A, C***). The same trend was found in the other ARID subunit of the PBAF complex: ARID2 (***Fig. 32D, E***). Importantly, SMARCA2, the only ATPase subunit that is present in all the SWI/SNF complexes in the A549 cell line, also showed lower expression levels after *ARID1A* KD (***Fig. 32F***). Overall, we discovered that knocking down *ARID1A* not only altered the protein levels of the BAF complexes that contain this subunit but also impaired the expression of other members of the rest of the different SWI/SNF complexes, including their catalytic subunit.

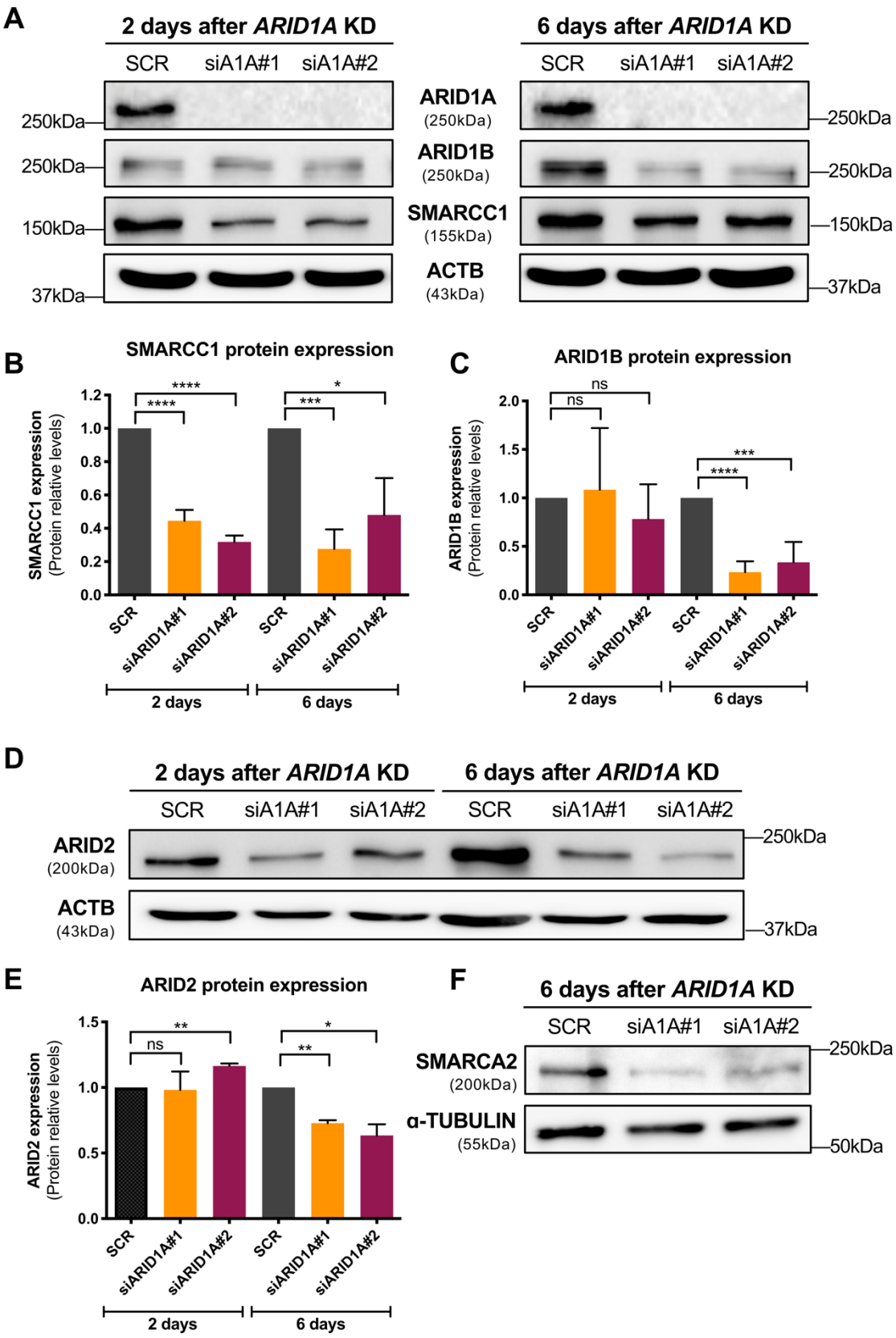


Figure 32: *ARID1A* loss reduces protein levels of other subunits of the SWI/SNF complex. (**A**) Western Blot analysis of ARID1A, ARID1B, and SMARCC1 protein levels in the A549 cell line at different time points after transfection with either nonsense scrambled siRNAs (SCR) or siRNAs

against *ARID1A* mRNA (siA1A#1 or siA1A#2) at different time points. ACTB was used as a loading control. (**B**) Quantification of SMARCC1 western blots normalized with ACTB values. (**C**) Quantification of ARID1B western blots normalized with ACTB values. (**D**) Western blot analysis of ARID2 in A549 at different time points after transfection with SCR, siA1A#1, or siA1A#2. ACTB was used as a loading control. (**E**) Quantification of ARID2 western blots normalized with ACTB values. (**F**) Western blot analysis of SMARCA2 in A549 after 6 days of transfection with SCR, siA1A#1, or siA1A#2. α-Tubulin was used as a loading control. Values represent mean ± SD (n ≥ 3). *Two-tailed t-test *p-value < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; ns: non-significant.

Discussion

5) DISCUSSION

Since the development of next-generation sequencing (NGS) techniques, large-scale genomic studies and projects, such as The Cancer Genome Atlas (TCGA), have shown an elevated incidence of mutations in epigenetic regulators, as well as epigenetic alterations in almost all cancer types (Lawrence et al. 2014; Network 2014). Among the epigenetic regulators currently described, the chromatin remodeling complex SWI/SNF has received great interest. Several studies have underlined the high frequency of mutations or other types of alterations that can affect the SWI/SNF subunits in many types of tumors (Savas and Skardasi 2018; Centore et al. 2020). These findings, plus the diverse range of biological functions in which this chromatin remodeling is involved, have made the SWI/SNF complex a promising candidate for both basic and applied research.

In this Ph.D. thesis, we have focused on the study of the SWI/SNF complex in lung adenocarcinoma (LUAD), one of the deadliest types of cancer today. Our aim was to provide an integral and comprehensive perspective of the whole complex in LUAD patients, as well as in LUAD cell models that are currently used for functional studies. With this thesis, we intended to shed light on the open-ended questions that are currently in this field of cancer research.

5.1. Chapter I: Characterization of the SWI/SNF complex in LUAD

Although tissue specificity is a widely known trait of the SWI/SNF complex, no studies have analyzed its composition in a lung epithelial cell model (Kadoch and Crabtree 2015; Alpsoy and Dykhuizen 2018; Mashtalir et al. 2018; Schick et al. 2019). For that reason, identifying the SWI/SNF subunits that are found in this particular tissue is crucial for a better characterization of the study area.

Specifically, in our analysis of a normal lung tissue model, we did not detect the following nine proteins currently defined as SWI/SNF subunits: ACTB, ACTL6B, BCL7B, BCL11A, BCL11B, BICRAL, DPF1, SS18, and SS18L1. Some of these absent subunits, such as ACTL6B, DPF1, and SS18L1, have exclusively been associated with

the neuronal SWI/SNF complex and, therefore, we did not expect to find them in lung (Olave et al. 2002; Aizawa et al. 2004; Lessard et al. 2007; Staahl et al. 2013). BCL11A and BCL11B have only been validated in human T cells (Kadoch et al. 2013), postnatal brain (Kadoch et al. 2013), and renal cancer cell lines (Alpsoy and Dykhuizen 2018). The same occurred with SS18, which has been associated with BAF complexes since 2002 when Kato and colleagues identified its direct interaction with the SWI/SNF complex in HeLa cells (Kato et al. 2002). Other studies have found it in other tissues, such as the analyzed in the previously cited publications (Kadoch et al. 2013; Mashtalir et al. 2018; Alpsoy and Dykhuizen 2018; Schick et al. 2019), but this subunit was not present in our results. In addition, the only non-canonical (nc) SWI/SNF subunit that we did not identify in our lung model was BICRAL. However, although two independent studies have shown that BICRAL can substitute its homolog BICRA (Mashtalir et al. 2018; Alpsoy and Dykhuizen 2018), a previous publication did not find BICRAL in their mass spectrometric analysis neither, even though the authors studied the same cell model as Mashtalir and colleagues (Middeljans et al. 2012). BCL7B was also absent in our results, but other authors have demonstrated that there are other homologous proteins such as BCL7A and BCL7C, which were also present in our lung model, that can substitute it (Middeljans et al. 2012; Kadoch et al. 2013). In fact, Middeljans and colleagues also described that BCL7B was only associated with SWI/SNF complexes that contained the fusion protein SS18-SSX, an oncogenic subunit involved in synovial sarcoma development (Middeljans et al. 2012). The most unexpected result was the absence of ACTB, which has always been depicted as part of all SWI/SNF complexes studied in different cell types including distinct developmental stages and tissue origin (reviewed in (Kadoch and Crabtree 2015), and recently observed in (Mashtalir et al. 2018; Alpsoy and Dykhuizen 2018)). However, other researchers did not find any interaction between ACTB and the different SWI/SNF complexes evaluated (Middeljans et al. 2012). These authors identified a member of other actin families: actin gamma, which was also present in our lung model. Nevertheless, there is a lack of studies that have analyzed the functional repercussion of this switch between ACTB and other actin families.

In short, we discriminated the lung SWI/SNF subunits among other tissue-specific components of this complex (***Figure 33***). We confirmed that only certain SWI/SNF subunits were selectively incorporated into lung SWI/SNF complexes, conferring the biological specificity of those chromatin remodeling complexes (Euskirchen et al. 2011; Hodges et al. 2018).

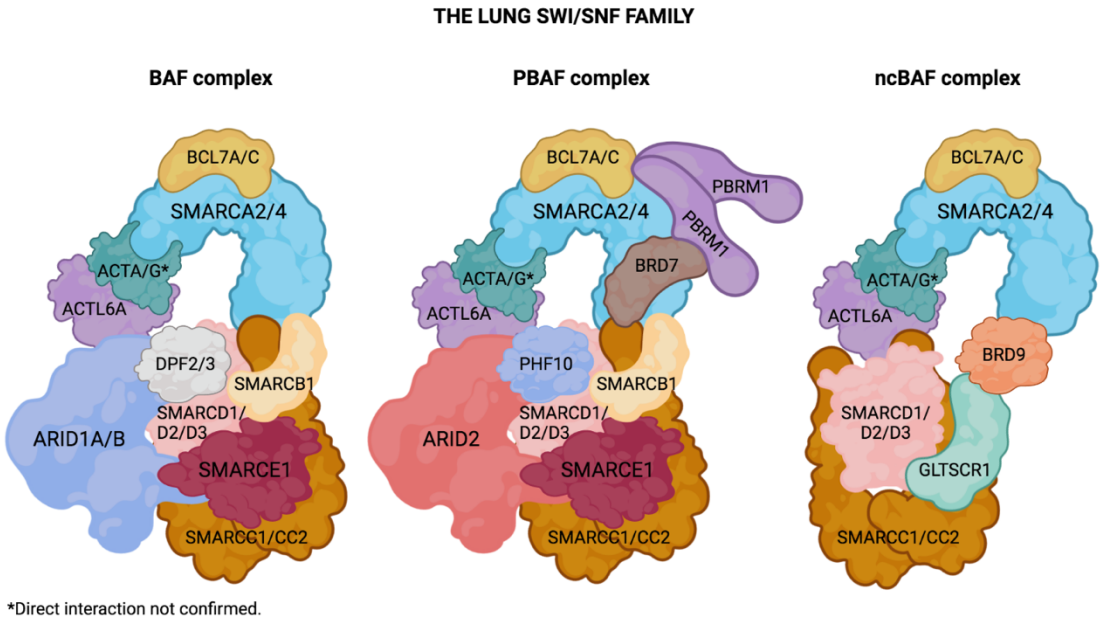


Figure 33: Graphical model of SWI/SNF complexes in lung cells. Compilation of lung SWI/SNF subunits that can compose the canonical BAF complex (BRM/BRG1 Associated Factors), PBAF (polybromo-associated BAF complexes), and ncBAF complex. Image made with BioRender.com and adapted from (Mashtalir et al. 2018) using our LC-MS/MS data of epithelial lung cells (***Annex 6***).

When we studied the lung SWI/SNF subunits in a cohort of 70 LUAD patients, we found that more than 41% of patients harbored a somatic mutation in any of those SWI/SNF subunits. Interestingly, we also observed that SWI/SNF-mutant tumors had an elevated Tumor Mutation Burden (TMB), supporting the relevant role of the SWI/SNF complex in the maintenance of genome integrity. Several articles have associated the SWI/SNF complex with functions in DNA repair (Harrod et al. 2020; Ribeiro-Silva et al. 2019) and replication stress (Gupta et al. 2020; Kurashima et al. 2020; Bayona-Feliu et al. 2021; Tsai et al. 2021), two processes that highly contribute to genome instability. Interestingly, this high TMB of SWI/SNF-mutant tumors could explain the worse prognosis that we observed in that subgroup of patients. In fact, other studies have also associated high TMB with poorer outcomes

in different cancer types (Hwang et al. 2019; Owada-Ozaki et al. 2018; Eder et al. 2019). As Valero and colleagues discussed in their work (Valero et al. 2021), high TMB could be responsible for (1) an increased likelihood of generating oncogenic drivers or mutations that could mediate therapeutic resistance (Bozic et al. 2013), and (2) an increased intratumor genetic heterogeneity that facilitates tumor growth under selective pressure (Morris et al. 2016; Andor et al. 2016). However, an elevated TMB can be a good prognostic biomarker when analyzing the outcomes of immunotherapy for cancer patients. Importantly, immune checkpoint inhibitors (ICIs) are currently one the most relevant therapeutic treatments for human tumors (He and Xu 2020). Nevertheless, despite the remarkable success of ICIs in the past decade, the efficacy and effectiveness of these therapies vary greatly not only among different cancer types but among individual patients. The overall response rate of ICIs is less than 30% of cancer patients, and even NSCLC patients can show lower response rates (≤20%) (Nishino et al. 2017). Moreover, researchers and clinicians have observed that a significant number of ICIs therapy recipients (up to two-thirds of them) exhibit either primary or acquired resistance (Wang and Wu 2017; Restifo et al. 2016). Therefore, the identification of biomarkers that predict patients who are more likely to respond to ICIs is crucial. That is why the discovery of increased effectiveness of ICIs in tumors with a high mutation burden has arisen much interest. Different studies have observed that cancers that arise from chronic exposure to DNA-damaging agents, such as tobacco carcinogens or ultraviolet radiation, or those that develop DNA repair defects present better response rates to ICI therapy (Rizvi et al. 2015; Mouw et al. 2017; Le et al. 2017; Valero et al. 2021). Specifically, a high TMB is associated with more variety of neoantigens produced by tumor cells that facilitate the recognition and activation of T cells, enhancing anti-tumor immune responses (Jardim et al. 2021). For this reason, our finding of the higher TMB of SWI/SNF-mutant LUAD tumors confers an additional clinical interest, not only for its connection with worse overall survival but for defining a clinical subgroup that could benefit from ICI therapy. To date, other studies have also found associations between some SWI/SNF subunits and a better response rate to ICI therapy for different tumor types highlighting this relevant connection (Shen et al. 2018; Braun et al. 2019; Naito et al. 2019; Zhu et al. 2021; Zhou et al. 2021). In addition, Pan and colleagues have proposed that the involvement of the SWI/SNF

complex in the response rate to ICIs can be explained due to its transcriptional role by modulating the expression of chemokines (Cxcl9 and Cxcl10) that result in more efficient recruitment of effector T cells into tumors (Pan et al. 2018). However, currently, more studies are necessary to translate these findings to the clinic. Importantly, the relevance of the SWI/SNF in the clinical practice is underlined when compared with any of the top 10 LUAD driver genes identified by Bailey and colleagues (ATM, BRAF, EGFR, KEAP1, KRAS, NF1, RB1, RBM10, STK11, and TP53). We found that there was not a significant correlation between their mutational status and the overall survival of LUAD patients. On the contrary, the mutant status of the lung SWI/SNF complex not only correlated with worse overall survival but was an independent prognostic factor when evaluated alongside other clinical variables commonly associated with LUAD survival.

Although many authors have focused on the study of the mutational status of the SWI/SNF complex due to the high prevalence of its genetic alterations, other researchers have observed significant expression changes of the members of this chromatin remodeler in several types of cancer (Marquez et al. 2015; Glaros et al. 2007; Zhang et al. 2018). For this reason, we also analyzed the expression levels of the lung SWI/SNF subunits. Interestingly, we found an extensive downregulation of the SWI/SNF complex regardless of its mutational status. As we have mentioned before, previous studies have also described expression alterations in some SWI/SNF subunits in many tumor types (Schallenberg et al. 2020; Li et al. 2016b; Reisman et al. 2003; Yamamichi et al. 2005; Kuo et al. 2006; Takao et al. 2017; Han et al. 2020; Park et al. 2014; Papp et al. 2013), highlighting the importance of the adequate regulation of SWI/SNF expression to prevent tumorigenesis. However, the novelty of our results lies in the general downregulation of the whole SWI/SNF complex that we observed in LUAD patients, suggesting a possible transcriptional coregulation that affects the SWI/SNF members. This observation coincides with the results obtained by Schick and colleagues when they performed specific knockout (KO) clones of different SWI/SNF subunits (Schick et al. 2019). In fact, they observed that the SMARCA2^{KO} model performed in the HAP1 cell line showed lower expression levels of the rest of 28 SWI/SNF subunits in comparison with the wild-type HAP1 cells. Interestingly, the most significantly downregulated SWI/SNF

subunit in our cohort of LUAD patients was SMARCA2 (adjusted *p*-value: $3.28 \cdot 10^{-21}$), and this downregulation coexisted with less expression of many other SWI/SNF subunits. However, to date, there are no additional studies that confirm the mechanism behind a transcriptional coregulation of the SWI/SNF complex.

This also shows the importance of having solid models that help to analyze the biochemical roles of the SWI/SNF complex in well-studied cellular contexts. For this reason, we also wanted to provide a resource for researchers where they can find a compilation of the most frequently used LUAD cell lines and their SWI/SNF profile at the genomic, transcriptomic, and protein levels of the five most characteristic and recurrently altered subunits. Specifically, we found that more than 76% of our panel of LUAD cell lines harbored a genetic alteration in at least one out of the twenty subunits of the SWI/SNF complex analyzed. This high mutation rate contrasts with the data obtained with our cohort of LUAD patients, which showed that only 41% of LUAD tumors had SWI/SNF mutations. But, interestingly, this increase in the mutation frequency in cell lines in comparison with primary tumors was also observed by other authors (Blanco et al. 2009). Blanco and colleagues reported that lung cancer cell lines harbored about twice as many mutations in any known cancer gene as those detected in primary lung tumors. They explained this effect as the result of the contamination of primary tumors with normal cells, which masks the actual mutation frequency of the tumors. In addition, other researchers have described genetic changes between cell lines and primary tumors derived from the culture procedures used for cell lines (Wilding and Bodmer 2014; Roschke et al. 2003). However, Wilding and colleagues concluded that there is no evidence of major genetic changes produced after long-term *in vitro* cultivation. They observed that cell lines and primary tumors share the same patterns of mutation, copy number variation, methylation, and mRNA expression. Indeed, we found that *SMARCA4*, *ARID1A*, and *ARID2* were the top mutated SWI/SNF genes in our panel of LUAD cell lines, as they were in our cohort of LUAD patients and in other studies (Bailey et al. 2018; Imielinski et al. 2012; Network 2014).

Interestingly, we also observed that most of the genetic alterations of our panel of LUAD cell lines could impact the functionality of the mutated subunit, even missense

mutations, according to SIFT predictions. This agrees with the recently published observations that validate the deleterious effect of missense mutations of SMARCA4, supporting *in silico* predictions (Fernando et al. 2020; Hodges et al. 2018). In addition, *SMARCA4* was the SWI/SNF gene that displayed the highest number of deleterious mutations considering splice-site, stop-gain variants, frameshift indels, and large deletions. Specifically, 66% percent (12/18) of the deleterious mutations found in the 20 SWI/SNF subunits of our study were located in *SMARCA4*. Furthermore, this gene concentrated 72% (13/18) of all of the homozygous alterations of our mutational analysis of LUAD cell lines. To date, it is not clear why SMARCA4 is the most deleteriously mutated SWI/SNF subunit in LUAD (Bailey et al. 2018; Imielinski et al. 2012; Network 2014). Some studies have linked the proximity of the *SMARCA4* locus to the *STK11* locus (another well-known LUAD tumor suppressor gene) and the loss of heterozygosity that recurrently affects the short arm of the chromosome 19 in LUAD (Rodriguez-Nieto and Sanchez-Cespedes 2009).

According to our mutational data, the most suitable LUAD models of a defective SWI/SNF complex with deleterious and homozygous mutations in *SMARCA4* are: A427, A549, H1568, H1573, H1793, H1944, H2030, H23, and H838 cell lines. The cell lines H441 and SKLU-1 are two examples of a model with homozygous and deleterious mutations in *ARID2* and *ARID1A*, respectively. On the contrary, identifying a LUAD cell line without any aberrant SWI/SNF complex was more difficult, especially after evaluating the expression levels of the subunits of this complex. Indeed, when we measured the mRNA levels of the lung SWI/SNF subunits in our panel of LUAD cell lines, and we found that they displayed either a transcriptional downregulation or upregulation. In this case, we performed the $-\Delta\Delta$Ct analysis using the median ΔCt for each of the measured genes. This analysis provided a transcriptional comparison among the LUAD cell lines of our panel of study, showing the differential expression regulation of the SWI/SNF subunits within these cell lines. In addition, as we also observed in LUAD tumors, the relevance of a regulatory mechanism behind the control of SWI/SNF expression was underlined when, regardless of the presence of genetic alterations, the expression of a particular SWI/SNF subunit could be reduced or enhanced. Intriguingly, several LUAD cell lines, such as H1395, displayed a transcriptional downregulation of most

of the SWI/SNF subunits, which could be indicative of a still unknown regulatory system of the SWI/SNF complex subunits. Many researchers have pointed to different epigenetic mechanisms to explain these changes in SWI/SNF expression. Those include modifications of methylation patterns in the CpG islands of promoters (Luo et al. 2020b; Wu et al. 2019; Khursheed et al. 2013), alterations of regulatory histone marks in the promoter region (Januario et al. 2017), post-translational modifications that change the stability of the proteins (Macher-Goeppinger et al. 2015; Bock et al. 2011; Jiang et al. 2015), and post-transcriptional inhibition by microRNAs (Coira et al. 2015; Arts et al. 2017; Taulli et al. 2014; Sakurai et al. 2011).

Importantly, according to several studies, mutation or loss of expression of a SWI/SNF subunit does not fully inactivate the SWI/SNF complex but creates alternative residual complexes that can drive genome regulation of tumor cells (Oike et al. 2013; Helming et al. 2014b; Wang et al. 2009; Hoffman et al. 2014; Schiaffino-Ortega et al. 2014). Schick and colleagues not only demonstrated that the depletion of some SWI/SNF subunits did not impair the assembly of the SWI/SNF complex, but more importantly, that the remaining complexes were aberrantly targeted to their corresponding genomic regions (Schick et al. 2019). Specifically, these authors showed that intact SWI/SNF complexes containing SMARCA4, ARID1A, and SMARCC1 were crucial for targeting cell-specific enhancer sites and enabling their activation in the HAP1 cell line. Moreover, they described different recognition specificities of the distinct BAF and PBAF complexes that coexist in a wild-type cellular model. Thus, impairing the formation of one of the SWI/SNF subcomplexes could modify chromatin accessibility, and consequently, transcription. For this reason, studying the residual SWI/SNF complexes inside tumor cells with altered SWI/SNF subunits provides new knowledge about tumor biology and novel tools for developing better cancer therapies (reviewed in (Mittal and Roberts 2020; Centore et al. 2020)). Therefore, we decided to analyze in our panel of LUAD cell lines the protein levels of four SWI/SNF subunits that determine the assembly of the different SWI/SNF subcomplexes, according to the model proposed by Masthalir and colleagues (Mashtalir et al. 2018). These subunits were the ATPases SMARCA4 and SMARCA2, which are essential for the remodeling activity of the complex, and the ARIDs subunits: ARID1A, ARID1B, and ARID2, which

are crucial for targeting the SWI/SNF complex to different genomic regions (Chandler et al. 2013). With our western blot analysis, we found that although SMARCA4 was the most mutated SWI/SNF subunit, ARID1A was the subunit with the highest incidence of protein loss in our 38 LUAD cell lines. In addition, ARID1A was the subunit with more cases of protein loss not caused by mutations in its gene. In fact, Hung and colleagues have recently noted that the mutation status of ARID1A is an unreliable predictor of ARID1A expression in NSLCL patients (Hung et al. 2020a). Importantly, this and other studies have demonstrated the clinicopathological associations of ARID1A expression in NSCLC. They observed the ARID1A loss of expression correlated with worse overall survival of NSCLC patients, although this correlation was only significant in the squamous cell carcinoma subtype (Jang et al. 2020; Hung et al. 2020a). Overall, these results show a new subset of LUAD patients with loss of ARDI1A that could be interesting for developing new therapies. To date, one of the most exploited synthetic lethal approaches in ARID1A-deficient tumors is targeting its paralog ARID1B (Kelso et al. 2017; Mathur et al. 2017; Helming et al. 2014b). However, we found a subgroup of LUAD cell lines that not only lacked ARID1A protein but also ARID1B and ARID2, the two alternative ARID subunits. This subgroup of LUAD cell lines that we classified as "ARID-deficient" could be especially interesting for studying the role of the recently described non-canonical BAF complex (ncBAF). Indeed, other researchers have already described the potential therapeutic vulnerabilities that are mediated by ncBAF complexes in other types of tumors (Michel et al. 2018).

Moreover, we found another subgroup of LUAD cell lines that we classified as "ATPase deficient", and that resembled a subgroup of LUAD patients with the loss of both SMARCA4 and SMARCA2 (Reisman et al. 2003; Herpel et al. 2017). This subgroup has gained additional interest since this specific profile leads to more aggressive outcomes (Marquez-Vilendrer et al. 2016). For this reason, having these ATPase-deficient cellular models can facilitate the understanding of this specific condition and lay the foundation for the development of new therapeutic strategies to treat these aggressive LUAD tumors.

5.2. Chapter II: Regulation of microRNA expression by SMARCA4, the catalytic subunit of the SWI/SNF complex

To date, several studies have experimentally demonstrated the implications of the SWI/SNF complex in chromatin regulation in diverse biological contexts, either related to tumorigenesis (Banine et al. 2005; Romero et al. 2012; Orvis et al. 2014; Hodges et al. 2018; Wang et al. 2017; Mathur et al. 2017; Karnezis et al. 2016; Schick et al. 2019) or cell development (Kaeser et al. 2008; Ho et al. 2009; Iurlaro et al. 2021). Many of these studies have also shown that aberrant SWI/SNF complexes, due to genetic alterations or loss of expression of some of their subunits, have a pronounced effect on the transcriptome of the cells and, therefore, on tumorigenesis too. However, those analyses have been focused on protein-coding genes without evaluating the impact on non-protein-coding genes. Here, we show that re-expression of SMARCA4 in a SMARCA4-deficient LUAD cell line affects other master regulators of gene expression: microRNAs (miRNAs). Previous articles have described that miRNA biogenesis is controlled by epigenetic mechanisms, such as methylation and changes of histone modifications (Liu et al. 2013; Sato et al. 2011; Moutinho and Esteller 2017). Since these epigenetic marks modify chromatin structure and accessibility, this supports a connection between chromatin remodeling and the so-called miRNome: the set of expressed miRNAs at a given time in a cellular context.

We observed that SMARCA4 re-expression in A549 cells induced significant changes in the miRNome, and many of those miRNAs were associated with tumor development. Importantly, the design of our study, by not using a stable cellular model, allowed us to detect the expression changes immediately triggered after SMARCA4 restoration. Specifically, Schick and colleagues have underlined the relevance of performing studies that analyze the cellular context after an immediate change in the SWI/SNF composition (Schick et al. 2019). They point out that during the acquisition of stable cellular models, rewiring mechanisms occur in tumor cells to support their viability and proliferation. Thus, our study model prevented those drawbacks without losing efficacy in restoring SMARCA4 expression and its incorporation into the endogenous SWI/SNF complexes of the A549 cell line.

Among the miRNAs that changed their expression upon SMARCA4 restoration, miR-222 was the most significant dysregulated. Interestingly, many studies have also found this miRNA dysregulated in several tumor types (reviewed in (Song et al. 2017; Amini et al. 2019)). However, the role of miR-222 is controversial. Some studies have shown that miR-222 has oncogenic properties and that it is upregulated in tumors (Visone et al. 2007; Garofalo et al. 2009; Zhang et al. 2010; Xue et al. 2017), but in other types of cancer it shows a tumor suppressor role (Fu et al. 2016; Liu et al. 2009; Medina et al. 2008b). Moreover, this contradictory role has been observed even within lung cancer (Yamashita et al. 2015). Nevertheless, several researchers have also described this dual effect of some miRNAs depending on the tumor type and the cellular contexts analyzed (Li et al. 2019; 2016a; Grossi et al. 2018; Ferrari and Gandellini 2020). In fact, the functional consequences of the dysregulation of a specific miRNA not only depend on its targets but also on the consequent alterations of the pathways where those targets participate. For that reason, we decided to validate the miR-222 function in our LUAD cell model. We observed that miR-222 overexpression impaired cell viability and clonogenicity. Interestingly, this phenotype was similar to that obtained after SMARCA4 restoration. Mallappa and colleagues had also observed that the loss of DICER, which is one of the key enzymes involved in miRNA biogenesis, phenocopied the depletion of SMARCA4 in zebrafish (Mallappa et al. 2010). Their work demonstrated for the first time, albeit indirectly, the implications of the SWI/SNF complex on the expression of miRNAs. Although we corroborated that miR-222 behaved as a tumor suppressor miRNA in our LUAD model, SMARCA4 is part of chromatin remodeling complex with a genome-wide regulatory activity, and therefore, miR-222 is not the only contributor to the tumor suppressor role of SMARCA4 in LUAD (Medina et al. 2005; Banine et al. 2005; Romero et al. 2012; Song et al. 2014; Orvis et al. 2014). Indeed, other authors have discovered that SMARCA4 is responsible for the regulation of 4.8-20% of the human genome (Zhang et al. 2014; Raab et al. 2016). Moreover, Schick and colleagues underlined that, among all the SWI/SNF subunits, SMARCA4, SMARCC1, and ARID1A showed the highest impact on chromatin accessibility and regulation of gene expression after their depletion (Schick et al. 2019).

Importantly, we have observed for the first time that the SWI/SNF complex modulates miR-222 by binding to its enhancer region. Although the SWI/SNF complex can bind to promoters (Tolstorukov et al. 2013; Lu and Roberts 2013), it is also well-known that it can regulate gene expression through enhancers (Lazar et al. 2020; Xue et al. 2019; Shi et al. 2013; Hodges et al. 2018; Wang et al. 2017; Alver et al. 2017; Mathur et al. 2017; Bossen et al. 2015; Vierbuchen et al. 2017; Hu et al. 2011). Moreover, a recent publication has shown that SMARCA4 restoration increases enhancer-associated histone marks, supporting the relevance of the SWI/SNF complex at these specific regulatory regions (Lazar et al. 2020). Other authors explain this preferential binding of the SWI/SNF complex to enhancer regions because each cell type tends to vary more enhancer utilization than promoter openness (Thurman et al. 2012; Song et al. 2011).

Interestingly, we only observed the upregulation of miR-222 when SMARCA4 levels were at their peak of expression, but when SMARCA4 levels decreased, the binding was replaced by SMARCA2 and miR-222 levels dropped. Thus, we discovered that this change of the catalytic subunit of the SWI/SNF complex generated an opposite effect on miR-222 expression. Our finding supports two previous observations: the elevated dynamism of the SWI/SNF complex and the non-redundant role of the mutually exclusive subunits of this chromatin remodeler. Regarding the dynamic function of the SWI/SNF complex, Iurlaro and colleagues have recently demonstrated the dynamism of the SWI/SNF complex in regulating chromatin accessibility (Iurlaro et al. 2021). Specifically, they showed how the alterations of the SWI/SNF complex could affect chromatin structure within minutes. In addition, other authors had previously expressed their support to studies that included analysis performed immediately after removing SWI/SNF subunits to study the direct implications of those changes in tumor development (Schick et al. 2019). Since we performed a transient restoration of SMARCA4 in a LUAD cell line model, our work shows the dynamism of a cellular context with a transition between SMARCA2– and SMARCA4–SWI/SNF complexes and vice versa. Consequently, the abrupt changes of miR-222 expression depending on the catalytic subunit that drives the SWI/SNF complex reflect the highly variable functions of these chromatin remodelers.

On the other hand, this result also shows the different effects that the ATPase-helicase subunit of the SWI/SNF complex can have on chromatin regulation. Although some studies have suggested complementary roles for SMARCA2 and SMARCA4 (Willis et al. 2012; Strobeck et al. 2002), other researchers have observed that in SMARCA4-mutant contexts, SMARCA2 acquires distinct roles during the oncogenic transformation and becomes an essential gene (Wilson et al. 2014). Moreover, Schick and colleagues have found that the depletion of SMARCA4 or SMARCA2 in an isogenic cell model showed different patterns of chromatin accessibility and gene expression (Schick et al. 2019). Some authors described that the different functional roles of both ATPases could be the result of their unique recruitment of transcription factors to the chromatin (Kadam and Emerson 2003).

All of these studies support our observation that the downregulation of miR-222 after the decrease of wild-type SMARCA4 could be the consequence of SMARCA2 acting as a part of a residual SWI/SNF complex in a SMARCA4-mutant context. As Mathur and Roberts have discussed, tumor cells that harbor mutations in the subunits of the SWI/SNF complex present residual complexes that maintain the function of this chromatin remodeler (Mathur and Roberts 2018). However, those residual SWI/SNF complexes can be mistargeted (Schick et al. 2019) or have other effects on gene expression as it happens with miR-222 expression after SMARCA4 restoration.

In conclusion, we report that the expression of miR-222 in LUAD cells is under control of the SWI/SNF complex and shows different expression patterns depending on the composition of the catalytic subunit of the complex. Thus, this finding suggests a change derived from the transformation of SMARCA2 to an essential gene in SMARCA4-mutant lung adenocarcinomas.

5.3. Chapter III: Functional study of the SWI/SNF subunit ARID1A in LUAD

For many years, most of the studies of the SWI/SNF complex in lung adenocarcinoma were focused on its catalytic subunits SMARCA4 and SMARCA2. This can be explained because, before the expansion of NGS, *SMARCA4* was the first SWI/SNF gene found mutated (Wong et al. 2000; Medina et al. 2004; 2008a; Rodriguez-Nieto et al. 2010) or with loss of expression (Reisman et al. 2003) in lung tumors. In addition, the discovery of synthetic lethalities in SMARCA4-deficient tumors reinforced the interest of many researchers to continue their study, and it opened a new field of therapeutic strategies for this type of cancer (Xue et al. 2019; Rago et al. 2018; Tagal et al. 2017; Lissanu Deribe et al. 2018; Hoffman et al. 2014; Oike et al. 2013). However, the development of NGS techniques and high-throughput analyses allowed a better characterization of the mutational and expression profile of LUAD tumors. Those studies revealed that *ARID1A* was the second most mutated SWI/SNF gene in LUAD with a mutation rate of 6-8% (Bailey et al. 2018; Network 2014; Imielinski et al. 2012). Moreover, our previous mutational analysis of the whole SWI/SNF complex in a different LUAD cohort (see section 4.1.2) also showed a similar mutation frequency in *ARID1A*, specifically, 8.7%. In addition, Bailey and colleagues also noted that *ARID1A* was among the top 20 LUAD driver genes, as was *SMARCA4* (Bailey et al. 2018). These high-throughput studies have described that between 60-69% of *ARID1A* alterations correspond to loss of function (LOF) mutations that include nonsense, frameshift, and splice site (Bailey et al. 2018; Hung et al. 2020b). However, the relationship between ARID1A mutations and expression levels is not trivial. Hung and colleagues observed that loss of function mutations not always correlated with loss of ARID1A expression in LUAD tumors (Hung et al. 2020b). Indeed, only 1-2% of non-small cell lung cancer (NSCLC) patients, which also include LUAD tumors, show loss of ARID1A protein expression (Naito et al. 2019; Hung et al. 2020a). This result supports our observation that, in general, LUAD primary tumors do not express significantly less ARID1A protein than their paired normal adjacent tissue. Given that three independent articles had described a dualistic role of ARID1A either as a tumor suppressor or as an oncogene in two different tumor types (Sun et al. 2017; Mathur et al. 2017; Sen et al. 2019), we evaluated the effect of silencing *ARID1A* in LUAD cell

lines that expressed this SWI/SNF subunit. Importantly, we analyzed diverse genetic backgrounds of LUAD because of the elevated context-dependency of the SWI/SNF complex and, specifically, of ARID1A.

On the one hand, we found that *ARID1A* knockdown impaired cell viability and clonogenicity in LUAD cell lines that harbored mutations in other members of the SWI/SNF complex. In particular, those were a frameshift deletion in *SMARCA4* in the A549 cell line and two missense mutations in *SMARCC2* and *SMARCD3* in the H2009 cell line. As we have previously discussed, many studies have demonstrated the dependencies that tumor cells develop towards residual SWI/SNF complexes after the alteration of one of their subunits (recently reviewed in (Wanior et al. 2021)). However, an ARID1A dependency in *SMARCA4*-mutant or *SMARCC2*/*SMARCD3*-mutant contexts has not been previously described. Intriguingly, A549 and H2009 cell lines also harbored oncogenic missense alterations in *KRAS*. In fact, *KRAS* was the only LUAD driver gene that was mutated in both cell lines. On top of that, a recent study had described that in colorectal cancer, the depletion of *ARID1A* impaired tumor growth in *KRAS*-mutant contexts (Sen et al. 2019). For that reason, we analyzed two additional LUAD cell lines that, although both of them expressed wild-type SWI/SNF complexes, one of them was *KRAS*-mutant (H358), and the other was wild-type *KRAS* (H1395). In both cases, the knockdown of *ARID1A* significantly decreased cell viability and clonogenicity. Thus, these results showed two relevant observations. First, we discovered that LUAD cells with an intact SWI/SNF complex relied on ARID1A for their survival, and therefore, it was not a result of a synthetic lethality of other members of the SWI/SNF complex. Second, these results demonstrated that, regardless of the *KRAS* status, LUAD cells showed ARID1A dependency.

On the other hand, we found that the impairment of cell viability and colony formation upon *ARID1A* knockdown was only characteristic of lung tumor contexts. We observed that NL20, a normal lung epithelial cell line, did not significantly decrease its growth after silencing *ARID1A*, as it happened in NL20-TA, the tumor cell line derived from NL20. This result is especially relevant when considering that a significant proportion of LUAD tumors does not lose ARID1A expression, and thus,

this could be an essential gene for tumor maintenance. In fact, previous studies have shown the crucial role of ARID1A in different cellular contexts as diverse as liver cancer (Sun et al. 2017), colorectal cancer (Sen et al. 2019), hematopoietic cells (Han et al. 2019), or intestinal stem cells (Hiramatsu et al. 2019). Moreover, ARID1A is not the only SWI/SNF subunit that presents a dual functionality in cancer depending on the context or the tumor type studied. For instance, although SMARCA4 is described as a tumor suppressor in many cancer types, such as lung and ovarian cancer (Chetty and Serra 2020), other researchers have found that its expression is essential for other tumors to grow (Watanabe et al. 2011; Buscarlet et al. 2014; Jubierre et al. 2016). Roy and colleagues also discovered that in pancreatic cancer, SMARCA4 played a dual role as a tumor suppressor or an oncogene depending on the tumor stage (Roy et al. 2015), as Sun and colleagues observed with ARID1A in liver cancer (Sun et al. 2017). Overall, there is increasing evidence of a characteristic pattern of alterations of the SWI/SNF complex in different tumor types, which underlines the relevant context- and subunit-specific effects of this chromatin remodeler in cancer.

Interestingly, when we studied the transcription changes produced in a cell model of LUAD after silencing *ARID1A*, we found an upregulation of c-MYC expression and enrichment of its pathway in our RNA-seq data. Initially, this result seemed to contradict the phenotype of less cell viability upon *ARID1A* knockdown, given that c-MYC is a transcription factor involved in proliferation, growth, as well as many other biosynthetic processes (reviewed in (Stine et al. 2015)). In fact, our RNA-seq data also showed enrichment of gene sets related to biosynthetic pathways, such as ribosome biogenesis and translation, two processes highly controlled by c-MYC (Dunn and Cowling 2015; van Riggelen et al. 2010; Destefanis et al. 2020). However, the involvement of c-MYC in protein synthesis is tightly connected with protein homeostasis, also known as proteostasis. Tumor cells need to carefully coordinate high translation rates to increase their biomass with their folding capacity to avoid proteotoxicity (Zhang et al. 2020). Tameire and colleagues have considered c-MYC activation as a source of intracellular stress that triggers unfolded protein response (UPR) and consequently, endoplasmic reticulum (ER) stress (Tameire et al. 2015). Importantly, we also found that the UPR pathway was significantly enriched after

silencing *ARID1A* in our LUAD model, supporting this connection between MYC activation and a proteostasis imbalance. In addition, we found a significant upregulation of *DDIT3* (also known as *CHOP* or *GADD153*) expression after *ARID1A* knockdown. This gene is one of the key participants of the so-called pro-apoptotic UPR (Rutkowski et al. 2006; McCullough et al. 2001; Yamaguchi and Wang 2004). Pro-apoptotic UPR is triggered when ER stress is chronically prolonged, and the protein load on the ER greatly exceeds its fold capacity (Sano and Reed 2013). Inside this specific cell death signaling program triggered by the UPR pathway, DDIT3 is a relevant ER stress-mediated apoptotic executor. This transcription factor is minimally expressed under physiological conditions, but it is strongly induced in chronic ER stress. When DDIT3 is overexpressed, it leads to apoptosis through the upregulation of the pro-apoptotic transcription factor ATF5 (Teske et al. 2013), whose expression was also significantly increased in our *ARID1A* knockdown model. In addition, DDIT3 induces the expression of many other pro-apoptotic factors (Nishitoh 2012). Specifically, we found the pro-apoptotic genes *BAK*, *BAX*, *BID*, and *HRK* significantly upregulated after silencing *ARID1A*. Finally, we also corroborated that there was a significant increase in apoptotic cells upon *ARID1A* knockdown. Overall, these results suggested that silencing *ARID1A* in our LUAD model modified cell proteostasis generating severe ER stress that induced apoptosis.

Interestingly, this MYC-UPR-ER stress-SWI/SNF axis has been confirmed by other authors in malignant rhabdoid tumors (MRTs) (Carugo et al. 2019). In that tumor type, they observed that their *Smarcb1*-deficient mice model showed an upregulation of c-MYC target genes, protein biosynthesis, and UPR pathways. Moreover, when they restored SMARCB1 expression, ER stress markers and MYC expression decreased, suggesting that the SWI/SNF complex tightly controlled ER stress responses through the regulation of c-MYC expression. In fact, several studies have also described this connection between MYC expression and the SWI/SNF complex in various contexts (Sims et al. 2007; Shi et al. 2013; Weissmiller et al. 2019; Nagl et al. 2007). Interestingly, Nagl and colleagues discovered that the SWI/SNF complex binds to *c-MYC* promoter, but depending on the ARID subunit that is assembled in the complex, it can either induce or repress c-MYC expression (Nagl et al. 2007). Two independent studies, which suppressed ARID1A expression in two

distinct tumor types (Wang et al. 2019; Luo et al. 2020a), also corroborated the observations of Nagl and colleagues that ARID1A had a repressor role on c-MYC expression. The results of our cell model of LUAD also support these findings, although, in this particular context, the overexpression of MYC exceeds the protein folding capacity of the cells and induces a pro-apoptotic UPR.

Furthermore, two recent publications performed in *Saccharomyces cerevisiae* have demonstrated that the SWI/SNF complex can play a direct role in regulating proteostasis by controlling expression of UPR genes either at a transcriptional (Sahu et al. 2021) or a post-transcriptional level (Sahu et al. 2020). Zundell and colleagues have confirmed this direct connection between the SWI/SNF complex and ER stress in ovarian tumors (Zundell et al. 2021). Interestingly, they found that ARID1A had a repressive function on the expression of some UPR genes by its binding to their promoters. Overall, these new studies, together with the results of our *ARID1A* KD model, underline a novel function of the SWI/SNF complex in proteostasis that may help to design new therapeutic approaches for tumors with aberrant SWI/SNF complexes. Although more research is needed in this particular field, Zundel and colleagues have shown the significant improvement of the therapy with B-109, a UPR-inhibitor, in ARID1A-deficient ovarian clear cell carcinoma.

Importantly, in our LUAD model, we observed that *ARID1A* silencing not only activated the UPR pathway but specifically its pro-apoptotic branch. As we have mentioned before, the role of DDIT3 is crucial in that process. Several studies have shown that DDIT3 overexpression acts as a positive feedback loop by enhancing protein synthesis to cause oxidative stress that ultimately leads to cell death (Han et al. 2013; Marciniak et al. 2004). Moreover, the UPR itself is a source of reactive oxygen species (ROS). Accumulating evidence demonstrates the oxidizing conditions generated by the protein folding process to favor the formation of disulfide bonds (Santos et al. 2009; Malhotra and Kaufman 2007). Thus, the generation of ROS can also be a byproduct of ER stress. On top of that, it is known that ER and oxidative stress can accentuate each other and ultimately that positive feedback interferes with cell function and activates pro-apoptosis signaling (Malhotra and Kaufman 2007), as we have observed in our model. Indeed, we also

found that the ROS pathway was among the significant gene sets after silencing *ARID1A* in our LUAD model. This connection between ROS generation and ARID1A has been considered as a synthetic lethality in some ARID1A-deficient tumors (Ogiwara et al. 2019). Due to the high proliferative and metabolic rate of tumor cells, regulating ROS homeostasis is an essential process for tumor survival. Since cancer cells rely on antioxidant pathways to eliminate the excess of ROS, therapeutic strategies that aim to target antioxidant defense systems may improve tumor elimination. Specifically, Ogiwara and colleagues discovered that ARID1A-deficient tumors presented higher ROS levels than ARID1A-proficient tumors. They found that the absence of ARID1A generated transcriptional repression of *SLC7A11*, a gene involved in the synthesis of glutathione (GSH), an essential antioxidant of the ER (Ogiwara et al. 2019). For that reason, they proposed that those ARID1A-deficient tumors, specifically ovarian, uterus, and biliary tract cancers, could benefit from the treatment with inhibitors of the GSH pathway.

One of the most damaging effects of ROS lies in their interaction with DNA. Several studies have shown that oxidative stress can induce DNA damage, causing single- or double-strand breaks, which activate apoptosis if they are not repaired (reviewed in (Kryston et al. 2011)). We observed that, upon *ARID1A* knockdown, different LUAD cell lines presented high levels of DNA damage, supporting those studies. Moreover, we found that *E2F1* was upregulated after silencing *ARID1A*. This gene code for a relevant transcription factor that is induced after DNA damage (Lin et al. 2001; Pediconi et al. 2003) and is able to trigger apoptosis (Kowalik et al. 1998; DeGregori et al. 1997; Hallstrom and Nevins 2003). Similarly, *JUN* and *ATF3*, which are two transcription factors with important functions in sensing DNA damage and activating apoptosis (Turchi et al. 2009; 2008; Lu et al. 2006; Devary et al. 1991; Kasibhatla et al. 1998; Shaulian et al. 2000), also showed significantly higher expression levels upon *ARID1A* knockdown. Overall, these results suggest that ARID1A depletion could act as a genotoxic agent for LUAD cells triggering DNA-damage apoptosis.

When we compared the effect of silencing *ARID1A* with other genotoxic agents used in cancer therapy, such as etoposide and doxorubicin, we found that depleting

ARID1A impaired cell viability at the same level as did the treatment with doxorubicin. Furthermore, knocking down *ARID1A* even showed greater cytotoxicity than the treatment with etoposide. Importantly, we also observed that combining *ARID1A* knockdown with any of these genotoxic therapies significantly improved the efficacy of both drugs in LUAD cells.

These results support previous studies that demonstrate that the loss of ARID1A or other SWI/SNF subunits increases chemosensitivity to DNA damaging agents (Watanabe et al. 2014; Agnes et al. 2006; Park et al. 2006; Ogiwara et al. 2011; Shen et al. 2015; de Castro et al. 2017). Many researchers have explained this effect because of the relevant role of the SWI/SNF complex in DNA repair. Specifically, some authors have found that ARID1A interacts with different components of the DNA repair machinery (Dykhuizen et al. 2013; Shen et al. 2018). Several studies have described that ARID1A-deficient tumors have a higher difficulty in repairing DNA damage, and consequently, present better responses to DNA damage-inducing-therapies than ARID1A-proficient tumors (Watanabe et al. 2014; Shen et al. 2015; Williamson et al. 2016; Park et al. 2019). However, we found that regardless of ARID1A expression, after the treatment with different genotoxic drugs, LUAD cells managed to decrease the levels of γ-H2AX, a characteristic marker of DNA damage. This suggested that silencing *ARID1A* did not interfere in the repair of DNA damage. Nevertheless, we observed that *ARID1A* KD drastically increased the γ-H2AX levels triggered by other genotoxic treatments. Importantly, that elevated induction of DNA damage correlated with the highest impairment of cell viability. Thus, these results led us to propose that in LUAD cells, ARID1A loss could be a source of generation of DNA damage rather than an obstacle in the DNA repair pathway. Intriguingly, this idea also agrees with the recent study of Bayona-Feliu and colleagues in which they found that the depletion of SMARCA4 or ARID1A increased DNA damage associated with replicative stress (Bayona-Feliu et al. 2021).

In short, the increase of DNA damage upon ARID1A loss could be a consequence of a proteostasis imbalance due to the transcriptional role of the SWI/SNF complex on the expression of essential genes in ER stress. However, as we have previously underlined, the SWI/SNF complex synergistically maintains cellular homeostasis

through several mechanisms, including the regulation of transcription-replication conflicts, which can lead to replicative stress and DNA damage (***Figure 34***). In any case, the depletion of ARID1A in LUAD may improve cancer therapy by enhancing the effect of other genotoxic agents and thus, allowing the reduction of their therapeutic dose to reduce their side effects. However, more studies are needed to explore this therapeutic approach for LUAD tumors.

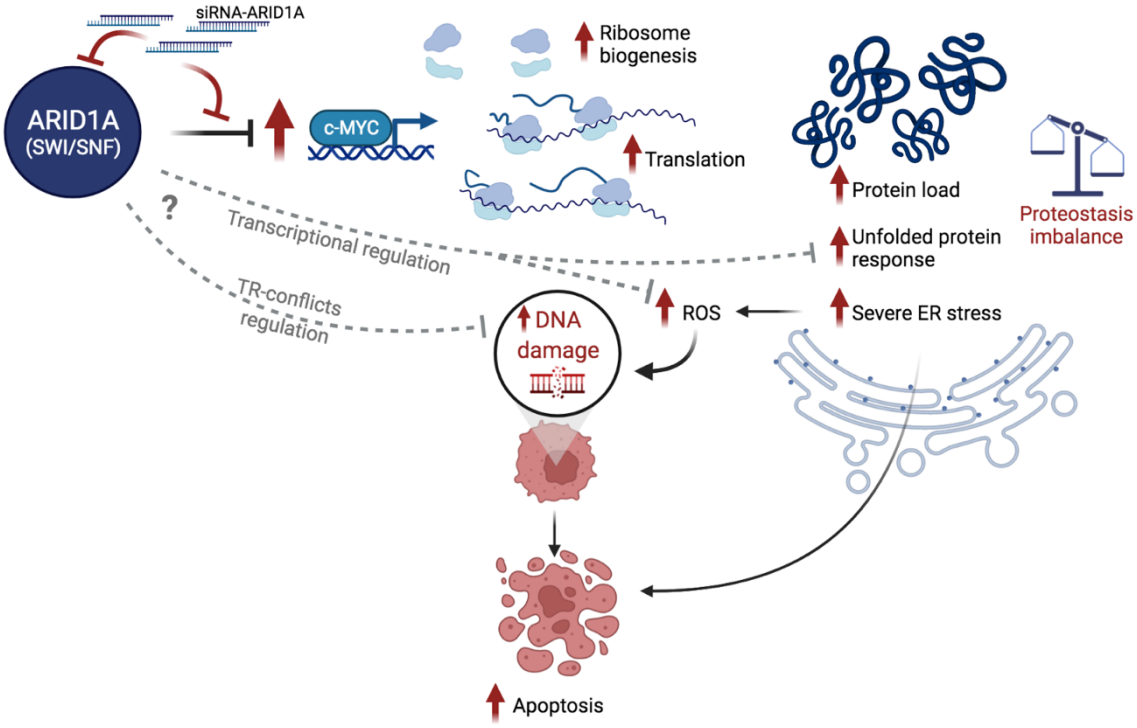


Figure 34: Schematic overview of the proposed pathways triggered by ARID1A loss in LUAD cells. After *ARID1A* knockdown, we found an increase in c-MYC expression levels and some of the biosynthetic pathways regulated by this transcription factor, such as ribosome biogenesis and translation. We also observed an upregulation of the unfolded protein response (UPR), suggesting that ARID1A loss caused proteostasis imbalance due to the increase of protein synthesis mediated by c-MYC. This imbalance could also generate ROS that explain the rise of DNA damage, which led to apoptosis. In grey, we depict alternative pathways that other authors have previously described in other tumor types. These include the direct transcriptional regulation of ARID1A of some genes involved in the UPR and ROS generation, as well as the role of ARID1A in modulating transcription-replication (TR) conflicts, which are also a source of DNA damage. Image made with BioRender.com

Lastly, we found that silencing *ARID1A* alters the expression of other subunits of the SWI/SNF complex. In particular, the two mutually exclusive subunits ARID1B and ARID2 were significantly downregulated upon *ARID1A* knockdown. Moreover, SMARCA2, the only catalytic subunit of the SWI/SNF complex in the A549 cell line

also showed lower protein levels after depleting ARID1A. Although structural studies are required to determine the composition of the residual SWI/SNF complexes in our LUAD model after *ARID1A* knockdown, a recent publication has demonstrated the relevant role of ARID1A as an essential scaffold of the SWI/SNF complex (He et al. 2020). Specifically, He and colleagues found that the assembly of the SWI/SNF complex was not properly maintained in the absence of ARID1A, although some subunits remained associated by the scaffold subunits, SMARCC1 and SMARCC2. This result agrees with the modular organization of the SWI/SNF complex previously proposed by Mashtalir et al (Mashtalir et al. 2018). These authors also observed that the depletion of ARID1A impaired the assembly of fully formed SWI/SNF complexes. Importantly, they underlined that the loss of ARID1A was not compensated by an increase of ARID1B expression, as we also observed in our LUAD model. Overall, ARID1A loss could alter the expression of other SWI/SNF subunits because of an impairment of their assembly to residual SWI/SNF complexes. This finding could demonstrate the importance of ARID1A for residual SWI/SNF complexes on which tumor cells rely, showing a potential synthetic lethality not explored yet. However, more structural analyses are required to validate this hypothesis.

In summary, our work has underlined the relevance of ARID1A in LUAD. We have discovered for the first time that some LUAD cell lines develop an essential ARID1A dependency. We have observed that ARID1A loss alters tumor homeostasis leading to an increase in DNA damage that causes cell death. Therefore, these findings present a potential therapeutic approach for LUAD tumors, which requires further research.

Conclusions

6) CONCLUSIONS

1. The study of the composition of the SWI/SNF complexes in lung epithelial cells revealed an absence of nine subunits that were found in other cell types, confirming the tissue-specific trait of this chromatin remodeler.

2. In LUAD primary tumors, the SWI/SNF complex is mutated in 41.4% of the cases, and according to *in silico* predictions, at least 70% of those mutations may affect the functionality of the complex.

3. The mutational status of the SWI/SNF complex, when considered as a functional unit, defines a clinical subgroup of LUAD patients with worse overall survival and a high tumor mutation burden.

4. There is a major downregulation of the whole SWI/SNF complex in LUAD that can only be partly attributed to mutations.

5. The characterization of a panel of 38 LUAD cell lines showed that more than 75% of them harbor mutations in lung SWI/SNF subunits and present other epigenetic alterations that change the expression levels of the different SWI/SNF complexes.

6. There are four groups of LUAD cell lines that share a specific signature regarding the SWI/SNF complex, providing different resources for researchers to choose the most suitable cellular models for their studies of this chromatin remodeler.

7. The restoration of SMARCA4 in a SMARCA4-deficient LUAD cell line not only modifies the miRNome of the cells but also changes the expression of microRNAs involved in tumorigenic pathways.

8. SMARCA4–SWI/SNF complexes allow the expression of the microRNA miR-222 through their direct binding to its enhancer region, while SMARCA2–SWI/SNF complexes have the opposite effect.

9. miR-222 is a tumor suppressor microRNA in LUAD whose upregulation decreases cell viability and impairs colony formation.

10. The miR-222 enhancer region, which we have defined as a target of the SWI/SNF complex in LUAD cells, resides in a topologically associated domain that does not contain any cancer-related protein-coding genes.

11. ARID1A plays an essential role in the survival of LUAD cell lines regardless of their genetic background, but that dependency is not observed in normal lung epithelial cells.

12. ARID1A fine-tunes the expression of relevant cellular pathways whose dysregulation unbalance cellular homeostasis, causing an increase of DNA damage that cannot be overcome by LUAD cells and triggers apoptosis.

13. Depleting ARID1A in LUAD cell lines resembles the effect of other genotoxic agents used in lung cancer therapy, such as doxorubicin and etoposide, and also improves the efficacy of such treatments.

14. ARID1A loss alters the protein levels of other SWI/SNF members, including the mutually exclusive subunits ARID1B and ARID2, as well as the ATPase-helicase subunit of the residual SWI/SNF complexes.

7) CONCLUSIONES

1. El estudio de la composición de los complejos SWI/SNF en células epiteliales de pulmón reveló la ausencia de nueve subunidades que se habían encontrado en otros tipos celulares, confirmando la especificidad de tejido característica de este complejo remodelador de la cromatina.

2. En tumores primarios de adenocarcinoma de pulmón, el complejo SWI/SNF se encuentra mutado en el 41.4% de los casos y, de acuerdo con las predicciones *in silico*, al menos el 70% de esas mutaciones podrían afectar a la funcionalidad del complejo.

3. El estado mutacional del complejo SWI/SNF, cuando se considera como una completa unidad funcional, podría definir un subgrupo clínico de pacientes de adenocarcinoma de pulmón que presenta peor supervivencia global y una mayor carga de mutación del tumor.

4. Existe una importante regulación a la baja de todo el complejo SWI/SNF en adenocarcinoma de pulmón que sólo puede atribuirse en parte a las mutaciones.

5. La caracterización de un panel de 38 líneas de adenocarcinoma de pulmón mostró que más del 75% de ellas albergan mutaciones en las subunidades pulmonares del SWI/SNF y presentan otras alteraciones epigenéticas que cambian los niveles de expresión de los diferentes complejos SWI/SNF.

6. Existen cuatro grupos de líneas celulares de adenocarcinoma de pulmón que comparten una firma específica en relación con el complejo SWI/SNF, lo que proporciona diferentes recursos a los investigadores para elegir los modelos celulares más adecuados para sus estudios de este remodelador de la cromatina.

7. La restauración de SMARCA4 en una línea de adenocarcinoma de pulmón deficiente en esta proteína no solo modifica el miRNoma de las células, sino que cambia la expresión de microARNs involucrados en rutas tumorigénicas.

8. Los complejos SWI/SNF con SMARCA4 permiten la expresión del microARN miR-222 a través de su unión directa a su región potenciadora, mientras que los complejos SWI/SNF con SMARCA2 tienen el efecto opuesto.

9. miR-222 es un microARNs supresor de tumores en adenocarcinoma de pulmón cuya regulación al alza disminuye la viabilidad celular e impide la formación de colonias.

10. La región potenciadora del miR-222, la cual hemos definido como diana del complejo SWI/SNF en células adenocarcinoma de pulmón, se encuentra en un dominio de asociación topológico que no contiene ningún gen codificante de proteína relacionado con cáncer.

11. ARID1A desempeña un papel esencial en la supervivencia de líneas celulares de adenocarcinoma de pulmón independientemente de su fondo genético, pero esa dependencia no se observa en células normales del epitelio pulmonar.

12. ARID1A ajusta la expresión de importantes rutas celulares cuya desregulación desequilibra la homeostasis celular provocando un aumento en el daño en el ADN que no puede ser superado por las células de adenocarcinoma de pulmón y activa la apoptosis.

13. Eliminar ARID1A en líneas celulares de adenocarcinoma de pulmón se asemeja al efecto que causan otros agentes genotóxicos usados en la terapia del cáncer de pulmón, tales como doxorubicina y etopósido, y también mejora la eficacia de dichos tratamientos.

14. La pérdida de ARID1A altera los niveles proteicos de otros miembros del SWI/SNF, incluyendo las subunidades mutuamente exclusivas ARID1B y ARID2, así como la subunidad ATPasa-helicasa de los complejos SWI/SNF residuales.

References

REFERENCES

Abrams E, Neigeborn L, Carlson M. 1986. Molecular analysis of SNF2 and SNF5, genes required for expression of glucose-repressible genes in Saccharomyces cerevisiae. *Molecular and cellular biology* **6**: 3643–3651.

Agnes K-Y, Eli P, Moshe Y. 2006. Increased DNA Damage Sensitivity and Apoptosis in Cells Lacking the Snf5/Ini1 Subunit of the SWI/SNF Chromatin Remodeling Complex. *Molecular and cellular biology* **26**: 2661–2674.

Aizawa H, Hu S-C, Bobb K, Balakrishnan K, Ince G, Gurevich I, Cowan M, Ghosh A. 2004. Dendrite development regulated by CREST, a calcium-regulated transcriptional activator. *Science* **303**: 197–202.

Alpsoy A, Dykhuizen EC. 2018. Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *Journal of Biological Chemistry* **293**: 3892–3903.

Alver BH, Kim KH, Lu P, Wang X, Manchester HE, Wang W, Haswell JR, Park PJ, Roberts CWM. 2017. The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nature communications* **8**: 14648.

Amini S, Abak A, Sakhinia E, Abhari A. 2019. MicroRNA-221 and MicroRNA-222 in Common Human Cancers: Expression, Function, and Triggering of Tumor Progression as a Key Modulator. *Lab Med* **50**: 333–347.

Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, Ji HP, Maley CC. 2016. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med* **22**: 105–113.

Anglesio MS, Wang Y, Yang W, Senz J, Wan A, Heravi-Moussavi A, Salamanca C, Maines-Bandiera S, Huntsman DG, Morin GB. 2013. Cancer-associated somatic DICER1 hotspot mutations cause defective miRNA processing and reverse-strand expression bias to predominantly mature 3p strands through loss of 5p strand cleavage. *J Pathol* **229**: 400–409.

Arts FA, Keogh L, Smyth P, O'Toole S, Ta R, Gleeson N, O'Leary JJ, Flavin R, Sheils O. 2017. miR-223 potentially targets SWI/SNF complex protein SMARCD1 in atypical proliferative serous tumor and high-grade ovarian serous carcinoma. *Hum Pathol* **70**: 98–104.

Asenjo HG, Gallardo A, López-Onieva L, Tejada I, Martorell-Marugán J, Carmona-Sáez P, Landeira D. 2020. Polycomb regulation is coupled to cell cycle transition in pluripotent stem cells. *Sci Adv* **6**: eaay4768.

Bai J, Mei P, Zhang C, Chen F, Li C, Pan Z, Liu H, Zheng J. 2013. BRG1 is a prognostic marker and potential therapeutic target in human breast cancer. *PLOS one* **8**: e59772.

Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, Colaprico A, Wendl MC, Kim J, Reardon B, et al. 2018. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **174**: 1034–1035.

Banine F, Bartlett C, Gunawardena R, Muchardt C, Yaniv M, Knudsen ES, Weissman BE, Sherman LS. 2005. SWI/SNF chromatin-remodeling factors induce changes in DNA methylation to promote transcriptional activation. *Cancer Research* **65**: 3542–3547.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297.

Batsché E, Yaniv M, Muchardt C. 2006. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13**: 22–29.

Baylin SB, Jones PA. 2011. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer* **11**: 726–734.

Baylin SB, Jones PA. 2016. Epigenetic Determinants of Cancer. *Cold Spring Harb Perspect Biol* **8**: a019505.

Bayona-Feliu A, Barroso S, Muñoz S, Aguilera A. 2021. The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription-replication conflicts. *Nat Genet* **53**: 1050–1063.

Beck S, Bernstein BE, Campbell RM, Costello JF, Dhanak D, Ecker JR, Greally JM, Issa J-P, Laird PW, Polyak K, et al. 2012. A blueprint for an international cancer epigenome consortium. A report from the AACR Cancer Epigenome Task Force. *Cancer Research* **72**: 6319–6324.

Becker PB, Workman JL. 2013. Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol* **5**: a017905.

Belandia B, Orford RL, Hurst HC, Parker MG. 2002. Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. *EMBO J* **21**: 4094–4103.

Beringer M, Pisano P, Di Carlo V, Blanco E, Chammas P, Vizán P, Gutiérrez A, Aranda S, Payer B, Wierer M, et al. 2016. EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells. *Molecular and cellular* **64**: 645–658.

Betz BL, Strobeck MW, Reisman DN, Knudsen ES, Weissman BE. 2002. Re-expression of hSNF5/INI1/BAF47 in pediatric tumor cells leads to G1 arrest associated with induction of p16ink4a and activation of RB. *Oncogene* **21**: 5193–5203.

Béguelin W, Popovic R, Teater M, Jiang Y, Bunting KL, Rosen M, Shen H, Yang SN, Wang L, Ezponda T, et al. 2013. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer cell* **23**: 677–692.

Birnbaum DJ, Adélaïde J, Mamessier E, Finetti P, Lagarde A, Monges G, Viret F, Gonçalvès A, Turrini O, Delpero J-R, et al. 2011. Genome profiling of pancreatic adenocarcinoma. *Genes, chromosomes & cancer* **50**: 456–465.

Biswas AK, Johnson DG. 2012. Transcriptional and nontranscriptional functions of E2F1 in response to DNA damage. *Cancer Research* **72**: 13–17.

Blanco R, Iwakawa R, Tang M, Kohno T, Angulo B, Pio R, Montuenga LM, Minna JD, Yokota J, Sanchez-Cespedes M. 2009. A gene-alteration profile of human lung cancer cell lines. *Human mutation* **30**: 1199–1206.

Blattner C, Sparks A, Lane D. 1999. Transcription factor E2F-1 is upregulated in response to DNA damage in a manner analogous to that of p53. *Molecular and cellular biology* **19**: 3704–3713.

Bock VL, Lyons JG, Huang XXJ, Jones AM, McDonald LA, Scolyer RA, Moloney FJ, Barnetson RS, Halliday GM. 2011. BRM and BRG1 subunits of the SWI/SNF chromatin remodelling complex are downregulated upon progression of benign skin lesions into invasive tumours. *Br J Dermatol* **164**: 1221–1227.

Bossen C, Murre CS, Chang AN, Mansson R, Rodewald H-R, Murre C. 2015. The chromatin remodeler Brg1 activates enhancer repertoires to establish B cell identity and modulate cell growth. *Nat Immunol* **16**: 775–784.

Bozic I, Reiter JG, Allen B, Antal T, Chatterjee K, Shah P, Moon YS, Yaqubie A, Kelly N, Le DT, et al. 2013. Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife* **2**: e00747.

Bögershausen N, Wollnik B. 2018. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front Mol Neurosci* **11**: 252.

Bracken AP, Brien GL, Verrijzer CP. 2019. Dangerous liaisons: interplay between SWI/SNF, NuRD, and Polycomb in chromatin regulation and cancer. *Genes Dev* **33**: 936–959.

Braun DA, Ishii Y, Walsh AM, Van Allen EM, Wu CJ, Shukla SA, Choueiri TK. 2019. Clinical Validation of PBRM1 Alterations as a Marker of Immune Checkpoint Inhibitor Response in Renal Cell Carcinoma. *JAMA Oncol* **5**: 1631–1633.

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**: 394–424.

Breeden L, Nasmyth K. 1987. Cell cycle control of the yeast HO gene: cis- and trans-acting regulators. *Cell* **48**: 389–397.

Brownlee PM, Chambers AL, Cloney R, Bianchi A, Downs JA. 2014. BAF180 promotes cohesion and prevents genome instability and aneuploidy. *CellReports* **6**: 973–981.

Bultman S, Gebuhr T, Yee D, La Mantia C, Nicholson J, Gilliam A, Randazzo F, Metzger D, Chambon P, Crabtree G, et al. 2000. A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Molecular and cellular* **6**: 1287–1295.

Bultman SJ, Bultman SJ, Herschkowitz JI, Herschkowitz JI, Godfrey V, Godfrey V, Gebuhr TC, Gebuhr TC, Yaniv M, Yaniv M, et al. 2008. Characterization of mammary tumors from Brg1 heterozygous mice. *Oncogene* **27**: 460–468.

Buscarlet M, Krasteva V, Ho L, Simon C, Hébert J, Wilhelm B, Crabtree GR, Sauvageau G, Thibault P, Lessard JA. 2014. Essential role of BRG, the ATPase subunit of BAF chromatin remodeling complexes, in leukemia maintenance. *Blood* **123**: 1720–1728.

Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, Shukla SA, Guo G, Brooks AN, Murray BA, et al. 2016. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* **48**: 607–616.

Carugo A, Minelli R, Sapio L, Soeung M, Carbone F, Robinson FS, Tepper J, Chen Z, Lovisa S, Svelto M, et al. 2019. p53 Is a Master Regulator of Proteostasis in SMARCB1-Deficient Malignant Rhabdoid Tumors. *Cancer cell* **35**: 204–220.e9.

Centore RC, Sandoval GJ, Soares LMM, Kadoch C, Chan HM. 2020. Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies. *Trends in Genetics* **36**: 936–950.

Chandler RL, Brennan J, Schisler JC, Serber D, Patterson C, Magnuson T. 2013. ARID1a-DNA interactions are required for promoter occupancy by SWI/SNF. *Molecular and cellular biology* **33**: 265–280.

Chang L, Azzolin L, Di Biagio D, Zanconato F, Battilana G, Lucon Xiccato R, Aragona M, Giulitti S, Panciera T, Gandin A, et al. 2018. The SWI/SNF complex is a mechanoregulated inhibitor of YAP and TAZ. *Nature* **563**: 265–269.

Chetty R, Serra S. 2020. SMARCA family of genes. *J Clin Pathol* **73**: 257–260.

Chiarugi A, Moskowitz MA. 2002. Cell biology. PARP-1--a perpetrator of apoptotic cell death? *Science* **297**: 200–201.

Cho H, Kim JS-Y, Chung H, Perry C, Lee H, Kim J-H. 2013. Loss of ARID1A/BAF250a expression is linked to tumor progression and adverse prognosis in cervical cancer. *Hum Pathol* **44**: 1365–1374.

Chodavarapu RK, Feng S, Bernatavichute YV, Chen P-Y, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, et al. 2010. Relationship between nucleosome positioning and DNA methylation. *Nature* **466**: 388–392.

Chun H-JE, Lim EL, Heravi-Moussavi A, Saberi S, Mungall KL, Bilenky M, Carles A, Tse K, Shlafman I, Zhu K, et al. 2016. Genome-Wide Profiles of Extra-cranial

Malignant Rhabdoid Tumors Reveal Heterogeneity and Dysregulated Developmental Pathways. *Cancer cell* **29**: 394–406.

Clapier CR, Iwasa J, Cairns BR, Peterson CL. 2017. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**: 407–422.

Clark J, Rocques PJ, Crew AJ, Gill S, Shipley J, Chan AM, Gusterson BA, Cooper CS. 1994. Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. *Nat Genet* **7**: 502–508.

Cohen SM, Chastain PD2, Rosson GB, Groh BS, Weissman BE, Kaufman DG, Bultman SJ. 2010. BRG1 co-localizes with DNA replication factors and is required for efficient replication fork progression. *Nucleic Acids Research* **38**: 6906–6919.

Coira IF, Rufino-Palomares EE, Romero OA, Peinado P, Metheetrairut C, Boyero-Corral L, Carretero J, Farez-Vidal E, Cuadros M, Reyes-Zurita FJ, et al. 2015. Expression inactivation of SMARCA4 by microRNAs in lung tumors. *Human molecular genetics* **24**: 1400–1409.

Collings CK, Anderson JN. 2017. Links between DNA methylation and nucleosome occupancy in the human genome. *Epigenetics Chromatin* **10**: 18.

Collings CK, Waddell PJ, Anderson JN. 2013. Effects of DNA methylation on nucleosome stability. *Nucleic Acids Research* **41**: 2918–2931.

Cornen S, Adelaide J, Bertucci F, Finetti P, Guille A, Birnbaum DJ, Birnbaum D, Chaffanet M. 2012. Mutations and deletions of ARID1A in breast tumors. *Oncogene* **31**: 4255–4256.

Croce CM. 2009. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* **10**: 704–714.

Dagogo-Jack I, Schrock AB, Kem M, Jessop N, Lee J, Ali SM, Ross JS, Lennerz JK, Shaw AT, Mino-Kenudson M. 2020. Clinicopathologic Characteristics of BRG1-Deficient NSCLC. *J Thorac Oncol* **15**: 766–776.

Dallas PB, Pacchione S, Wilsker D, Bowrin V, Kobayashi R, Moran E. 2000. The human SWI-SNF complex protein p270 is an ARID family member with non-sequence-specific DNA binding activity. *Molecular and cellular biology* **20**: 3137–3146.

Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, Hilton JA, Jain K, Baymuradov UK, Narayanan AK, et al. 2018. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Research* **46**: D794–D801.

de Castro RO, Previato L, Goitea V, Felberg A, Guiraldelli MF, Filiberti A, Pezza RJ. 2017. The chromatin-remodeling subunit Baf200 promotes homology-directed DNA repair and regulates distinct chromatin-remodeling complexes. *The Journal of biological chemistry* **292**: 8459–8471.

Dedes KJ, Natrajan R, Lambros MB, Geyer FC, Lopez-Garcia MA, Savage K, Jones RL, Reis-Filho JS. 2011. Down-regulation of the miRNA master regulators Drosha and Dicer is associated with specific subgroups of breast cancer. *European Journal of Cancer* **47**: 138–150.

DeGregori J, Leone G, Miron A, Jakoi L, Nevins JR. 1997. Distinct roles for E2F proteins in cell growth control and apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 7245–7250.

Destefanis F, Manara V, Bellosta P. 2020. Myc as a Regulator of Ribosome Biogenesis and Cell Competition: A Link to Cancer. *Int J Mol Sci* **21**.

Devary Y, Gottlieb RA, Lau LF, Karin M. 1991. Rapid and Preferential Activation of the c-jun Gene during the Mammalian UV Response. *Molecular and cellular biology* **11**: 2804–2811.

Dharap A, Pokrzywa C, Murali S, Pandi G, Vemuganti R. 2013. MicroRNA miR-324-3p induces promoter-mediated expression of RelA gene. *PLOS one* **8**: e79467.

Dizdaroglu M, Jaruga P, Birincioglu M, Rodriguez H. 2002. Free radical-induced damage to DNA: mechanisms and measurement1, 2 1This article is part of a series of reviews on "Oxidative DNA Damage and Repair." The full list of papers may be found on the homepage of the journal. 2Guest Editor: Miral Dizdaroglu. *Free Radical Biology and Medicine* **32**: 1102–1115.

Doan DN, Veal TM, Yan Z, Wang W, Jones SN, Imbalzano AN. 2004. Loss of the INI1 tumor suppressor does not impair the expression of multiple BRG1-dependent genes or the assembly of SWI/SNF enzymes. *Oncogene* **23**: 3462–3473.

Dunaief JL, Strober BE, Guha S, Khavari PA, Alin K, Luban J, Begemann M, Crabtree GR, Goff SP. 1994. The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell* **79**: 119–130.

Dunn S, Cowling VH. 2015. Myc and mRNA capping. *Biochim Biophys Acta* **1849**: 501–505.

Dykhuizen EC, Hargreaves DC, Miller EL, Cui K, Korshunov A, Kool M, Pfister S, Cho Y-J, Zhao K, Crabtree GR. 2013. BAF complexes facilitate decatenation of DNA by topoisomerase IIα. *Nature* **497**: 624–627.

Eder T, Hess AK, Konschak R, Stromberger C, Jöhrens K, Fleischer V, Hummel M, Balermpas P, Grün von der J, Linge A, et al. 2019. Interference of tumour mutational burden with outcome of patients with head and neck cancer treated with definitive chemoradiation: a multicentre retrospective study of the German Cancer Consortium Radiation Oncology Group. *European Journal of Cancer* **116**: 67–76.

Endo M, Yasui K, Zen Y, Gen Y, Zen K, Tsuji K, Dohi O, Mitsuyoshi H, Tanaka S, Taniwaki M, et al. 2013. Alterations of the SWI/SNF chromatin remodelling subunit-BRG1 and BRM in hepatocellular carcinoma. *Liver Int* **33**: 105–117.

Esteller M. 2008. Epigenetics in cancer. *N Engl J Med* **358**: 1148–1159.

Euskirchen GM, Auerbach RK, Davidov E, Gianoulis TA, Zhong G, Rozowsky J, Bhardwaj N, Gerstein MB, Snyder M. 2011. Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. *PLoS Genet* **7**: e1002008.

Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**: 3047–3048.

Fan F, Jin S, Amundson SA, Tong T, Fan W, Zhao H, Zhu X, Mazzacurati L, Li X, Petrik KL, et al. 2002. ATF3 induction following DNA damage is regulated by distinct signaling pathways and over-expression of ATF3 protein suppresses cells growth. *Oncogene* **21**: 7488–7496.

Feinberg AP, Koldobskiy MA, Göndör A. 2016. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet* **17**: 284–299.

Felsenfeld G. 1992. Chromatin as an essential part of the transcriptional mechanim. *Nature* **355**: 219–224.

Feng J, Xu X, Fan X, Yi Q, Tang L. 2021. BAF57/SMARCE1 Interacting with Splicing Factor SRSF1 Regulates Mechanical Stress-Induced Alternative Splicing of Cyclin D1. *Genes (Basel)* **12**: 306.

Fernando TM, Piskol R, Bainer R, Sokol ES, Trabucco SE, Zhang Q, Trinh H, Maund S, Kschonsak M, Chaudhuri S, et al. 2020. Functional characterization of SMARCA4 variants identified by targeted exome-sequencing of 131,668 cancer patients. *Nature communications* **11**: 5551.

Ferrari E, Gandellini P. 2020. Unveiling the ups and downs of miR-205 in physiology and cancer: transcriptional and post-transcriptional mechanisms. *Cell Death and Disease* **11**: 980.

Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, Rosen N, Kohn A, Twik M, Safran M, et al. 2017. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)* **2017**: bax028.

Flavahan WA, Gaskell E, Bernstein BE. 2017. Epigenetic plasticity and the hallmarks of cancer. *Science* **357**: eaal2380.

Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, et al. 2005. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* **37**: 391–400.

Friedländer MR, Lizano E, Houben AJ, Bezdan D, Báñez-Coronel M, Kudla G, Mateu-Huertas E, Kagerbauer B, González J, Chen KC, et al. 2014. Evidence for the biogenesis of more than 1,000 novel human microRNAs. *Genome Biol* **15**: R57.

Friedman JM, Liang G, Liu C-C, Wolff EM, Tsai YC, Ye W, Zhou X, Jones PA. 2009. The putative tumor suppressor microRNA-101 modulates the cancer epigenome by repressing the polycomb group protein EZH2. *Cancer Research* **69**: 2623–2629.

Fu X, Li Y, Alvero A, Li J, Wu Q, Xiao Q, Peng Y, Hu Y, Li X, Yan W, et al. 2016. MicroRNA-222-3p/GNAI2/AKT axis inhibits epithelial ovarian cancer cell growth and associates with good overall survival. *Oncotarget* **7**: 80633–80654.

Fukuoka J, Fujii T, Shih JH, Dracheva T, Meerzaman D, Player A, Hong K, Settnek S, Gupta A, Buetow K, et al. 2004. Chromatin remodeling factors and BRM/BRG1 expression as prognostic indicators in non-small cell lung cancer. *Clin Cancer Res* **10**: 4314–4324.

García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A. 2012. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**: 2678–2679.

Garofalo M, Di Leva G, Romano G, Nuovo G, Suh S-S, Ngankeu A, Taccioli C, Pichiorri F, Alder H, Secchiero P, et al. 2009. miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation. *Cancer cell* **16**: 498–509.

Gibson WJ, Hoivik EA, Halle MK, Taylor-Weiner A, Cherniack AD, Berg A, Holst F, Zack TI, Werner HMJ, Staby KM, et al. 2016. The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis. *Nat Genet* **48**: 848–855.

Gillette MA, Satpathy S, Cao S, Dhanasekaran SM, Vasaikar SV, Krug K, Petralia F, Li Y, Liang W-W, Reva B, et al. 2020. Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma. *Cell* **182**: 200–225.e35.

Glaros S, Cirrincione GM, Muchardt C, Kleer CG, Michael CW, Reisman D. 2007. The reversible epigenetic silencing of BRM: implications for clinical targeted therapy. *Oncogene* **26**: 7058–7066.

Glaros S, Cirrincione GM, Palanca A, Metzger D, Reisman D. 2008. Targeted knockout of BRG1 potentiates lung cancer development. *Cancer Research* **68**: 3689–3696.

Gong F, Fahy D, Smerdon MJ. 2006. Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair. *Nat Struct Mol Biol* **13**: 902–907.

Grossi I, Salvi A, Baiocchi G, Portolani N, De Petro G. 2018. Functional Role of microRNA-23b-3p in Cancer Biology. *Microrna* **7**: 156–166.

Grunstein M. 1990. Histone Function in Transcription. *Annu Rev Cell Biol* **6**: 643–676.

Gu Y-F, Cohn S, Christie A, McKenzie T, Wolff N, Do QN, Madhuranthakam AJ, Pedrosa I, Wang T, Dey A, et al. 2017. Modeling Renal Cell Carcinoma in Mice: Bap1 and Pbrm1 Inactivation Drive Tumor Grade. *Cancer Discov* **7**: 900–917.

Guan B, Wang T-L, Shih I-M. 2011. ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers. *Cancer Research* **71**: 6718–6727.

Gupta M, Concepcion CP, Fahey CG, Keshishian H, Bhutkar A, Brainson CF, Sanchez-Rivera FJ, Pessina P, Kim JY, Simoneau A, et al. 2020. BRG1 Loss Predisposes Lung Cancers to Replicative Stress and ATR Dependency. *Cancer Research* **80**: 3841–3854.

Hallstrom TC, Nevins JR. 2003. Specificity in the activation and control of transcription factor E2F-dependent apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 10848–10853.

Han J, Back SH, Hur J, Lin Y-H, Gildersleeve R, Shan J, Yuan CL, Krokowski D, Wang S, Hatzoglou M, et al. 2013. ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. *Nat Cell Biol* **15**: 481–490.

Han L, Madan V, Mayakonda A, Dakle P, Woon TW, Shyamsunder P, Nordin HBM, Cao Z, Sundaresan J, Lei I, et al. 2019. Chromatin remodeling mediated by ARID1A is indispensable for normal hematopoiesis in mice. *Leukemia* **33**: 2291–2305.

Han X, Chen W, Chen P, Zhou W, Rong Y, Lv Y, Li J-A, Ji Y, Chen W, Lou W, et al. 2020. Aberration of ARID1A Is Associated With the Tumorigenesis and Prognosis of Sporadic Nonfunctional Pancreatic Neuroendocrine Tumors. *Pancreas* **49**: 514–523.

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell* **144**: 646–674.

Hargreaves DC, Crabtree GR. 2011. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res* **21**: 396–420.

Harrod A, Lane KA, Downs JA. 2020. The role of the SWI/SNF chromatin remodelling complex in the response to DNA double strand breaks. *DNA Repair* **93**: 102919.

He S, Wu Z, Tian Y, Yu Z, Yu J, Wang X, Li J, Liu B, Xu Y. 2020. Structure of nucleosome-bound human BAF complex. *Science* **367**: 875–881.

He X, Xu C. 2020. Immune checkpoint signaling and cancer immunotherapy. *Cell Res* **30**: 660–669.

Heery DM, Kalkhoven E, Hoare S, Parker MG. 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**: 733–736.

Helming KC, Wang X, Roberts CWM. 2014a. Vulnerabilities of mutant SWI/SNF complexes in cancer. *Cancer cell* **26**: 309–317.

Helming KC, Wang X, Wilson BG, Vazquez F, Haswell JR, Manchester HE, Kim Y, Kryukov GV, Ghandi M, Aguirre AJ, et al. 2014b. ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat Med* **20**: 251–254.

Heravi-Moussavi A, Anglesio MS, Cheng SWG, Senz J, Yang W, Prentice L, Fejes AP, Chow C, Tone A, Kalloger SE, et al. 2011. Recurrent Somatic DICER1 Mutations in Nonepithelial Ovarian Cancers. *N Engl J Med* **366**: 234–242.

Herpel E, Rieker RJ, Dienemann H, Muley T, Meister M, Hartmann A, Warth A, Agaimy A. 2017. SMARCA4 and SMARCA2 deficiency in non-small cell lung cancer: immunohistochemical survey of 316 consecutive specimens. *Annals of Diagnostic Pathology* **26**: 47–51.

Hiramatsu Y, Fukuda A, Ogawa S, Goto N, Ikuta K, Tsuda M, Matsumoto Y, Kimura Y, Yoshioka T, Takada Y, et al. 2019. Arid1a is essential for intestinal stem cells through Sox9 regulation. *Proceedings of the National Academy of Sciences of the United States of America* **116**: 1704–1713.

Hirschhorn JN, Brown SA, Clark CD, Winston F. 1992. Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes Dev* **6**: 2288–2298.

Ho L, Jothi R, Ronan JL, Cui K, Zhao K, Crabtree GR. 2009. An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *PNAS* **106**: 5187–5191.

Hoadley KA, Hinoue T, Wolf DM, Lazar AJ, Drill E, Shen R, Taylor AM, Cherniack AD, Thorsson V, Akbani R, et al. 2018. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**: 291–304.e6.

Hodges C, Kirkland JG, Crabtree GR. 2016. The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. *Cold Spring Harb Perspect Med* **6**: a026930–25.

Hodges HC, Stanton BZ, Cermakova K, Chang C-Y, Miller EL, Kirkland JG, Ku WL, Veverka V, Zhao K, Crabtree GR. 2018. Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. *Nat Struct Mol Biol* **25**: 61–72.

Hodis E, Watson IR, Kryukov GV, Arold ST, Imielinski M, Theurillat J-P, Nickerson E, Auclair D, Li L, Place C, et al. 2012. A landscape of driver mutations in melanoma. *Cell* **150**: 251–263.

Hoffman GR, Rahal R, Buxton F, Xiang K, McAllister G, Frias E, Bagdasarian L, Huber J, Lindeman A, Chen D, et al. 2014. Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *PNAS* **111**: 3128–3133.

Hu G, Schones DE, Cui K, Ybarra R, Northrup D, Tang Q, Gattinoni L, Restifo NP, Huang S, Zhao K. 2011. Regulation of nucleosome landscape and transcription

factor targeting at tissue-specific enhancers by BRG1. *Genome Res* **21**: 1650–1658.

Hung YP, Redig A, Hornick JL, Sholl LM. 2020a. ARID1A mutations and expression loss in non-small cell lung carcinomas: clinicopathologic and molecular analysis. *Mod Pathol* **33**: 2256–2268.

Hung YP, Redig A, Hornick JL, Sholl LM. 2020b. ARID1A mutations and expression loss in non-small cell lung carcinomas: clinicopathologic and molecular analysis. *Mod Pathol* **33**: 2256–2268.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* **12**: 99–110.

Hwang WL, Wolfson RL, Niemierko A, Marcus KJ, DuBois SG, Haas-Kogan D. 2019. Clinical Impact of Tumor Mutational Burden in Neuroblastoma. *J Natl Cancer Inst* **111**: 695–699.

Imbalzano AN, Kwon H, Green MR, Kingston RE. 1994. Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature* **370**: 481–485.

Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, Cho J, Suh J, Capelletti M, Sivachenko A, et al. 2012. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**: 1107–1120.

Inoue H, Furukawa T, Giannakopoulos S, Zhou S, King DS, Tanese N. 2002. Largest subunits of the human SWI/SNF chromatin-remodeling complex promote transcriptional activation by steroid hormone receptors. *The Journal of biological chemistry* **277**: 41674–41685.

Iurlaro M, Stadler MB, Masoni F, Jagani Z, Galli GG, beler DSX. 2021. Mammalian SWI/SNF continuously restores local accessibility to chromatin. *Nat Genet* **53**: 279-287.

Jang S-H, Lee J-H, Lee HJ, Cho H, Ahn H, Song IH, Oh M-H. 2020. Loss of ARID1A expression is associated with poor prognosis in non-small cell lung cancer. *Pathol Res Pract* **216**: 153156.

Januario T, Ye X, Bainer R, Alicke B, Smith T, Haley B, Modrusan Z, Gould S, Yauch RL. 2017. PRC2-mediated repression of SMARCA2 predicts EZH2 inhibitor activity in SWI/SNF mutant tumors. *Proceedings of the National Academy of Sciences of the United States of America* **114**: 12249–12254.

Jardim DL, Goodman A, de Melo Gagliato D, Kurzrock R. 2021. The Challenges of Tumor Mutational Burden as an Immunotherapy Biomarker. *Cancer cell* **39**: 154–173.

Jelinic P, Mueller JJ, Olvera N, Dao F, Scott SN, Shah R, Gao J, Schultz N, Gonen M, Soslow RA, et al. 2014. Recurrent SMARCA4 mutations in small cell carcinoma of the ovary. *Nat Genet* **46**: 424–426.

Jiang C, Pugh BF. 2009. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* **10**: 161–172.

Jiang Z-H, Dong X-W, Shen Y-C, Qian H-L, Yan M, Yu Z-H, He H-B, Lu C-D, Qiu F. 2015. DNA damage regulates ARID1A stability via SCF ubiquitin ligase in gastric cancer cells. *Eur Rev Med Pharmacol Sci* **19**: 3194–3200.

Jones PA, Laird PW. 1999. Cancer epigenetics comes of age. *Nat Genet* **21**: 163–167.

Jones S, Wang T-L, Shih I-M, Mao T-L, Nakayama K, Roden R, Glas R, Slamon D, Diaz LAJ, Vogelstein B, et al. 2010. Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science* **330**: 228–231.

Jubierre L, Soriano A, Planells-Ferrer L, París-Coderch L, Tenbaum SP, Romero OA, Moubarak RS, Almazán-Moga A, Molist C, Roma J, et al. 2016. BRG1/SMARCA4 is essential for neuroblastoma cell viability through modulation of cell death and survival pathways. *Oncogene* **35**: 5179–5190.

Kadam S, Emerson BM. 2003. Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes. *Molecular and cellular* **11**: 377–389.

Kadoch C, Crabtree GR. 2015. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv* **1**: e1500447–18.

Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J, Crabtree GR. 2013. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* **45**: 592–601.

Kadoch C, Williams RT, Calarco JP, Miller EL, Weber CM, Braun SMG, Pulice JL, Chory EJ, Crabtree GR. 2017. Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nat Genet* **49**: 213–222.

Kaeser MD, Aslanian A, Dong M-Q, Yates 3rd JR, Emerson BM. 2008. BRD7, a novel PBAF-specific SWI/SNF subunit, is required for target gene activation and repression in embryonic stem cells. *Journal of Biological Chemistry* **283**: 32254–32263.

Kahali B, Yu J, Marquez SB, Thompson KW, Liang SY, Lu L, Reisman D. 2014. The silencing of the SWI/SNF subunit and anticancer gene BRM in Rhabdoid tumors. *Oncotarget* **5**: 3316–3332.

Kakarougkas A, Ismail A, Chambers AL, Riballo E, Herbert AD, Künzel J, Löbrich M, Jeggo PA, Downs JA. 2014. Requirement for PBAF in transcriptional repression and repair at DNA breaks in actively transcribed regions of chromatin. *Molecular and cellular* **55**: 723–732.

Kanai Y, Ushijima S, Nakanishi Y, Sakamoto M, Hirohashi S. 2003. Mutation of the DNA methyltransferase (DNMT) 1 gene in human colorectal cancers. *Cancer Letters* **192**: 75–82.

Kang H, Cui K, Zhao K. 2004. BRG1 controls the activity of the retinoblastoma protein via regulation of p21CIP1/WAF1/SDI. *Molecular and cellular biology* **24**: 1188–1199.

Karachaliou N, Paulina Bracht JW, Rosell R. 2018. ARID1A Gene Driver Mutations in Lung Adenocarcinomas. *J Thorac Oncol* **13**: e255–e257.

Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G, et al. 2018. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Research* **46**: D239–D245.

Karnezis AN, Wang Y, Ramos P, Hendricks WP, Oliva E, D'Angelo E, Prat J, Nucci MR, Nielsen TO, Chow C, et al. 2016. Dual loss of the SWI/SNF complex ATPases SMARCA4/BRG1 and SMARCA2/BRM is highly sensitive and specific for small cell carcinoma of the ovary, hypercalcaemic type. *J Pathol* **238**: 389–400.

Karube Y, Tanaka H, Osada H, Tomida S, Tatematsu Y, Yanagisawa K, Yatabe Y, Takamizawa J, Miyoshi S, Mitsudomi T, et al. 2005. Reduced expression of Dicer associated with poor prognosis in lung cancer patients. *Cancer Sci* **96**: 111–115.

Kasibhatla S, Brunner T, Genestier L, Echeverri F, Mahboubi A, Green DR. 1998. DNA Damaging Agents Induce Expression of Fas Ligand and Subsequent Apoptosis in T Lymphocytes via the Activation of NF-κB and AP-1. *Molecular and cellular* **1**: 543–551.

Kasinski AL, Slack FJ. 2011. Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. *Nat Rev Cancer* **11**: 849–864.

Kato H, Tjernberg A, Zhang W, Krutchinsky AN, An W, Takeuchi T, Ohtsuki Y, Sugano S, de Bruijn DR, Chait BT, et al. 2002. SYT associates with human SNF/SWI complexes and the C-terminal region of its fusion partner SSX1 targets histones. *The Journal of biological chemistry* **277**: 5498–5505.

Kelso TWR, Porter DK, Amaral ML, Shokhirev MN, Benner C, Hargreaves DC. 2017. Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. *eLife* **6**: e30506.

Khare SP, Habib F, Sharma R, Gadewal N, Gupta S, Galande S. 2012. HIstome--a relational knowledgebase of human histone proteins and histone modifying enzymes. *Nucleic Acids Research* **40**: D337–D342.

Khursheed M, Kolla JN, Kotapalli V, Gupta N, Gowrishankar S, Uppin SG, Sastry RA, Koganti S, Sundaram C, Pollack JR, et al. 2013. ARID1B, a member of the human SWI/SNF chromatin remodeling complex, exhibits tumour-suppressor activities in pancreatic cancer cell lines. *Br J Cancer* **108**: 2056–2062.

Kim KH, Roberts CWM. 2016. Targeting EZH2 in cancer. *Nat Med* **22**: 128–134.

Kinzler KW, Vogelstein B. 1997. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* **386**: 761–763.

Knudson AGJ. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **68**: 820–823.

Kontomanolis EN, Koutras A, Syllaios A, Schizas D, Mastoraki A, Garmpis N, Diakosavvas M, Angelou K, Tsatsaris G, Pagkalos A, et al. 2020. Role of Oncogenes and Tumor-suppressor Genes in Carcinogenesis: A Review. *Anticancer research* **40**: 6009–6015.

Kowalik TF, DeGregori J, Leone G, Jakoi L, Nevins JR. 1998. E2F1-specific induction of apoptosis and p53 accumulation, which is blocked by Mdm2. *Cell Growth Differ* **9**: 113–118.

Kozomara A, Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research* **39**: D152–7.

Kryston TB, Georgiev AB, Pissis P, Georgakilas AG. 2011. Role of oxidative stress and DNA damage in human carcinogenesis. *Mutat Res* **711**: 193–201.

Kuo K-T, Liang C-W, Hsiao C-H, Lin C-H, Chen C-A, Sheu B-C, Lin M-C. 2006. Downregulation of BRG-1 repressed expression of CD44s in cervical neuroendocrine carcinoma and adenocarcinoma. *Mod Pathol* **19**: 1570–1577.

Kurashima K, Kashiwagi H, Shimomura I, Suzuki A, Takeshita F, Mazevet M, Harata M, Yamashita T, Yamamoto Y, Kohno T, et al. 2020. SMARCA4 deficiency-associated heterochromatin induces intrinsic DNA replication stress and susceptibility to ATR inhibition in lung adenocarcinoma. *NAR Cancer* **2**: zcaa005.

Kwon H, Imbalzano AN, Khavari PA, Kingston RE, Green MR. 1994. Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. *Nature* **370**: 477–481.

Lai WKM, Pugh BF. 2017. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol* **18**: 548–562.

Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**: 495–501.

Lazar JE, Stehling-Sun S, Nandakumar V, Wang H, Chee DR, Howard NP, Acosta R, Dunn D, Diegel M, Neri F, et al. 2020. Global Regulatory DNA Potentiation by SMARCA4 Propagates to Selective Gene Expression Programs via Domain-Level Remodeling. *CellReports* **31**: 107676.

Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, et al. 2017. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357**: 409–413.

Le Loarer F, Watson S, Pierron G, de Montpreville VT, Ballet S, Firmin N, Auguste A, Pissaloux D, Boyault S, Paindavoine S, et al. 2015. SMARCA4 inactivation defines a group of undifferentiated thoracic malignancies transcriptionally related to BAF-deficient sarcomas. *Nat Genet* **47**: 1200–1205.

Lei K, Davis RJ. 2003. JNK phosphorylation of Bim-related members of the Bcl2 family induces Bax-dependent apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 2432–2437.

Lessard J, Wu JI, Ranish JA, Wan M, Winslow MM, Staahl BT, Wu H, Aebersold R, Graef IA, Crabtree GR. 2007. An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55**: 201–215.

Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, et al. 2010. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* **363**: 2424–2433.

Li J, Ju J, Ni B, Wang H. 2016a. The emerging role of miR-506 in cancer. *Oncotarget* **7**: 62778–62788.

Li J, Lu S, Lombardo K, Monahan R, Amin A. 2016b. ARID1A alteration in aggressive urothelial carcinoma and variants of urothelial carcinoma. *Hum Pathol* **55**: 17–23.

Li J, Zhang Z, Chen F, Hu T, Peng W, Gu Q, Sun Y. 2019. The Diverse Oncogenic and Tumor Suppressor Roles of microRNA-105 in Cancer. *Front Oncol* **9**: 518.

Lickert H, Takeuchi JK, Both Von I, Walls JR, McAuliffe F, Adamson SL, Henkelman RM, Wrana JL, Rossant J, Bruneau BG. 2004. Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* **432**: 107–112.

Lin S, Gregory RI. 2015. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer* **15**: 321–333.

Lin WC, Lin FT, Nevins JR. 2001. Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. *Genes Dev* **15**: 1833–1844.

Lissanu Deribe Y, Sun Y, Terranova C, Khan F, Martinez-Ledesma J, Gay J, Gao G, Mullinax RA, Khor T, Feng N, et al. 2018. Mutations in the SWI/SNF complex induce a targetable dependence on oxidative phosphorylation in lung cancer. *Nat Med* **24**: 1047–1057.

Liu X, Chen X, Yu X, Tao Y, Bode AM, Dong Z, Cao Y. 2013. Regulation of microRNAs by epigenetics and their interplay involved in cancer. *J Exp Clin Cancer Res* **32**: 96.

Liu X, Yu J, Jiang L, Wang A, Shi F, Ye H, Zhou X. 2009. MicroRNA-222 regulates cell invasion by targeting matrix metalloproteinase 1 (MMP1) and manganese

superoxide dismutase 2 (SOD2) in tongue squamous cell carcinoma cell lines. *Cancer Genomics Proteomics* **6**: 131–139.

Liu Z, Turkoz A, Jackson EN, Corbo JC, Engelbach JA, Garbow JR, Piwnica-Worms DR, Kopan R. 2011. Notch1 loss of heterozygosity causes vascular tumors and lethal hemorrhage in mice. *J Clin Invest* **121**: 800–808.

Love C, Sun Z, Jima D, Li G, Zhang J, Miles R, Richards KL, Dunphy CH, Choi WWL, Srivastava G, et al. 2012. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet* **44**: 1321–1325.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Lu D, Wolfgang CD, Hai T. 2006. Activating transcription factor 3, a stress-inducible gene, suppresses Ras-stimulated tumorigenesis. *The Journal of biological chemistry* **281**: 10473–10481.

Lu P, Roberts CWM. 2013. The SWI/SNF tumor suppressor complex: Regulation of promoter nucleosomes and beyond. *nucleus* **4**: 374–378.

Luger K, Dechassa ML, Tremethick DJ. 2012. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol* **13**: 436–447.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251–260.

Lunning MA, Green MR. 2015. Mutation of chromatin modifiers; an emerging hallmark of germinal center B-cell lymphomas. *Blood Cancer J* **5**: e361.

Luo Q, Wu X, Chang W, Zhao P, Nan Y, Zhu X, Katz JP, Su D, Liu Z. 2020a. ARID1A prevents squamous cell carcinoma initiation and chemoresistance by antagonizing pRb/E2F1/c-Myc-mediated cancer stemness. *Cell Death Differ* **27**: 1981–1997.

Luo Q, Wu X, Chang W, Zhao P, Zhu X, Chen H, Nan Y, Luo A, Zhou X, Su D, et al. 2020b. ARID1A Hypermethylation Disrupts Transcriptional Homeostasis to Promote Squamous Cell Carcinoma Progression. *Cancer Research* **80**: 406–417.

Macher-Goeppinger S, Keith M, Tagscherer KE, Singer S, Winkler J, Hofmann TG, Pahernik S, Duensing S, Hohenfellner M, Kopitz J, et al. 2015. PBRM1 (BAF180) protein is functionally regulated by p53-induced protein degradation in renal cell carcinomas. *J Pathol* **237**: 460–471.

Malhotra JD, Kaufman RJ. 2007. Endoplasmic reticulum stress and oxidative stress: a vicious cycle or a double-edged sword? *Antioxid Redox Signal* **9**: 2277–2293.

Mallappa C, Nasipak BT, Etheridge L, Androphy EJ, Jones SN, Sagerstrom CG, Ohkawa Y, Imbalzano AN. 2010. Myogenic microRNA expression requires ATP-

dependent chromatin remodeling enzyme function. *Molecular and cellular biology* **30**: 3176–3186.

Marciniak SJ, Yun CY, Oyadomari S, Novoa I, Zhang Y, Jungreis R, Nagata K, Harding HP, Ron D. 2004. CHOP induces death by promoting protein synthesis and oxidation in the stressed endoplasmic reticulum. *Genes Dev* **18**: 3066–3077.

Marmorstein R. 2001. Protein modules that manipulate histone tails for chromatin regulation. *Nat Rev Mol Cell Biol* **2**: 422–432.

Marquez SB, Thompson KW, Lu L, Reisman D. 2015. Beyond Mutations: Additional Mechanisms and Implications of SWI/SNF Complex Inactivation. *Front Oncol* **4**: 372–372.

Marquez-Vilendrer SB, Rai SK, Gramling SJ, Lu L, Reisman DN. 2016. Loss of the SWI/SNF ATPase subunits BRM and BRG1 drives lung cancer development. *Oncoscience* **3**: 322–336.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**: 3–12.

Mashtalir N, D'Avino AR, Michel BC, Luo J, Pan J, Otto JE, Zullow HJ, McKenzie ZM, Kubiak RL, Pierre RS, et al. 2018. Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell* **175**: 1272–1288.

Massagué J, Obenauf AC. 2016. Metastatic colonization by circulating tumour cells. *Nature* **529**: 298–306.

Mathur R, Alver BH, San Roman AK, Wilson BG, Wang X, Agoston AT, Park PJ, Shivdasani RA, Roberts CWM. 2017. ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat Genet* **49**: 296–302.

Mathur R, Roberts CWM. 2018. SWI/SNF (BAF) Complexes: Guardians of the Epigenome. *Annu Rev Cancer Biol* **2**: 413–427.

McCullough KD, Martindale JL, Klotz LO, Aw TY, Holbrook NJ. 2001. Gadd153 sensitizes cells to endoplasmic reticulum stress by down-regulating Bcl2 and perturbing the cellular redox state. *Molecular and cellular biology* **21**: 1249–1259.

Medina PP, Carretero J, Ballestar E, Angulo B, Lopez-Rios F, Esteller M, Sanchez-Cespedes M. 2005. Transcriptional targets of the chromatin-remodelling factor SMARCA4/BRG1 in lung cancer cells. *Human molecular genetics* **14**: 973–982.

Medina PP, Carretero J, Fraga MF, Esteller M, Sidransky D, Sanchez-Cespedes M. 2004. Genetic and Epigenetic screening for gene alterations of the chromatin-remodeling factor, SMARCA4/BRG1, in lung tumors. *Genes, chromosomes & cancer* **41**: 170–177.

Medina PP, Romero OA, Kohno T, Montuenga LM, Pio R, Yokota J, Sanchez-Cespedes M. 2008a. Frequent BRG1/SMARCA4-inactivating mutations in human lung cancer cell lines. *Human mutation* **29**: 617–622.

Medina PP, Slack FJ. 2008. microRNAs and cancer: an overview. *Cell Cycle* **7**: 2485–2492.

Medina R, Zaidi SK, Liu C-G, Stein JL, van Wijnen AJ, Croce CM, Stein GS. 2008b. MicroRNAs 221 and 222 bypass quiescence and compromise cell survival. *Cancer Research* **68**: 2773–2780.

Merritt WM, Lin YG, Han LY, Kamat AA, Spannuth WA, Schmandt R, Urbauer D, Pennacchio LA, Cheng J-F, Nick AM, et al. 2008. Dicer, Drosha, and outcomes in patients with ovarian cancer. *N Engl J Med* **359**: 2641–2650.

Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S, et al. 2017. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* **49**: 1779–1784.

Michel BC, D'Avino AR, Cassel SH, Mashtalir N, McKenzie ZM, McBride MJ, Valencia AM, Zhou Q, Bocker M, Soares LMM, et al. 2018. A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nat Cell Biol* **20**: 1410–1420.

Middeljans E, Wan X, Jansen PW, Sharma V, Stunnenberg HG, Logie C. 2012. SS18 together with animal-specific factors defines human BAF-type SWI/SNF complexes. *PLOS one* **7**: e33834.

Mittal P, Roberts CWM. 2020. The SWI/SNF complex in cancer - biology, biomarkers and therapy. *Nature Reviews Clinical Oncology* **17**: 435–448.

Mizutani T, Ito T, Nishina M, Yamamichi N, Watanabe A, Iba H. 2002. Maintenance of integrated proviral gene expression requires Brm, a catalytic subunit of SWI/SNF complex. *The Journal of biological chemistry* **277**: 15859–15864.

Morris LGT, Riaz N, Desrichard A, Şenbabaoğlu Y, Hakimi AA, Makarov V, Reis-Filho JS, Chan TA. 2016. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget* **7**: 10051–10063.

Moutinho C, Esteller M. 2017. MicroRNAs and Epigenetics. *Adv Cancer Res* **135**: 189–220.

Mouw KW, Goldberg MS, Konstantinopoulos PA, D'Andrea AD. 2017. DNA Damage and Repair Biomarkers of Immunotherapy Response. *Cancer Discov* **7**: 675–693.

Nagl NGJ, Wang X, Patsialou A, Van Scoy M, Moran E. 2007. Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *EMBO J* **26**: 752–763.

Naito T, Udagawa H, Umemura S, Sakai T, Zenke Y, Kirita K, Matsumoto S, Yoh K, Niho S, Tsuboi M, et al. 2019. Non-small cell lung cancer with loss of expression of the SWI/SNF complex is associated with aggressive clinicopathological features, PD-L1-positive status, and high tumor mutation burden. *Lung Cancer* **138**: 35–42.

Neigeborn L, Carlson M. 1984. Genes affecting the regulation of SUC2 gene expression by glucose repression in Saccharomyces cerevisiae. *Genetics* **108**: 845–858.

Network CGAR. 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**: 543–550.

Network CGAR. 2011. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**: 609–615.

Nie Z, Xue Y, Yang D, Zhou S, Deroo BJ, Archer TK, Wang W. 2000. A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex. *Molecular and cellular biology* **20**: 8879–8888.

Nishino M, Ramaiya NH, Hatabu H, Hodi FS. 2017. Monitoring immune-checkpoint blockade: response evaluation and biomarker development. *Nature Reviews Clinical Oncology* **14**: 655–668.

Nishitoh H. 2012. CHOP is a multifunctional transcription factor in the ER stress response. *J Biochem* **151**: 217–219.

O'Brien J, Hayder H, Zayed Y, Peng C. 2018. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)* **9**: 402.

O'Connor DJ, Lu X. 2000. Stress signals induce transcriptionally inactive E2F-1 independently of p53 and Rb. *Oncogene* **19**: 2369–2376.

Ogiwara H, Takahashi K, Sasaki M, Kuroda T, Yoshida H, Watanabe R, Maruyama A, Makinoshima H, Chiwaki F, Sasaki H, et al. 2019. Targeting the Vulnerability of Glutathione Metabolism in ARID1A-Deficient Cancers. *Cancer cell* **35**: 177–190.e8.

Ogiwara H, Ui A, Otsuka A, Satoh H, Yokomi I, Nakajima S, Yasui A, Yokota J, Kohno T. 2011. Histone acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors. *Oncogene* **30**: 2135–2146.

Oike T, Ogiwara H, Tominaga Y, Ito K, Ando O, Tsuta K, Mizukami T, Shimada Y, Isomura H, Komachi M, et al. 2013. A Synthetic Lethality-Based Strategy to Treat Cancers Harboring a Genetic Deficiency in the Chromatin Remodeling Factor BRG1. *Cancer Research* **73**: 5508–5518.

Olave I, Wang W, Xue Y, Kuo A, Crabtree GR. 2002. Identification of a polymorphic, neuron-specific chromatin remodeling complex. *Genes Dev* **16**: 2509–2517.

Ooi L, Belyaev ND, Miyake K, Wood IC, Buckley NJ. 2006. BRG1 chromatin remodeling activity is required for efficient chromatin binding by repressor element 1-silencing transcription factor (REST) and facilitates REST-mediated repression. *The Journal of biological chemistry* **281**: 38974–38980.

Orlando KA, Nguyen V, Raab JR, Walhart T, Weissman BE. 2019. Remodeling the cancer epigenome: mutations in the SWI/SNF complex offer new therapeutic opportunities. *Expert Rev Anticancer Ther* **19**: 375–391.

Orvis T, Hepperla A, Walter V, Song S, Simon J, Parker J, Wilkerson MD, Desai N, Major MB, Hayes DN, et al. 2014. BRG1/SMARCA4 inactivation promotes non-small cell lung cancer aggressiveness by altering chromatin organization. *Cancer Research* **74**: 6486–6498.

Owada-Ozaki Y, Muto S, Takagi H, Inoue T, Watanabe Y, Fukuhara M, Yamaura T, Okabe N, Matsumura Y, Hasegawa T, et al. 2018. Prognostic Impact of Tumor Mutation Burden in Patients With Completely Resected Non-Small Cell Lung Cancer: Brief Report. *J Thorac Oncol* **13**: 1217–1221.

Pan D, Kobayashi A, Jiang P, Ferrari de Andrade L, Tay RE, Luoma AM, Tsoucas D, Qiu X, Lim K, Rao P, et al. 2018. A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science* **359**: 770–775.

Papp G, Changchien Y-C, Péterfia B, Pecsenka L, Krausz T, Stricker TP, Khoor A, Donner L, Sápi Z. 2013. SMARCB1 protein and mRNA loss is not caused by promoter and histone hypermethylation in epithelioid sarcoma. *Mod Pathol* **26**: 393–403.

Park J-H, Park E-J, Lee H-S, Kim SJ, Hur S-K, Imbalzano AN, Kwon J. 2006. Mammalian SWI/SNF complexes facilitate DNA double-strand break repair by promoting gamma-H2AX induction. *EMBO J* **25**: 3986–3997.

Park Y, Chui MH, Suryo Rahmanto Y, Yu Z-C, Shamanna RA, Bellani MA, Gaillard S, Ayhan A, Viswanathan A, Seidman MM, et al. 2019. Loss of ARID1A in Tumor Cells Renders Selective Vulnerability to Combined Ionizing Radiation and PARP Inhibitor Therapy. *Clin Cancer Res* **25**: 5584–5594.

Park Y-A, Lee J-W, Kim H-S, Lee Y-Y, Kim T-J, Choi CH, Choi J-J, Jeon H-K, Cho YJ, Ryu JY, et al. 2014. Tumor suppressive effects of bromodomain-containing protein 7 (BRD7) in epithelial ovarian carcinoma. *Clin Cancer Res* **20**: 565–575.

Pascal JM. 2018. The comings and goings of PARP-1 in response to DNA damage. *DNA Repair* **71**: 177–182.

Pediconi N, Ianari A, Costanzo A, Belloni L, Gallo R, Cimino L, Porcellini A, Screpanti I, Balsano C, Alesse E, et al. 2003. Differential regulation of E2F1 apoptotic target genes in response to DNA damage. *Nat Cell Biol* **5**: 552–558.

Peterson CL, Herskowitz I. 1992. Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. *Cell* **68**: 573–583.

Petesch SJ, Lis JT. 2012. Overcoming the nucleosome barrier during transcript elongation. *Trends in Genetics* **28**: 285–294.

Phelan ML, Sif S, Narlikar GJ, Kingston RE. 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Molecular and cellular* **3**: 247–253.

Pulice JL, Kadoch C. 2017. Composition and Function of Mammalian SWI/SNF Chromatin Remodeling Complexes in Human Disease. *Cold Spring Harb Symp Quant Biol* **81**: 53–60.

Raab JR, Resnick S, Magnuson T. 2016. Genome-Wide Transcriptional Regulation Mediated by Biochemically Distinct SWI/SNF Complexes. *PLoS Genet* **11**: e1005748.

Rago F, DiMare MT, Elliott G, Ruddy DA, Sovath S, Kerr G, Bhang H-EC, Jagani Z. 2018. Degron mediated BRM/SMARCA2 depletion uncovers novel combination partners for treatment of BRG1/SMARCA4-mutant cancers. *Biochemical and Biophysical Research Communications* **1**: 109-116.

Rakheja D, Chen KS, Liu Y, Shukla AA, Schmid V, Chang T-C, Khokhar S, Wickiser JE, Karandikar NJ, Malter JS, et al. 2014. Somatic mutations in DROSHA and DICER1 impair microRNA biogenesis through distinct mechanisms in Wilms tumours. *Nature communications* **2**: 4802.

Reisman DN, Sciarrotta J, Wang W, Funkhouser WK, Weissman BE. 2003. Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. *Cancer Research* **63**: 560–566.

Restifo NP, Smyth MJ, Snyder A. 2016. Acquired resistance to immunotherapy and future challenges. *Nat Rev Cancer* **16**: 121–126.

Ribeiro-Silva C, Vermeulen W, Lans H. 2019. SWI/SNF_ Complex complexes in genome stability and cancer. *DNA Repair* **77**: 87–95.

Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, et al. 2015. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**: 124–128.

Roberts CWM, Leroux MM, Fleming MD, Orkin SH. 2002. Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. *Cancer cell* **2**: 415–425.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**: R25.

Rodriguez-Nieto S, Cañada A, Pros E, Pinto AI, Torres-Lanzas J, Lopez-Rios F, Sanchez-Verde L, Pisano DG, Sanchez-Cespedes M. 2010. Massive parallel DNA pyrosequencing analysis of the tumor suppressor BRG1/SMARCA4 in lung primary tumors. *Human mutation* **32**: E1999–E2017.

Rodriguez-Nieto S, Sanchez-Cespedes M. 2009. BRG1 and LKB1: tales of two tumor suppressor genes on chromosome 19p and lung cancer. *Carcinogenesis* **30**: 547–554.

Romero OA, Setien F, John S, Gimenez-Xavier P, Gómez-López G, Pisano D, Condom E, Villanueva A, Hager GL, Sanchez-Cespedes M. 2012. The tumour suppressor and chromatin-remodelling factor BRG1 antagonizes Myc activity and promotes cell differentiation in human cancer. *EMBO Mol Med* **4**: 603–616.

Roschke AV, Tonon G, Gehlhaus KS, McTyre N, Bussey KJ, Lababidi S, Scudiero DA, Weinstein JN, Kirsch IR. 2003. Karyotypic Complexity of the NCI-60 Drug-Screening Panel. *Cancer Res* **63**: 8634.

Roy N, Malik S, Villanueva KE, Urano A, Lu X, Figura von G, Seeley ES, Dawson DW, Collisson EA, Hebrok M. 2015. Brg1 promotes both tumor-suppressive and oncogenic activities at distinct stages of pancreatic cancer formation. *Genes Dev* **29**: 658–671.

Rutkowski DT, Arnold SM, Miller CN, Wu J, Li J, Gunnison KM, Mori K, Sadighi Akha AA, Raden D, Kaufman RJ. 2006. Adaptation to ER stress is mediated by differential stabilities of pro-survival and pro-apoptotic mRNAs and proteins. *Plos Biol* **4**: e374.

Ryan BM, Robles AI, Harris CC. 2010. Genetic variation in microRNA networks: the implications for cancer research. *Nat Rev Cancer* **10**: 389–402.

Ryme J, Asp P, Böhm S, Cavellán E, Farrants A-KO. 2009. Variations in the composition of mammalian SWI/SNF chromatin remodelling complexes. *Journal of cellular biochemistry* **108**: 565–576.

Ryujiro H, Aziz S. 2002. The SWI/SNF Chromatin-Remodeling Factor Stimulates Repair by Human Excision Nuclease in the Mononucleosome Core Particle. *Molecular and cellular biology* **22**: 6779–6787.

Sahu RK, Saha N, Das L, Sahu PK, Sariki SK, Tomar RS. 2020. SWI/SNF chromatin remodelling complex contributes to clearance of cytoplasmic protein aggregates and regulates unfolded protein response in Saccharomyces cerevisiae. *FEBS J* **287**: 3024–3041.

Sahu RK, Singh S, Tomar RS. 2021. The ATP-dependent SWI/SNF and RSC chromatin remodelers cooperatively induce unfolded protein response genes during endoplasmic reticulum stress. *Biochim Biophys Acta Gene Regul Mech* **1864**: 194748.

Saito Y, Liang G, Egger G, Friedman JM, Chuang JC, Coetzee GA, Jones PA. 2006. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer cell* **9**: 435–443.

Sakurai K, Furukawa C, Haraguchi T, Inada K-I, Shiogama K, Tagawa T, Fujita S, Ueno Y, Ogata A, Ito M, et al. 2011. MicroRNAs miR-199a-5p and -3p target the Brm subunit of SWI/SNF to generate a double-negative feedback loop in a variety of human cancers. *Cancer Research* **71**: 1680–1689.

Sand M, Gambichler T, Skrygan M, Sand D, Scola N, Altmeyer P, Bechara FG. 2010. Expression levels of the microRNA processing enzymes Drosha and dicer in epithelial skin cancer. *Cancer Invest* **28**: 649–653.

Sano R, Reed JC. 2013. ER stress-induced cell death mechanisms. *Biochim Biophys Acta* **1833**: 3460–3470.

Santos CXC, Tanaka LY, Wosniak J, Laurindo FRM. 2009. Mechanisms and implications of reactive oxygen species generation during the unfolded protein response: roles of endoplasmic reticulum oxidoreductases, mitochondrial electron transport, and NADPH oxidase. *Antioxid Redox Signal* **11**: 2409–2427.

Sato F, Tsuchiya S, Meltzer SJ, Shimizu K. 2011. MicroRNAs and epigenetics. *FEBS J* **278**: 1598–1609.

Savas S, Skardasi G. 2018. The SWI/SNF complex subunit genes: Their functions, variations, and links to risk and survival outcomes in human cancers. *Critical Reviews in Oncology / Hematology* **123**: 114–131.

Schallenberg S, Bork J, Essakly A, Alakus H, Buettner R, Hillmer AM, Bruns C, Schroeder W, Zander T, Loeser H, et al. 2020. Loss of the SWI/SNF-ATPase subunit members SMARCF1 (ARID1A), SMARCA2 (BRM), SMARCA4 (BRG1) and SMARCB1 (INI1) in oesophageal adenocarcinoma. *BMC cancer* **20**: 12.

Schiaffino-Ortega S, Balinas C, Cuadros M, Medina PP. 2014. SWI/SNF proteins as targets in cancer therapy. *J Hematol Oncol* **7**: 81.

Schick S, Rendeiro AF, Runggatscher K, Ringler A, Boidol B, Hinkel M, Májek P, Vulliard L, Penz T, Parapatics K, et al. 2019. Systematic characterization of BAF mutations provides insights into intracomplex synthetic lethalities in human cancers. *Nat Genet* **51**: 1399–1410.

Schübeler D. 2015. Function and information content of DNA methylation. *Nature* **517**: 321–326.

Sen M, Wang X, Hamdan FH, Rapp J, Eggert J, Kosinsky RL, Wegwitz F, Kutschat AP, Younesi FS, Gaedcke J, et al. 2019. ARID1A facilitates KRAS signaling- regulated enhancer activity in an AP1- dependent manner in colorectal cancer cells. *Clinical Epigenetics* **1**: 92.

Shain AH, Pollack JR. 2013. The spectrum of SWI/SNF mutations, ubiquitous in human cancers. *PLOS one* **8**: e55119.

Sharma A, Singh K, Almasan A. 2012. Histone H2AX phosphorylation: a marker for DNA damage. *Methods Mol Biol* **920**: 613–626.

Shaulian E, Schreiber M, Piu F, Beeche M, Wagner EF, Karin M. 2000. The mammalian UV response: c-Jun induction is required for exit from p53-imposed growth arrest. *Cell* **103**: 897–907.

Shen H, Laird PW. 2013. Interplay between the cancer genome and epigenome. *Cell* **153**: 38–55.

Shen J, Ju Z, Zhao W, Wang L, Peng Y, Ge Z, Nagel ZD, Zou J, Wang C, Kapoor P, et al. 2018. ARID1A deficiency promotes mutability and potentiates therapeutic antitumor immunity unleashed by immune checkpoint blockade. *Nat Med* **24**: 556–562.

Shen J, Peng Y, Wei L, Zhang W, Yang L, Lan L, Kapoor P, Ju Z, Mo Q, Shih I-M, et al. 2015. ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors. *Cancer Discov* **5**: 752–767.

Shi J, Whyte WA, Zepeda-Mendoza CJ, Milazzo JP, Shen C, Roe J-S, Minder JL, Mercan F, Wang E, Eckersley-Maslin MA, et al. 2013. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev* **27**: 2648–2662.

Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. 2012. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research* **40**: W452–7.

Sims HI, Lane JM, Ulyanova NP, Schnitzler GR. 2007. Human SWI/SNF drives sequence-directed repositioning of nucleosomes on C-myc promoter DNA minicircles. *Biochemistry* **46**: 11377–11388.

Son EY, Crabtree GR. 2014. The role of BAF (mSWI/SNF) complexes in mammalian neural development. *Am J Med Genet C Semin Med Genet* **166C**: 333–349.

Song J, Ouyang Y, Che J, Li X, Zhao Y, Yang K, Zhao X, Chen Y, Fan C, Yuan W. 2017. Potential Value of miR-221/222 as Diagnostic, Prognostic, and Therapeutic Biomarkers for Diseases. *Front Immunol* **8**: 56.

Song L, Zhang Z, Grasfeder LL, Boyle AP, Giresi PG, Lee B-K, Sheffield NC, Gräf S, Huss M, Keefe D, et al. 2011. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res* **21**: 1757–1767.

Song S, Walter V, Karaca M, Li Y, Bartlett CS, Smiraglia DJ, Serber D, Sproul CD, Plass C, Zhang J, et al. 2014. Gene silencing associated with SWI/SNF complex loss during NSCLC development. *Molecular cancer research: MCR* **12**: 560–570.

Soria G, Polo SE, Almouzni G. 2012. Prime, Repair, Restore: The Active Role of Chromatin in the DNA Damage Response. *Molecular and cellular* **46**: 722–734.

Staahl BT, Tang J, Wu W, Sun A, Gitler AD, Yoo AS, Crabtree GR. 2013. Kinetic analysis of npBAF to nBAF switching reveals exchange of SS18 with CREST and integration with neural developmental pathways. *J Neurosci* **33**: 10348–10361.

Stanton BZ, Hodges C, Calarco JP, Braun SMG, Ku WL, Kadoch C, Zhao K, Crabtree GR. 2017. Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat Genet* **49**: 282–288.

Stern M, Jensen R, Herskowitz I. 1984. Five SWI genes are required for expression of the HO gene in yeast. *J Mol Biol* **178**: 853–868.

Stine ZE, Walton ZE, Altman BJ, Hsieh AL, Dang CV. 2015. MYC, Metabolism, and Cancer. *Cancer Discov* **5**: 1024–1039.

Stratton MR, Campbell PJ, Futreal PA. 2009. The cancer genome. *Nature* **458**: 719–724.

Strobeck MW, Reisman DN, Gunawardena RW, Betz BL, Angus SP, Knudsen KE, Kowalik TF, Weissman BE, Knudsen ES. 2002. Compensation of BRG-1 function by Brm: insight into the role of the core SWI-SNF subunits in retinoblastoma tumor suppressor signaling. *The Journal of biological chemistry* **277**: 4782–4789.

Sun X, Wang SC, Wei Y, Luo X, Jia Y, Li L, Gopal P, Zhu M, Nassour I, Chuang J-C, et al. 2017. Arid1a Has Context-Dependent Oncogenic and Tumor Suppressor Functions in Liver Cancer. *Cancer cell* **32**: 574–589.e6.

Swift LH, Golsteyn RM. 2014. Genotoxic anti-cancer agents and their relationship to DNA damage, mitosis, and checkpoint adaptation in proliferating cancer cells. *Int J Mol Sci* **15**: 3403–3431.

Tabas I, Ron D. 2011. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nat Cell Biol* **13**: 184–190.

Tagal V, Wei S, Zhang W, Brekken RA, Posner BA, Peyton M, Girard L, Hwang T, Wheeler DA, Minna JD, et al. 2017. SMARCA4-inactivating mutations increase sensitivity to Aurora kinase A inhibitor VX-680 in non-small cell lung cancers. *Nature communications* **8**: 14098.

Takao C, Morikawa A, Ohkubo H, Kito Y, Saigo C, Sakuratani T, Futamura M, Takeuchi T, Yoshida K. 2017. Downregulation of ARID1A, a component of the SWI/SNF chromatin remodeling complex, in breast cancer. *J Cancer* **8**: 1–8.

Tameire F, Verginadis II, Koumenis C. 2015. Cell intrinsic and extrinsic activators of the unfolded protein response in cancer: Mechanisms and targets for therapy. *Seminars in Cancer Biology* **33**: 3–15.

Tang Z, Kang B, Li C, Chen T, Zhang Z. 2019. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Research* **47**: W556–W560.

Taulli R, Foglizzo V, Morena D, Coda DM, Ala U, Bersani F, Maestro N, Ponzetto C. 2014. Failure to downregulate the BAF53a subunit of the SWI/SNF chromatin remodeling complex contributes to the differentiation block in rhabdomyosarcoma. *Oncogene* **33**: 2354–2362.

Teske BF, Fusakio ME, Zhou D, Shan J, McClintick JN, Kilberg MS, Wek RC. 2013. CHOP induces activating transcription factor 5 (ATF5) to trigger apoptosis in response to perturbations in protein homeostasis. *Mol Biol Cell* **24**: 2477–2490.

Tessarz P, Kouzarides T. 2014. Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol* **15**: 703–708.

Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al. 2012. The accessible chromatin landscape of the human genome. *Nature* **489**: 75–82.

Tolstorukov MY, Sansam CG, Lu P, Koellhoffer EC, Helming KC, Alver BH, Tillman EJ, Evans JA, Wilson BG, Park PJ, et al. 2013. Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. *Proc Natl Acad Sci USA* **110**: 10165.

Tremethick DJ. 2007. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* **128**: 651–654.

Tsai S, Fournier L-A, Chang EY-C, Wells JP, Minaker SW, Zhu YD, Wang AY-H, Wang Y, Huntsman DG, Stirling PC. 2021. ARID1A regulates R-loop associated DNA replication stress. *PLoS Genet* **17**: e1009238.

Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, Kaname T, Naritomi K, Kawame H, Wakui K, et al. 2012. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* **44**: 376–378.

Turchi L, Aberdam E, Mazure N, Pouysségur J, Deckert M, Kitajima S, Aberdam D, Virolle T. 2008. Hif-2alpha mediates UV-induced apoptosis through a novel ATF3-dependent death pathway. *Cell Death Differ* **15**: 1472–1480.

Turchi L, Fareh M, Aberdam E, Kitajima S, Simpson F, Wicking C, Aberdam D, Virolle T. 2009. ATF3 and p15PAF are novel gatekeepers of genomic integrity upon UV stress. *Cell Death Differ* **16**: 728–737.

Tyagi A, Ryme J, Brodin D, Ostlund Farrants AK, Visa N. 2009. SWI/SNF associates with nascent pre-mRNPs and regulates alternative pre-mRNA processing. *PLoS Genet* **5**: e1000470.

Underhill C, Qutob MS, Yee SP, Torchia J. 2000. A novel nuclear receptor corepressor complex, N-CoR, contains components of the mammalian SWI/SNF complex and the corepressor KAP-1. *The Journal of biological chemistry* **275**: 40463–40470.

Valencia AM, Kadoch C. 2019. Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat Cell Biol* **21**: 152–161.

Valero C, Lee M, Hoen D, Wang J, Nadeem Z, Patel N, Postow MA, Shoushtari AN, Plitas G, Balachandran VP, et al. 2021. The association between tumor mutational burden and prognosis is dependent on treatment context. *Nat Genet* **53**: 11–15.

Valletta M, Russo R, Baglivo I, Russo V, Ragucci S, Sandomenico A, Iaccarino E, Ruvo M, De Feis I, Angelini C, et al. 2020. Exploring the Interaction between the SWI/SNF Chromatin Remodeling Complex and the Zinc Finger Factor CTCF. *Int J Mol Sci* **21**: 8950.

van de Haar J, Canisius S, Yu MK, Voest EE, Wessels LFA, Ideker T. 2019. Identifying Epistasis in Cancer Genomes: A Delicate Affair. *Cell* **177**: 1375–1383.

van Riggelen J, Yetil A, Felsher DW. 2010. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer* **10**: 301–309.

Versteege I, Sévenet N, Lange J, Rousseau-Merck MF, Ambros P, Handgretinger R, Aurias A, Delattre O. 1998. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394**: 203–206.

Vierbuchen T, Ling E, Cowley CJ, Couch CH, Wang X, Harmin DA, Roberts CWM, Greenberg ME. 2017. AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. *Molecular and cellular* **68**: 1067–1082.e12.

Visone R, Russo L, Pallante P, De Martino I, Ferraro A, Leone V, Borbone E, Petrocca F, Alder H, Croce CM, et al. 2007. MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle. *Endocr Relat Cancer* **14**: 791–798.

Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T, Hatzigeorgiou AG. 2015. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucleic Acids Research* **43**: W460–W466.

Wade PA, Gegonne A, Jones PL, Ballestar E, Aubry F, Wolffe AP. 1999. Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat Genet* **23**: 62–66.

Wang K, Kan J, Yuen ST, Shi ST, Chu KM, Law S, Chan TL, Kan Z, Chan ASY, Tsui WY, et al. 2011. Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat Genet* **43**: 1219–1223.

Wang Q, Wu X. 2017. Primary and acquired resistance to PD-1/PD-L1 blockade in cancer treatment. *Int Immunopharmacol* **46**: 210–219.

Wang SC, Nassour I, Xiao S, Zhang S, Luo X, Lee J, Li L, Sun X, Nguyen LH, Chuang J-C, et al. 2019. SWI/SNF component ARID1A restrains pancreatic neoplasia formation. *Gut* **68**: 1259–1270.

Wang W, Côté J, Xue Y, Zhou S, Khavari PA, Biggar SR, Muchardt C, Kalpana GV, Goff SP, Yaniv M, et al. 1996. Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J* **15**: 5370–5382.

Wang X, Lee RS, Alver BH, Haswell JR, Wang S, Mieczkowski J, Drier Y, Gillespie SM, Archer TC, Wu JN, et al. 2017. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat Genet* **49**: 289–295.

Wang X, Sansam CG, Thom CS, Metzger D, Evans JA, Nguyen PTL, Roberts CWM. 2009. Oncogenesis caused by loss of the SNF5 tumor suppressor is dependent on activity of BRG1, the ATPase of the SWI/SNF chromatin remodeling complex. *Cancer Research* **69**: 8094–8101.

Wanior M, Krämer A, Knapp S, Joerger AC. 2021. Exploiting vulnerabilities of SWI/SNF chromatin remodelling complexes for cancer therapy. *Oncogene* **40**: 3637–3654.

Watanabe R, Ui A, Kanno SI, Ogiwara H, Nagase T, Kohno T, Yasui A. 2014. SWI/SNF Factors Required for Cellular Resistance to DNA Damage Include ARID1A and ARID1B and Show Interdependent Protein Stability. *Cancer Research* **74**: 2465–2475.

Watanabe T, Semba S, Yokozaki H. 2011. Regulation of PTEN expression by the SWI/SNF chromatin-remodelling protein BRG1 in human colorectal carcinoma cells. *Br J Cancer* **104**: 146–154.

Weber CM, Hafner A, Kirkland JG, Braun SMG, Stanton BZ, Boettiger AN, Crabtree GR. 2021. mSWI/SNF promotes Polycomb repression both directly and through genome-wide redistribution. *Nat Struct Mol Biol* **28**: 501–511.

Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. 2013. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45**: 1113–1120.

Weissmiller AM, Wang J, Lorey SL, Howard GC, Martinez E, Liu Q, Tansey WP. 2019. Inhibition of MYC by the SMARCB1 tumor suppressor. *Nature communications* **10**: 2014.

Whitehouse I, Tsukiyama T. 2006. Antagonistic forces that position nucleosomes in vivo. *Nat Struct Mol Biol* **13**: 633–640.

Wieczorek D, Bögershausen N, Beleggia F, Steiner-Haldenstätt S, Pohl E, Li Y, Milz E, Martin M, Thiele H, Altmüller J, et al. 2013. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Human molecular genetics* **22**: 5121–5135.

Wiegand KC, Shah SP, Al-Agha OM, Zhao Y, Tse K, Zeng T, Senz J, McConechy MK, Anglesio MS, Kalloger SE, et al. 2010. ARID1A mutations in endometriosis-associated ovarian carcinomas. *N Engl J Med* **363**: 1532–1543.

Wilding JL, Bodmer WF. 2014. Cancer Cell Lines for Drug Discovery and Development. *Cancer Res* **74**: 2377.

Williamson CT, Miller R, Pemberton HN, Jones SE, Campbell J, Konde A, Badham N, Rafiq R, Brough R, Gulati A, et al. 2016. ATR inhibitors as a synthetic lethal therapy for tumours deficient in ARID1A. *Nature communications* **7**: 13837.

Willis MS, Homeister JW, Rosson GB, Annayev Y, Holley D, Holly SP, Madden VJ, Godfrey V, Parise LV, Bultman SJ. 2012. Functional redundancy of SWI/SNF catalytic subunits in maintaining vascular endothelial cells in the adult heart. *Circ Res* **111**: e111–22.

Wilsker D, Patsialou A, Zumbrun SD, Kim S, Chen Y, Dallas PB, Moran E. 2004. The DNA-binding properties of the ARID-containing subunits of yeast and mammalian SWI/SNF complexes. *Nucleic Acids Research* **32**: 1345–1353.

Wilson BG, Helming KC, Wang X, Kim Y, Vazquez F, Jagani Z, Hahn WC, Roberts CWM. 2014. Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation. *Molecular and cellular biology* **34**: 1136–1144.

Winter J, Jung S, Keller S, Gregory RI, Diederichs S. 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* **11**: 228–234.

Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, Grynspan D, Saloustros E, Nadaf J, Rivera B, et al. 2014. Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. *Nat Genet* **46**: 438–443.

Wong AK, Shanahan F, Chen Y, Lian L, Ha P, Hendricks K, Ghaffari S, Iliev D, Penn B, Woodland AM, et al. 2000. BRG1, a component of the SWI-SNF complex, is mutated in multiple human tumor cell lines. *Cancer Research* **60**: 6171–6177.

Wu J, He K, Zhang Y, Song J, Shi Z, Chen W, Shao Y. 2019. Inactivation of SMARCA2 by promoter hypermethylation drives lung cancer development. *Gene* **687**: 193–199.

Wu JN, Roberts CWM. 2013. ARID1A mutations in cancer: another epigenetic tumor suppressor? *Cancer Discov* **3**: 35–43.

Wu R-C, Wang T-L, Shih I-M. 2014. The emerging roles of ARID1A in tumor suppression. *Cancer Biol Ther* **15**: 655–664.

Xiao Y, Hsiao T-H, Suresh U, Chen H-IH, Wu X, Wolf SE, Chen Y. 2014. A novel significance score for gene selection and ranking. *Bioinformatics* **30**: 801–807.

Xu F, Flowers S, Moran E. 2012. Essential role of ARID2 protein-containing SWI/SNF complex in tissue-specific gene expression. *The Journal of biological chemistry* **287**: 5033–5041.

Xu W, San Lucas A, Wang Z, Liu Y. 2014. Identifying microRNA targets in different gene regions. *BMC Bioinformatics* **15 Suppl 7**: S4.

Xue L, Wang Y, Yue S, Zhang J. 2017. The expression of miRNA-221 and miRNA-222 in gliomas patients and their prognosis. *Neurol Sci* **38**: 67–73.

Xue Y, Canman JC, Lee CS, Nie Z, Yang D, Moreno GT, Young MK, Salmon ED, Wang W. 2000. The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 13015–13020.

Xue Y, Meehan B, Fu Z, Wang XQD, Fiset PO, Rieker R, Levins C, Kong T, Zhu X, Morin G, et al. 2019. SMARCA4 loss is synthetic lethal with CDK4/6 inhibition in non-small cell lung cancer. *Nature communications* **10**: 557.

Yamaguchi H, Wang H-G. 2004. CHOP is involved in endoplasmic reticulum stress-induced apoptosis by enhancing DR5 expression in human carcinoma cells. *The Journal of biological chemistry* **279**: 45495–45502.

Yamamichi N, Yamamichi-Nishina M, Mizutani T, Watanabe H, Minoguchi S, Kobayashi N, Kimura S, Ito T, Yahagi N, Ichinose M, et al. 2005. The Brm gene suppressed at the post-transcriptional level in various human cell lines is inducible by transient HDAC inhibitor treatment, which exhibits antioncogenic potential. *Oncogene* **24**: 5471–5481.

Yamashita R, Sato M, Kakumu T, Hase T, Yogo N, Maruyama E, Sekido Y, Kondo M, Hasegawa Y. 2015. Growth inhibitory effects of miR-221 and miR-222 in non-small cell lung cancer cells. *Cancer medicine* **4**: 551–564.

Yan X-J, Xu J, Gu Z-H, Pan C-M, Lu G, Shen Y, Shi J-Y, Zhu Y-M, Tang L, Zhang X-W, et al. 2011. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet* **43**: 309–315.

Yan Z, Wang Z, Sharova L, Sharov AA, Ling C, Piao Y, Aiba K, Matoba R, Wang W, Ko MSH. 2008. BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. *Stem Cells* **26**: 1155–1165.

You JS, Jones PA. 2012. Cancer genetics and epigenetics: two sides of the same coin? *Cancer cell* **22**: 9–20.

Zhang B, Chambers KJ, Faller DV, Wang S. 2007. Reprogramming of the SWI/SNF complex for co-activation or co-repression in prohibitin-mediated estrogen receptor regulation. *Oncogene* **26**: 7153–7157.

Zhang C, Zhang J, Zhang A, Wang Y, Han L, You Y, Pu P, Kang C. 2010. PUMA is a novel target of miR-221/222 in human epithelial cancers. *Int J Oncol* **37**: 1621–1626.

Zhang L, Nemzow L, Chen H, Hu JJ, Gong F. 2014. Whole Genome Expression Profiling Shows that BRG1 Transcriptionally Regulates UV Inducible Genes and Other Novel Targets in Human Cells. *PLOS one* **9**: e105764.

Zhang L, Wang C, Yu S, Jia C, Yan J, Lu Z, Chen J. 2018. Loss of ARID1A Expression Correlates with Tumor Differentiation and Tumor Progression Stage in Pancreatic Ductal Adenocarcinoma. *Technol Cancer Res Treat* **17**: 1533034618754475.

Zhang T, Li N, Sun C, Jin Y, Sheng X. 2020. MYC and the unfolded protein response in cancer: synthetic lethal partners in crime? *EMBO Mol Med* **12**: e11845.

Zhao J, Chen J, Lin H, Jin R, Liu J, Liu X, Meng N, Cai X. 2016. The Clinicopathologic Significance of BAF250a (ARID1A) Expression in Hepatocellular Carcinoma. *Pathol Oncol Res* **22**: 453–459.

Zhao S, Gordon W, Du S, Zhang C, He W, Xi L, Mathur S, Agostino M, Paradis T, Schack von D, et al. 2017. QuickMIRSeq: a pipeline for quick and accurate quantification of both known miRNAs and isomiRs by jointly processing multiple samples from microRNA sequencing. *BMC Bioinformatics* **18**: 180.

Zhou M, Yuan J, Deng Y, Fan X, Shen J. 2021. Emerging role of SWI/SNF complex deficiency as a target of immune checkpoint blockade in human cancers. *Oncogenesis* **10**: 3–3.

Zhu G, Shi R, Li Y, Zhang Z, Xu S, Chen C, Cao P, Zhang H, Liu M, Pan Z, et al. 2021. ARID1A, ARID1B, and ARID2 Mutations Serve as Potential Biomarkers for Immune Checkpoint Blockade in Patients with Non-Small Cell Lung Cancer. *Front Immunol* **12**: 670040.

Zundell JA, Fukumoto T, Lin J, Fatkhudinov N, Nacarelli T, Kossenkov AV, Liu Q, Cassel J, Hu C-CA, Wu S, et al. 2021. Targeting the IRE1α/XBP1 Endoplasmic Reticulum Stress Response Pathway in ARID1A-Mutant Ovarian Cancers. *Cancer Res* **81**: 5325.

ANNEX

ANNEX

Table of content

Annex I: Compilation of the subunits of the mammalian SWI/SNF complexes

Supplementary Table 1: SWI/SNF subunits described in mammalian cells. Alternative names of each subunit are displayed in brackets.

	SWI/SNF complex configurations		
Role	**BAF**	**PBAF**	**ncBAF**
ATPase module	SMARCA4 (BRG1) or SMARCA2 (BRM)		
ATPase module	BCL7A/B/C		
ATPase module	ACTB		
ATPase module	ACTL6A (BAF53A) or ACTL6B (BAF53B) *		
ATPase module	SS18 (SSXT) * or SS18L1 (CREST) *		
Core	SMARCC1 (BAF155)		
Core	SMARCC2 (BAF170)		
Core	SMARCD1 (BAF60A) or SMARCD2 (BAF60B) or SMARCD3 (BAF60C)		
Core			BICRA (GLTSCR1) or BICRAL (GLTSCR1L)
Core	SMARCB1 (BAF47, INI1, SNF5)		
Core	SMARCE1 (BAF57)		
Accessory	ARID1A (BAF250A) or ARID1B (BAF250B)	ARID2 (BAF200)	
Accessory		BRD7	BRD9
Accessory	DPF1 (BAF45B) or DPF2 (BAF45C) or DPF3 (BAF45D)	PHF10 (BAF45A)	
Accessory		PBRM1 (BAF180)	

*ACTL6B, SS18 and SS18L1 are only found in neural-specific SWI/SNF complexes

Annex II: Lung adenocarcinoma cell lines used in this Ph.D. thesis

Supplementary Table 2: List of LUAD cell lines. The second column shows a classification of the mutational status of the SWI/SNF complex. The third column compilates other mutated genes that are among the top five LUAD driver genes described by Bailey et al *(Bailey et al. 2018)*. The fourth column indicates the growth media recommended by the supplier.

Cell line	SWI/SNF status*	Other mutated LUAD driver genes	Growth media
A427	Mutated	KRAS	DMEM
A549	Mutated	KRAS, STK11	DMEM
CALU3	Mutated	TP53	DMEM
H1373	Wild type	KRAS, TP53	RPMI-1640
H1395	Wild type	BRAF, STK11	RPMI-1640
H1435	Mutated	TP53	DMEM
H1437	Wild type	TP53	RPMI-1640
H1568	Mutated	EGFR, TP53	RPMI-1640
H1573	Mutated	KRAS, TP53, STK11	RPMI-1640
H1623	Mutated	TP53, STK11	RPMI-1640
H1648	Mutated	TP53	DMEM
H1650	Mutated	EGFR, TP53	RPMI-1640
H1734	Mutated	KRAS, TP53, STK11	RPMI-1640
H1792	Mutated	KRAS, TP53	RPMI-1640
H1793	Mutated	EGFR, TP53	DMEM
H1944	Mutated	KRAS, STK11	RPMI-1640
H1975	Mutated	EGFR, TP53	RPMI-1640
H2009	Mutated	KRAS, EGFR, TP53	DMEM
H2030	Mutated	KRAS, TP53, STK11	RPMI-1640
H2087	Wild type	BRAF, TP53	RPMI-1640
H2122	Mutated	KRAS, STK11	RPMI-1640
H2126	Mutated	TP53	DMEM
H2228	Mutated	TP53	RPMI-1640
H23	Mutated	KRAS, TP53, STK11	RPMI-1640
H322	Mutated	TP53	RPMI-1640
H358	Wild type	KRAS	RPMI-1640
H441	Mutated	KRAS, TP53	RPMI-1640
H522	Mutated	TP53	RPMI-1640
H650	Mutated	KRAS, TP53	DMEM
H838	Mutated	TP53, STK11	RPMI-1640
HCC4006	Mutated	EGFR, TP53	RPMI-1640
HCC44	Wild type	KRAS, TP53, STK11	RPMI-1640

Cell line	SWI/SNF status*	Other mutated LUAD driver genes	Growth media
HCC827	Wild type	EGFR, TP53	RPMI-1640
LC319	Wild type	BRAF, TP53	RPMI-1640
LXF289	Mutated	TP53	RPMI-1640
PC14	Wild type	EGFR, TP53	RPMI-1640
PC9	Mutated	EGFR, TP53	RPMI-1640
SKLU1	Mutated	KRAS, TP53	DMEM

*We considered as 'Wild type' the cell line with any mutations in the 20 lung SWI/SNF subunits that we sequenced in our study.

Annex III: List of genes analyzed in this Ph.D. thesis

Supplementary Table 3: List of the lung SWI/SNF subunits and the top 10 LUAD driver genes. The genes are sorted in alphabetical order.

SWI/SNF genes	LUAD driver genes
ACTL6A	*ATM*
ARID1A	*BRAF*
ARID1B	*EGFR*
ARID2	*KEAP1*
BCL7A	*KRAS*
BCL7C	*NF1*
BRD9	*RB1*
DPF2	*RBM10*
DPF3	*STK11*
PBRM1	*TP53*
PHF10	
SMARCA2	
SMARCA4	
SMARCB1	
SMARCC1	
SMARCC2	
SMARCD1	
SMARCD2	
SMARCD3	
SMARCE1	

Annex IV: Performance of different mutational analysis approaches

To evaluate the performance of our non-paired analysis approaches, first we compared the results of the two non-paired pipelines. The agreement between the two approaches was high for primary tumors (***Supplementary Table 4***). We thoroughly examined the mutations uniquely found by one of the two approaches. Mutations detected by Mutect2 but not by our BCFtools-based approach were usually present at very low frequencies, as low as 1-2 total mutant reads. However, 17 mutations of Mutect2 showed high depth and high mutant allele frequency and we manually 'rescued' them for our final mutation list. The 'rescued' mutations were mostly indels, suggesting that Mutect2 has more power than BCFtools for indel detection. On the other hand, mutations found by our BCFtools-based pipeline but not by Mutect2 were all confirmed on IGV. Most of them had been flagged as 'germline' by Mutect2, possibly due to their presence in gnomAD, albeit at frequencies far below 0.01. We concluded that there was insufficient evidence to exclude such mutations. Therefore, for our final analysis, we combined the full results from our BCFtools-based pipeline with the mutations we 'rescued' from Mutect2.

We also compared the results from our nonpaired and paired analyses in our 27 tumor-normal pairs after the 'rescuing' step. Out of the 65 mutations from our nonpaired analysis, 56 (86%) were also found by the paired analysis. Only 13 of the 69 (19%) paired mutations were uniquely found by the paired analysis. They were all mutations present at very low frequencies (as low as 1-2 total mutant reads) and, therefore, they were filtered out by our non-paired analysis.

Supplementary Table 4: Agreement between the non-paired approaches applied in this study.

Primary tumors (N = 70)		
	# Mutations in common	**# Unique mutations**
BCFtools-based pipeline	105 (76%)	34 (24%)
Mutect2-based pipeline	105 (66%)	55 (34%)

Annex V: Primers used in this Ph.D. thesis

Supplementary Table 5: List of primers used for the characterization of the SWI/SNF complex.

Gene	Forward Sequence (5' to 3')	Reverse Sequence (5' to 3')
ACTL6A	TGGAGGCCATTTCACCTCTAA	TCTTTGCTCTAGTATTCCACGGT
ARID1A	TCCCAGCAAACTGCCTATTC	CATATCTTCTTGCCCTCCCTTAC
ARID1B	GGCCGTCCCGGAGTTTAATAA	CGGAGTGCATCATCCCCAT
ARID2	CAGTGTGTCGGATTATCTGCG	GCATGACGTGCTTGCTTTCATT
BCL7A	GGTGACCGTTGGTGACACAT	CACTTCTCGTCCTTGCCTTTTT
BCL7C	CGACACTTCCCTTCGTATCTTC	GAACCTTCCGAATGGAAACTCT
BRD9	GCAATGACATACAATAGGCCAGA	GAGCTGCCTGTTTGCTCATCA
DPF2	TATGCCTGTGACATTTGTGGAAAA	GAGTCTTCCTTGTCCTCGCC
DPF3	AAGAAGCGACGATTGCACC	AGCAAGGCTTCCAGTGTGGTAC
PBRM1	AGGAGGAGACTTTCCAATCTTCC	CTTCGCTTTGGTGCCCTAATG
PHF10	GCACTCTAGGCTTAACAGCATT	AGCATGTTTGGCTGGATATTCTT
SMARCA2	AGGGGATTGTAGAAGACATCCA	TTGGCTGTGTTGATCCATTGG
SMARCA4	AATGCCAAGCAAGATGTCGAT	GTTTGAGGACACCATTGACCATA
SMARCB1	GCGAGTTCTACATGATCGGCT	CACAGTGGCTAGTCGCCTC
SMARCC1	AGCTGTTTATCGACGGAAGGA	GCATCCGCATGAACATACTTCTT
SMARCC2	CCGTGACCCAGTTCGACAAC	CGGCAGTTTAGTGAGCGGT
SMARCD1	ATCCGGCTAAGTCAGATGCC	CAGATGGTTGTCTGGCCCAT
SMARCD2	GATCCATTCCGAAAACGCCTG	TGAGGTAGAACCTTATCTGCCA
SMARCD3	ACTGGATCAAACCATCATGCG	CAATGCTGCCGTCGGAATC
SMARCE1	CCTGCAACACAAATGCCCAG	TTTTGGAATCGTGATACCAGAGG
GAPDH	GAAGGTGAAGGTCGGAGTC	GAAGATGGTGATGGGATTTC

Supplementary Table 6: List of primers used for the analysis of miRNAs expression.

Primer name	Sequence (5' to 3')
Poly-T-adapter	GCGAGCACAGAATTAATACGACTCACTATAGGTTTTTTTTTTTT
miR-222-Fw	CTACATCTGGCTACTGGGTAAA
U6-Fw	CAAGGATGACACGCAAATTCG
Universal-Reverse	GCGAGCACAGAATTAATACGAC

Supplementary Table 7: List of primers used for the study of ARID1A.

Gene	Forward Sequence (5' to 3')	Reverse Sequence (5' to 3')
ATF3	GAAGGAGAAGACGGAGTGCC	GCTGCTTCTCGTTCTTGAGC
ATF5	CCTCCTCCTTCTCCACCTC	CTTCTTTTGCTTGCGGTCC
BAK1	AACTGGGCTCCCACTCAG	GACGGGATCAGCCTGCC
BAX	CCATCTTTGTGGCGGGAG	AACACAGTCCAAGGCAGC
BID	AAAGACATGAAGCCACGGG	GATGAGCTGAGCGTATGGC
DDIT3	CCGAGCTCTGATTGACCG	GGAAAGGTGGGTAGTGTGG
E2F1	GCCATCCAGGAAAAGGTGTG	GAAGCGCTTGGTGGTCAGAT
HRK	AGCGAGCAACAGGTTGGTG	CCTTCTCGAAGTGCCAACCG
JUN	GCCAGGTCGGCAGTATAGTC	TCTGGACACTCCCGAAACAC
c-MYC	CCACCAGCAGCGACTCTGAG	CCAGCAGAAGGTGATCCAGAC

Annex VI: SMARCA4 interactors of the SWI/SNF complex in lung epithelial cells

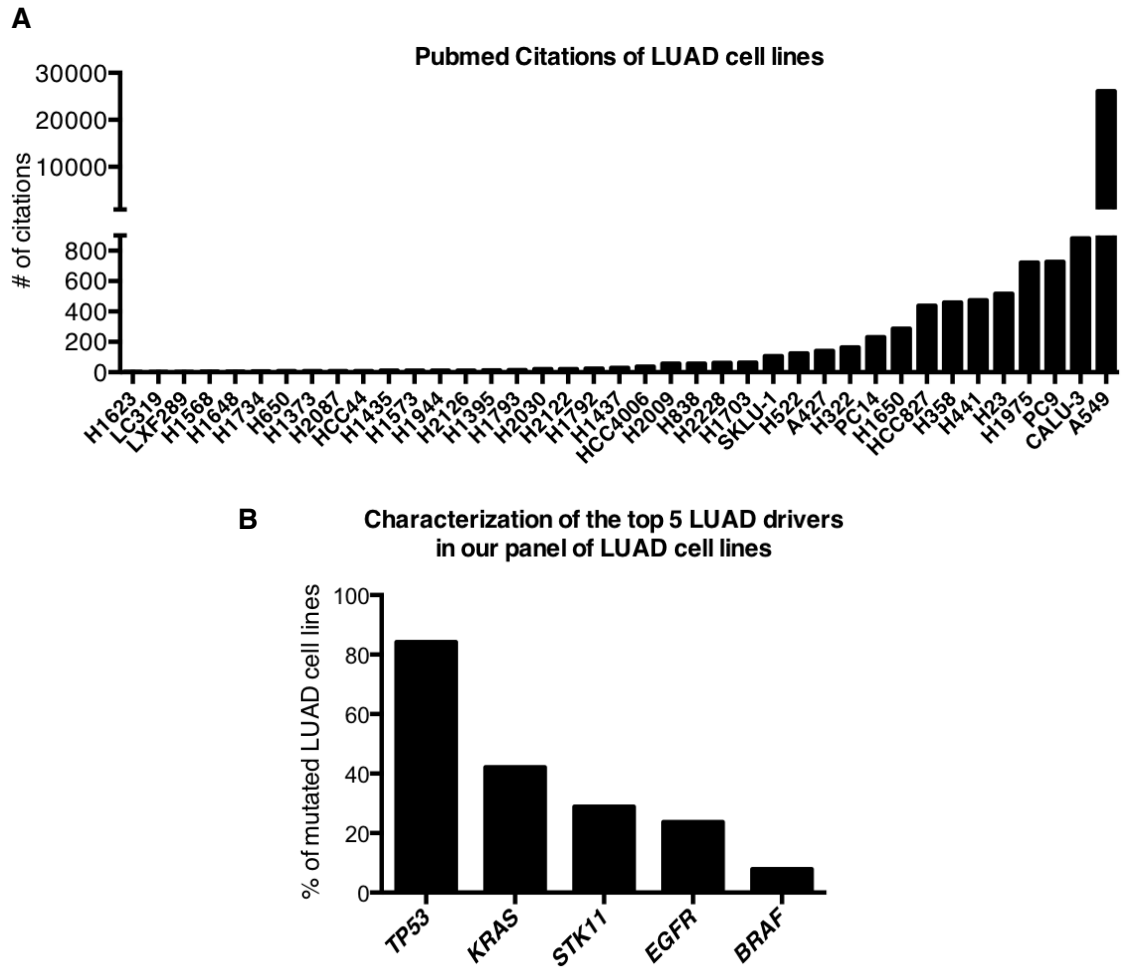
Supplementary Table 8: SWI/SNF subunits pulled down with SMARCA4 in NL20.

Proteins	Peptide average counts (rounded up)*
SMARCA4	46
SMARCC2	27
ACTL6A	21
ARID1A	21
SMARCC1	19
SMARCE1	19
SMARCD2	10
SMARCD1	8
PBRM1	6
SMARCB1	6
GLTSCR1	4
ARID2	3
DPF2	3
BRD7	3
BCL7C	2
ACTA1;ACTC1;ACTG2;ACTA2★	2
BRD9	2
DPF3	2
BCL7A	1
PHF10	1
ARID1B	1
SMARCD3	1

*Peptide number of all SWI/SNF subunits identified after pulling down SMARCA4 in NL20. An average value has been calculated using two biological replicates with two technical replicates each.

★Potential new subunits of the SWI/SNF complex in lung cells (as substitutes for ACTB).

Annex VII: Descriptive characterization of the panel of LUAD cell lines used in this Ph.D. thesis



Supplementary Figure 1: Descriptive analysis of the panel of LUAD cell lines used in this study. (**A**) Number of citations of each of the LUAD cell lines used in this study according to PubMed until September 2019. (**B**) Mutation profile of the top 5 LUAD driver genes identified by Bailey and colleagues in our panel of 38 LUAD cell lines.

Supplementary Table 9: Clinical features of our panel of 38 LUAD cell lines. Data provided by the ATCC (American Type Culture Collection) and Pubmed. (NA: Not available).

Disease	Adenocarcinoma	Broncoalveolar carcinoma	Carcinoma	Papillary adenocarcinoma
N=38	31	4	2	1
Tissue	Lung	Derivated from metastatic site	NA	
N=38	16	18	4	
Smoking status	Smokers	Non-smokers	NA	
N=38	15	7	16	

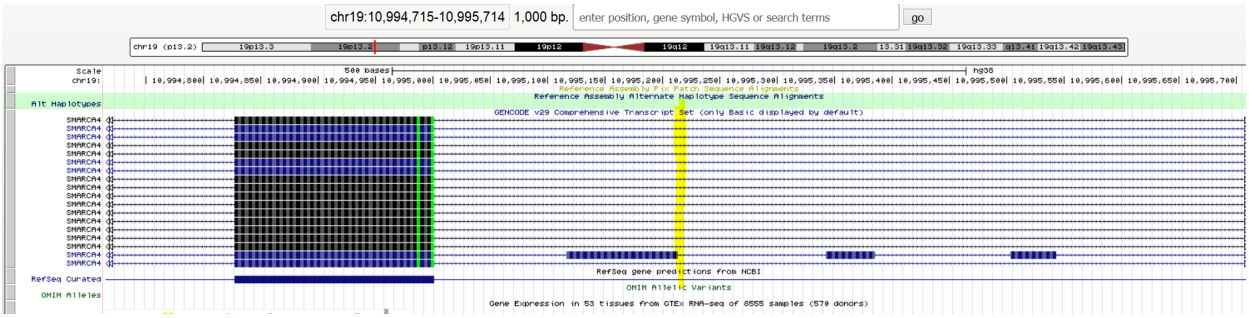
Sex	Male	Female	NA
N=38	22	14	2
Age	>45	<45	NA
N=38	23	7	8
Ethnicity	Caucasian	Black	NA
N=38	22	3	13

Annex VIII: CCLE comparison with our mutation data

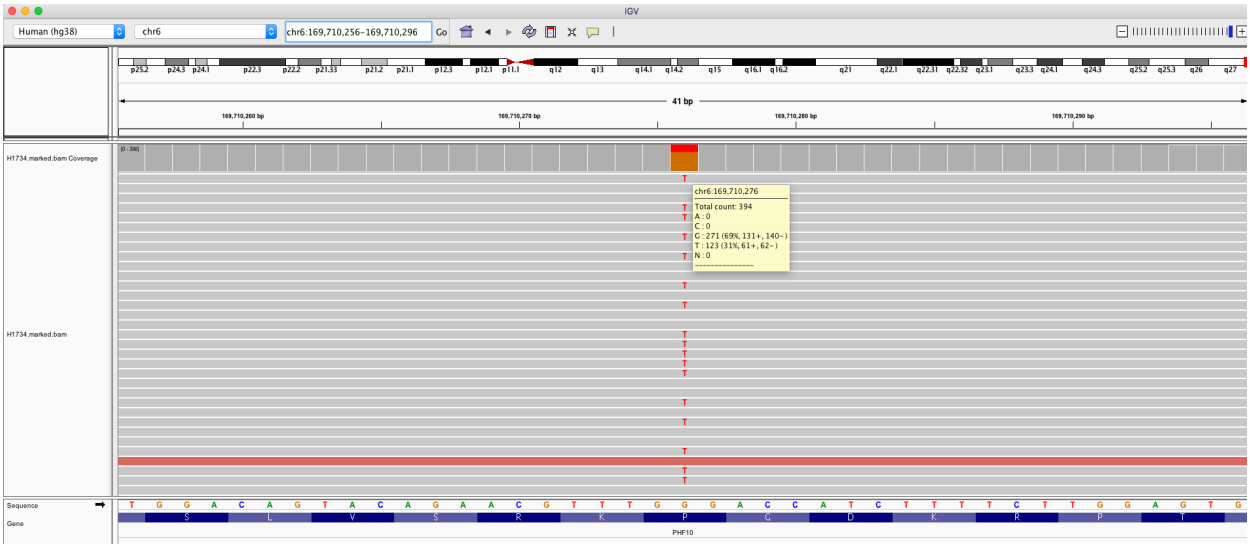
The sequencing data of 31 of our 38 LUAD cell lines was available in the CCLE database. Within these cell lines, 46 alterations were found in the 20 SWI/SNF subunits in the CCLE database in contrast with the 44 mutations identified in our study. In general, the results of our analysis were highly concordant with the mutations of CCLE. However, we detected and confirmed by Sanger sequencing three mutations that were not previously reported in CCLE: a splice site mutation in *SMARCA4* in H2122, which affects the isoform NM_001128845, and two non-synonymous mutations in *PHF10* and *BRD9* in H1734 and H1975, respectively (see ***Supplementary Fig. 2***).

On the other hand, in our analysis we did not find 9 mutations described in CCLE, after accounting for differences in the annotation reference (***Supplementary Fig. S3***). Four of these mutations were inframe indels or missense mutations that were filtered out in our analysis because they affected low complexity regions of *ARID1B* (Calu-3 and H1648), *ARID1A* (H441), or *SMARCC2* (H1623) (***Supplementary Fig. S4***). In addition, we did not find 5 non-synonymous mutations in *PBRM1* (PC14 and H1975), *SMARCC1* (A549), *SMARCD2* (H2122), and *DPF3* (H23). In the last case, we detected the mutation but it did not pass our filtering criteria due to low sequencing depth (***Supplementary Fig. S5***).

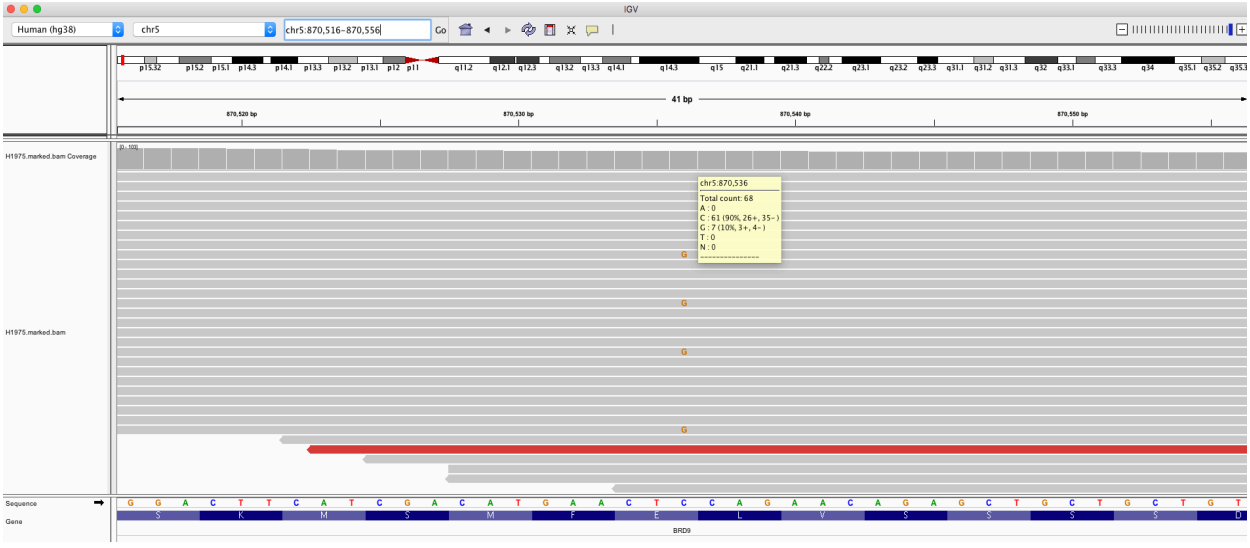
A


chr19:10,994,715-10,995,714
1,000 bp.
enter position, gene symbol, HGVS or search terms
go
Alt Haplotypes
SMARCA4
RefSeq Curated
OMIM Alleles
RefSeq gene predictions from NCBI
OMIM Allelic Variants

B

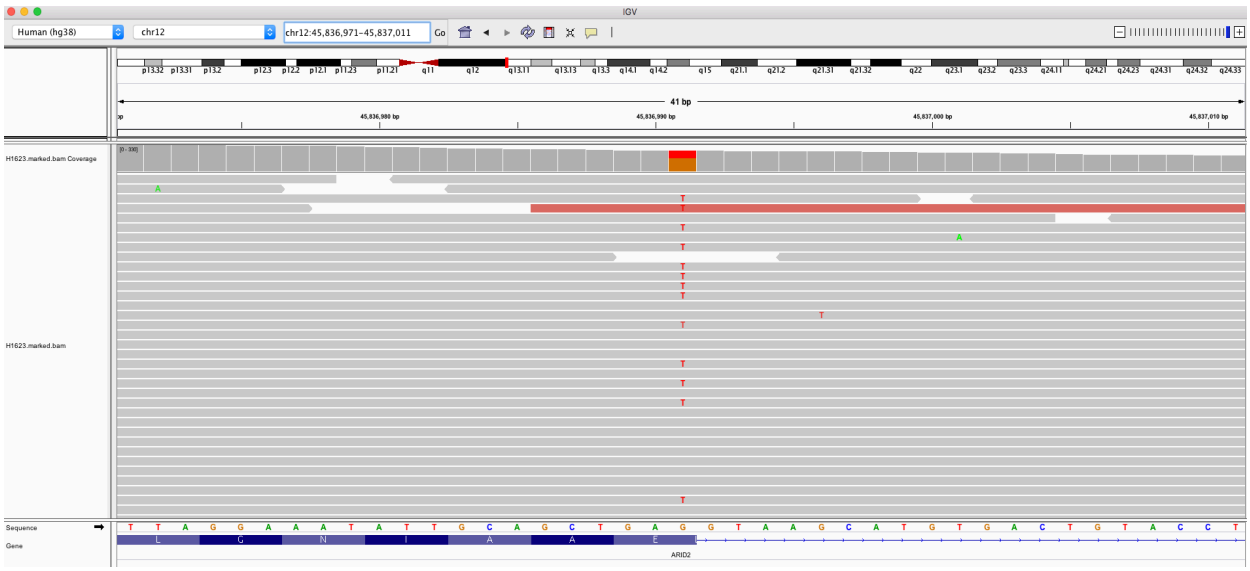

IGV
Human (hg38)
chr6
chr6:169,710,256-169,710,296
Go
41 bp
H1734.marked.bam Coverage
H1734.marked.bam
chr6:169,710,276
Total count: 394
A : 0
C : 0
G : 271 (69%, 131+, 140-)
T : 123 (31%, 61+, 62-)
N : 0
Sequence
Gene
PHF10

C

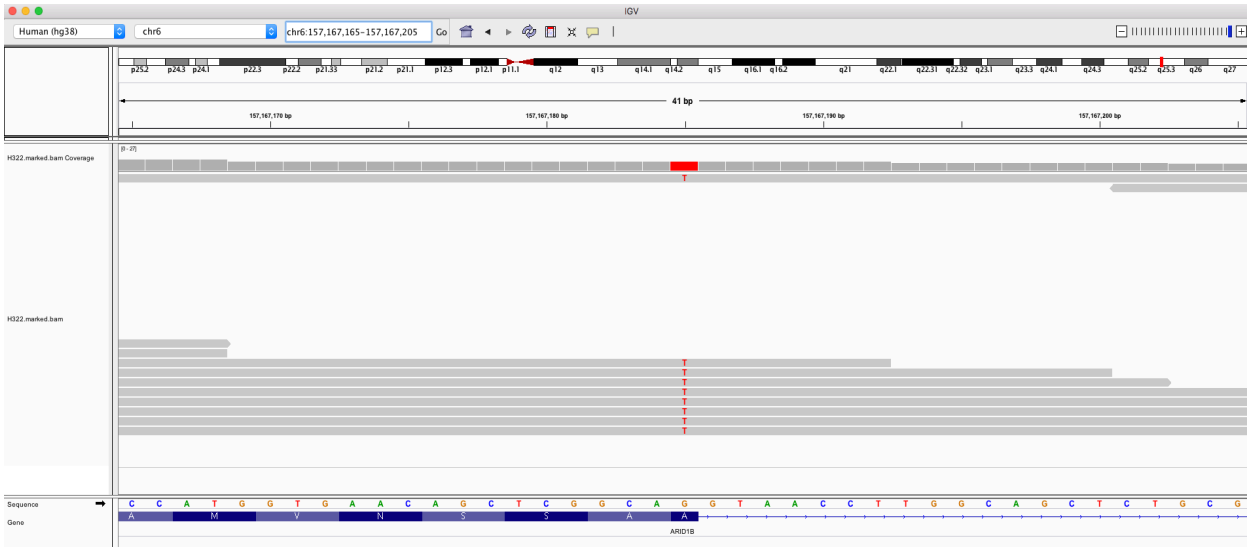

IGV
Human (hg38)
chr5
chr5:870,516-870,556
Go
41 bp
H1975.marked.bam Coverage
H1975.marked.bam
chr5:870,536
Total count: 68
A : 0
C : 61 (90%, 26+, 35-)
G : 7 (10%, 3+, 4-)
T : 0
N : 0
Sequence
Gene
BRD9

Supplementary Figure 2: *De novo* SWI/SNF complex mutations identified by our study in LUAD cell lines. (**A**) UCSC genome browser screenshot of the *SMARCA4* isoform that is mutated in the H2122 cell line (NM_001128845) according to our sequencing data. (**B**) *PHF10* mutation in the H1734 cell line (chr6:169710276 G-T). (**C**) *BRD9* mutation in the H1975 cell line (chr5:870536 C>G). All genomic coordinates are given with hg38 annotation.

A

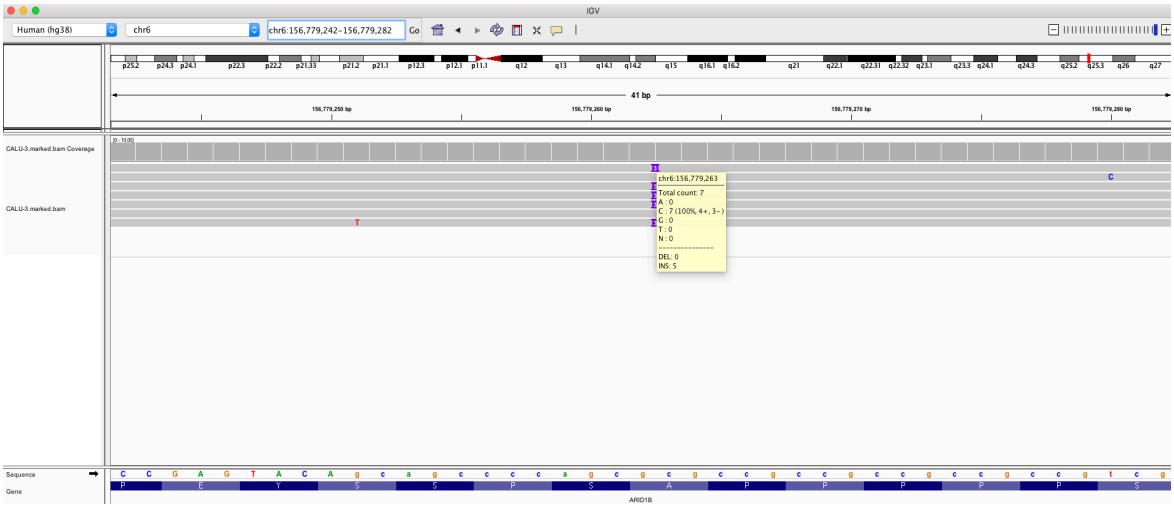


B

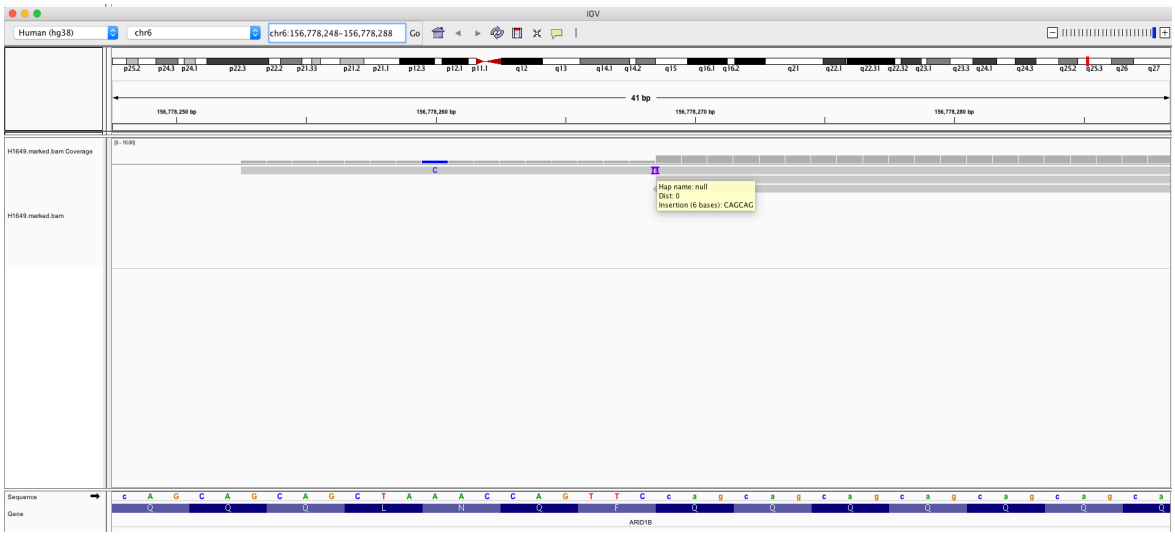


Supplementary Figure 3: SWI/SNF complex mutations with difference in annotation between CCLE and our data. (**A**) *ARID2* mutation in the H1623 cell line (chr12:45836991 G>T). (**B**) *ARID1B* mutation in the H322 cell line (chr6:157167185 G>T). All genomic coordinates are given with hg38 annotation.

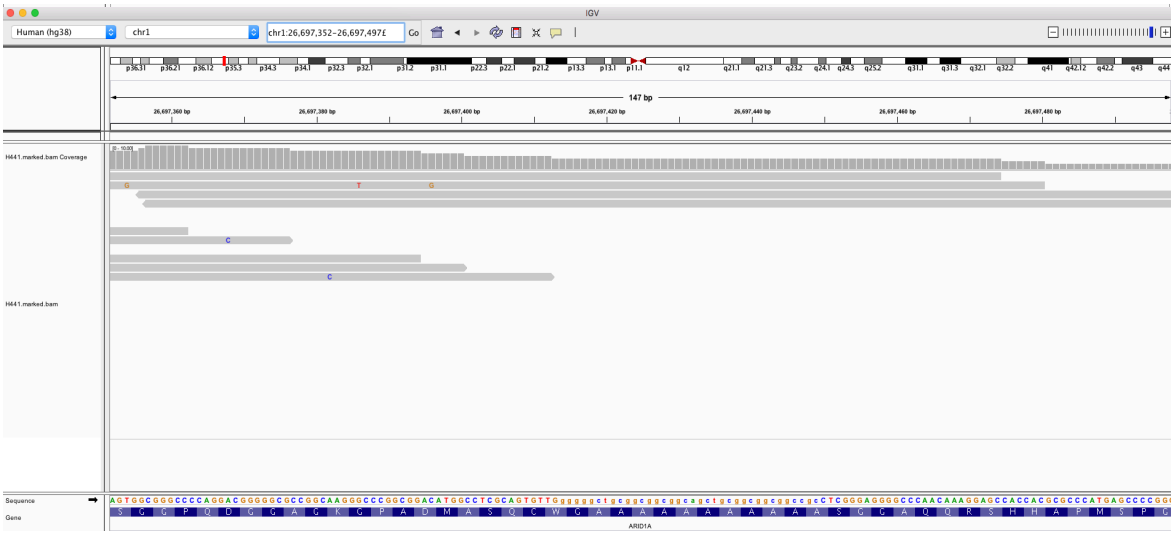
A


IGV
Human (hg38)
chr6
chr6:156,779,242-156,779,282
Go
41 bp
CALU-3.marked.bam Coverage
CALU-3.marked.bam
chr6:156,779,263
Total count: 7
A : 0
C : 7 (100%, 4+, 3-)
G : 0
T : 0
N : 0
DEL: 0
INS: 5
Sequence
Gene
ARID1B

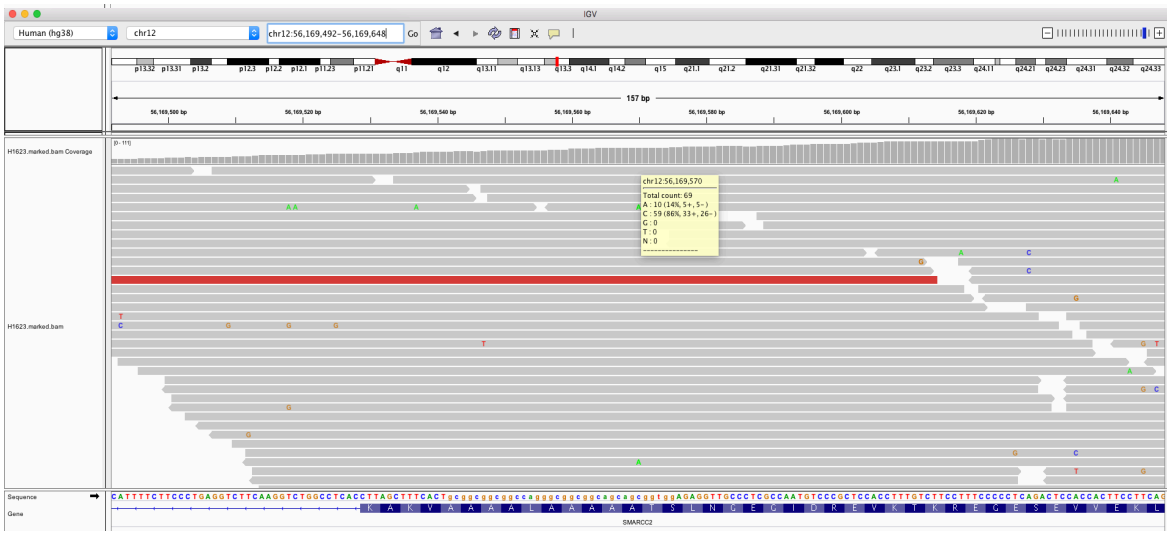
B


IGV
Human (hg38)
chr6
chr6:156,778,248-156,778,288
Go
41 bp
H1649.marked.bam Coverage
H1649.marked.bam
Hap name: null
Dist: 0
Insertion (6 bases): CAGCAG
Sequence
Gene
ARID1B

C

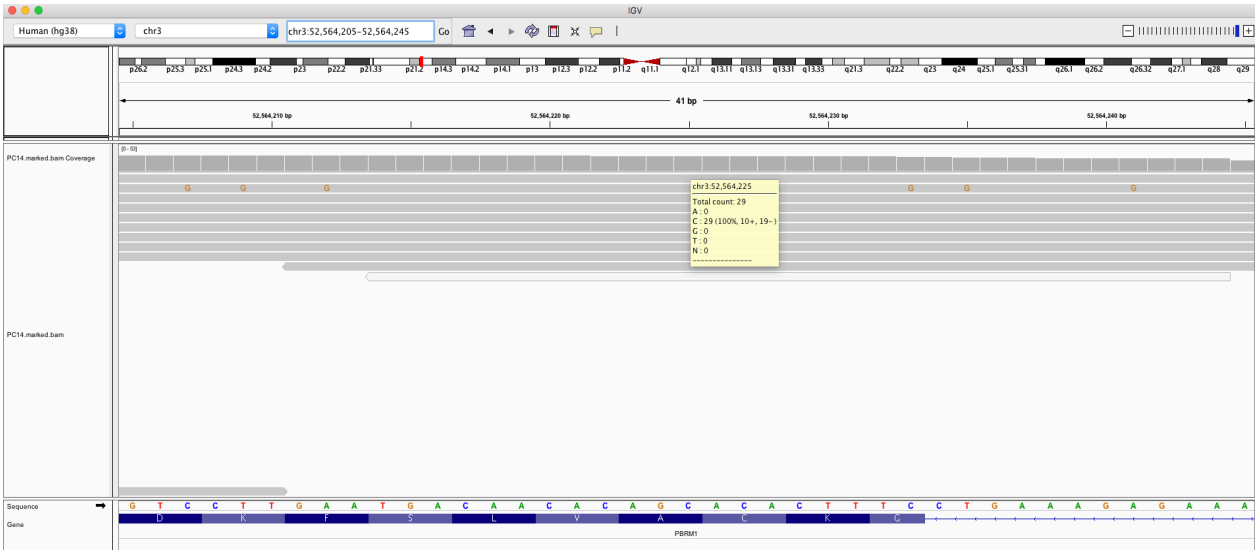

IGV
Human (hg38)
chr1
chr1:26,697,352-26,697,497
Go
147 bp
H441.marked.bam Coverage
H441.marked.bam
Sequence
Gene
ARID1A

D

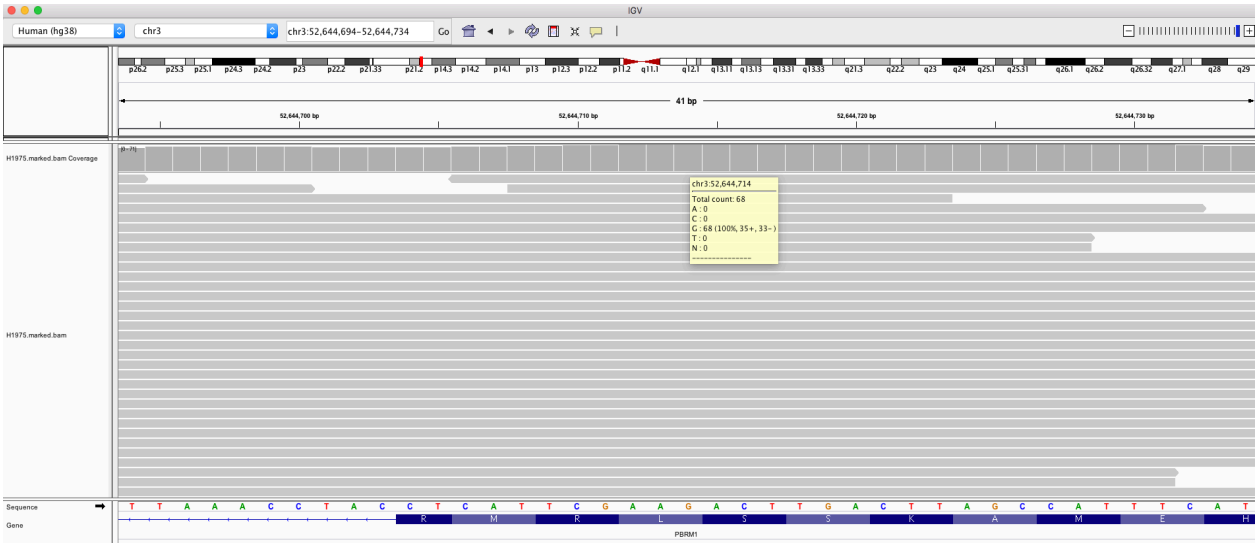


Supplementary Figure 4: CCLE mutations in the lung SWI/SNF complex excluded in our analysis because of affecting low complexity genomic regions. (**A**) *ARID1B* insertion in the Calu-3 cell line (chr6:156779262-156779263 ins CGC). (**B**) *ARID1B* insertion in the H1648 cell line (chr6:156778268-156778269 ins CAGCAG). (**C**) *ARID1A* deletion in the H441 cell line (chr1: 26697418-26697432 inframe del). (**D**) *SMARCC2* mutation in the H1623 cell line (chr12:56169570 C>A). All genomic coordinates are given with hg38 annotation.

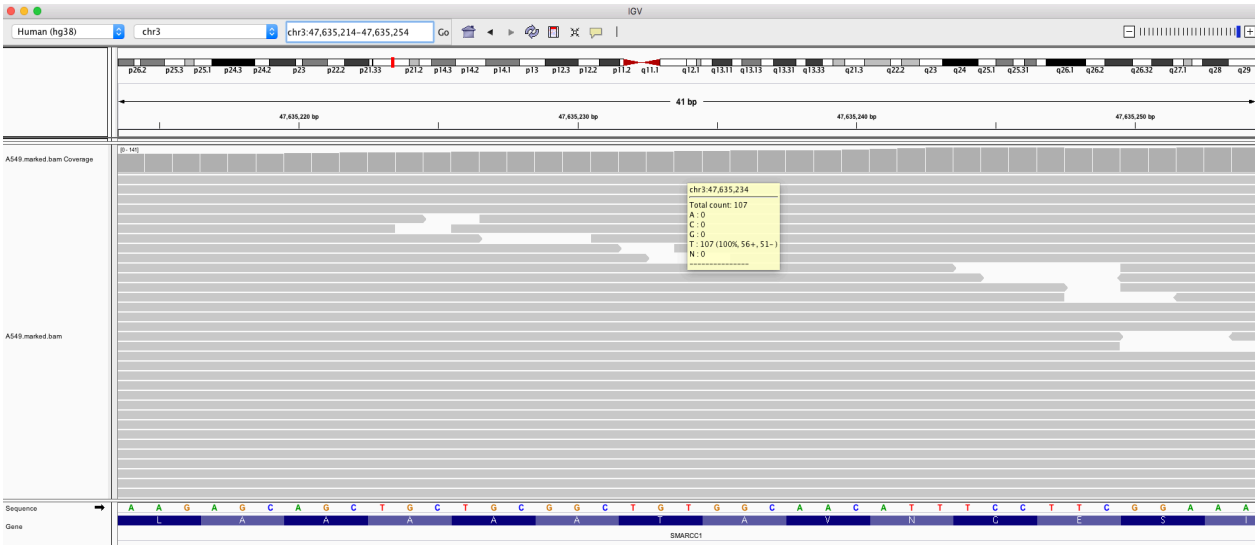
A


IGV
Human (hg38)
chr3
chr3:52,564,205-52,564,245
Go
41 bp
PC14.marked.bam Coverage
PC14.marked.bam
chr3:52,564,225
Total count: 29
A : 0
C : 29 (100%, 10+, 19-)
G : 0
T : 0
N : 0
Sequence
Gene
PBRM1

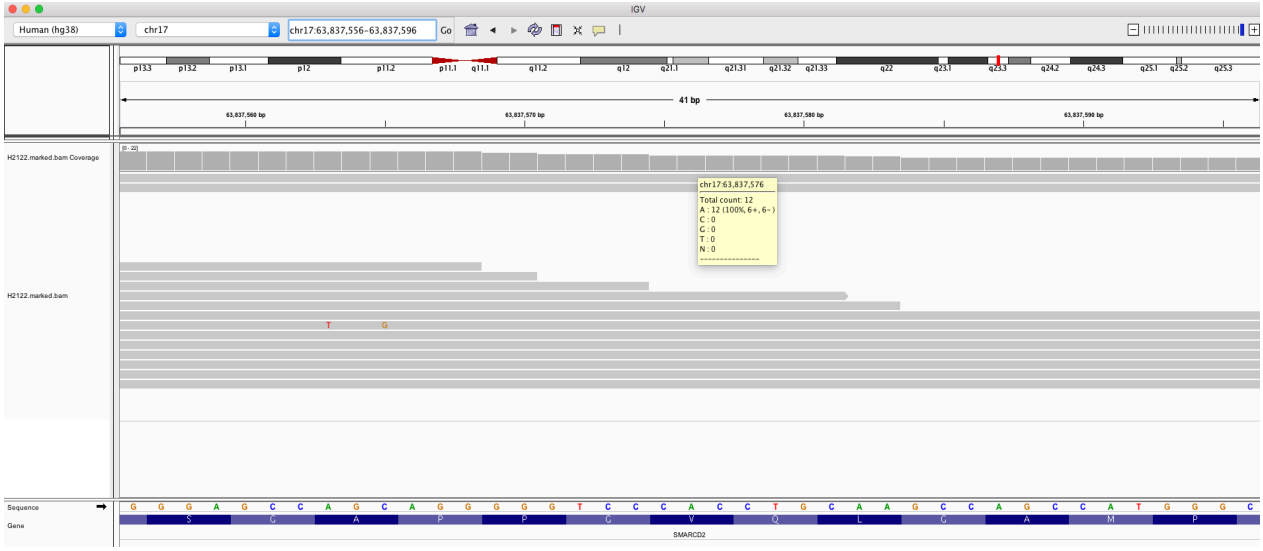
B


IGV
Human (hg38)
chr3
chr3:52,644,694-52,644,734
Go
41 bp
H1975.marked.bam Coverage
H1975.marked.bam
chr3:52,644,714
Total count: 68
A : 0
C : 0
G : 68 (100%, 35+, 33-)
T : 0
N : 0
Sequence
Gene
PBRM1

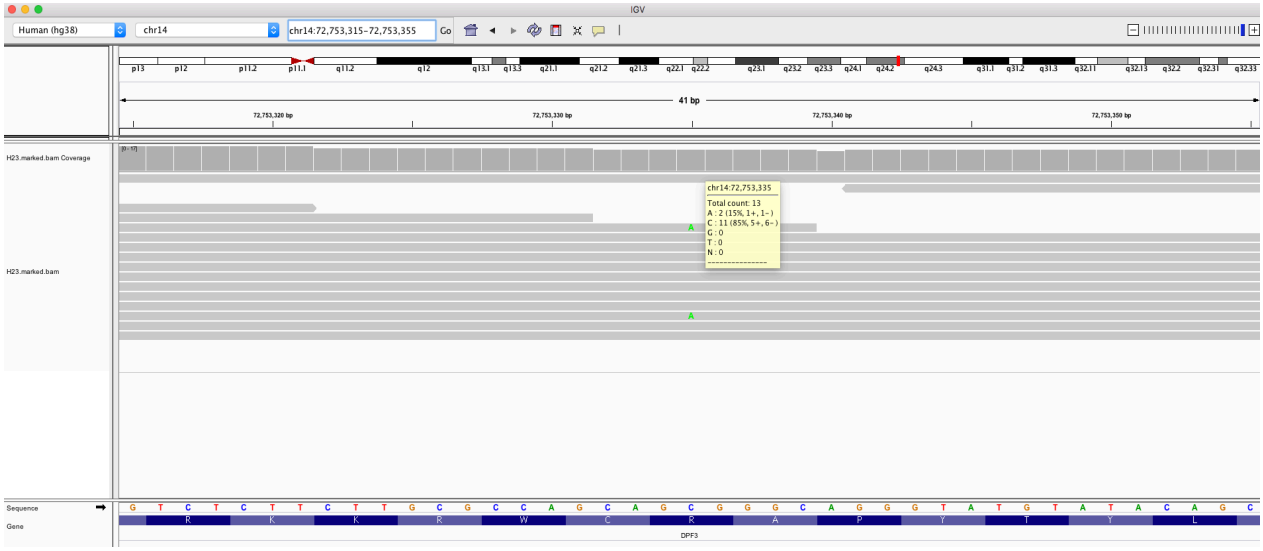
C


IGV
Human (hg38)
chr3
chr3:47,635,214-47,635,254
Go
41 bp
A549.marked.bam Coverage
A549.marked.bam
chr3:47,635,234
Total count: 107
A : 0
C : 0
G : 0
T : 107 (100%, 56+, 51-)
N : 0
Sequence
Gene
SMARCC1

D

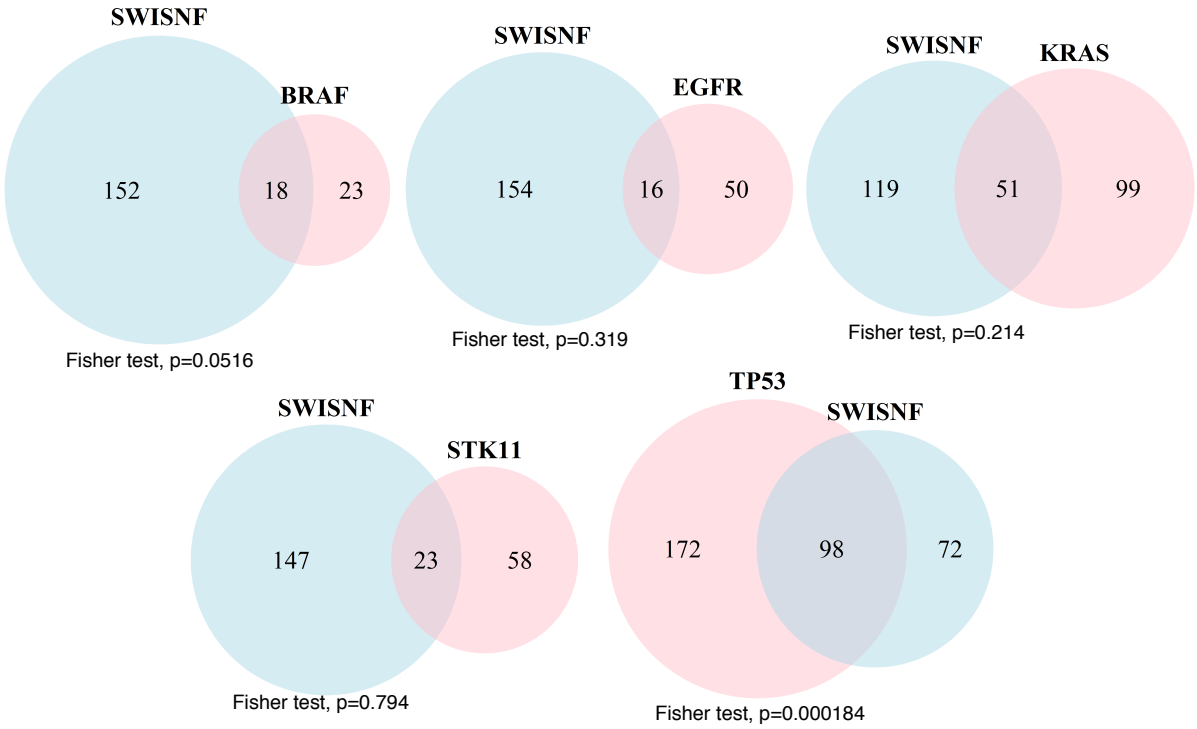


E



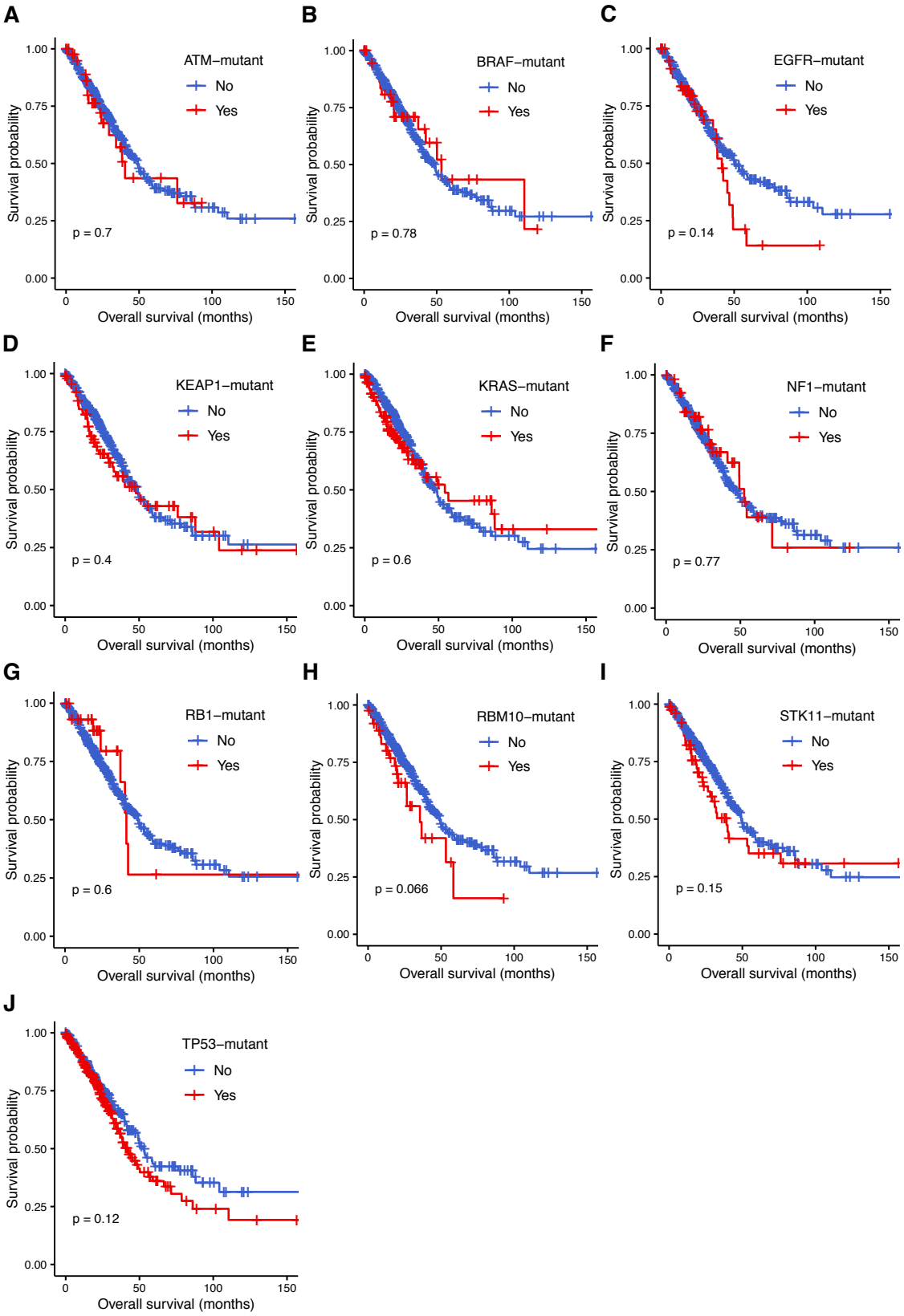
Supplementary Figure 5: Lung SWI/SNF complex mutations exclusively described in CCLE. (**A**) *PBRM1* mutation in the PC14 cell line (chr3:52564225 C>T). (**B**) *PBRM1* mutation in the H1975 cell line (chr3:52644714 G>T). (**C**) *SMARCC1* mutation in the A549 cell line (chr3:47635234 T>G). (**D**) *SMARCD2* mutation in the H2122 cell line (chr17:63837576 A>C). (**E**) *DPF3* mutation in the H23 cell line (chr14:72753335 C>A). All genomic coordinates are given with hg38 annotation.

Annex IX: Co-occurrence and mutual exclusion analysis



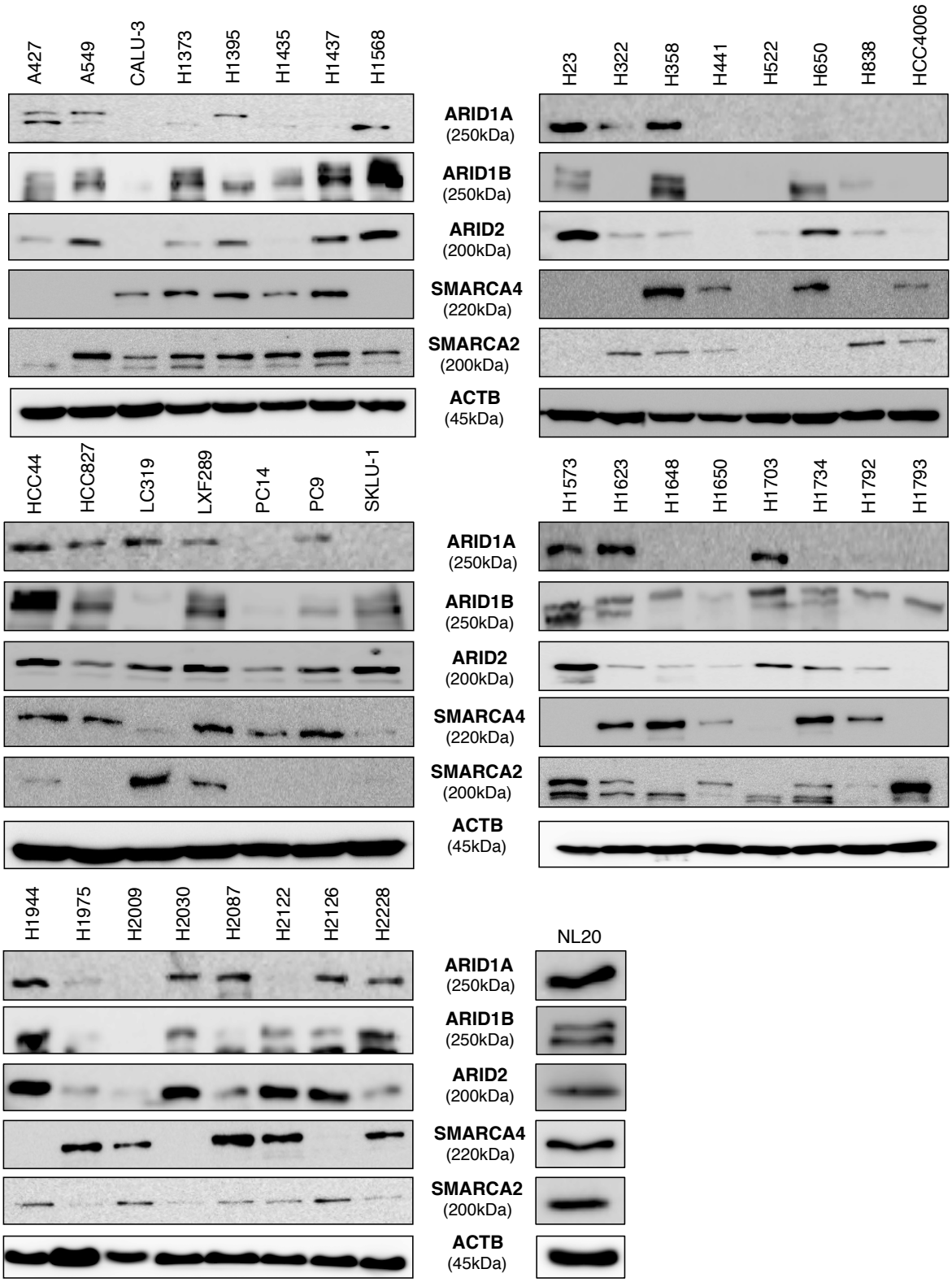
Supplementary Figure 6: Co-occurrence and mutual exclusion analysis of SWI/SNF mutations and the top 5 LUAD driver genes. Venn Diagrams with the mutational status of the top 5 LUAD driver genes and the SWI/SNF complex in LUAD patients using TCGA-LUAD data.

Annex X: Overall survival analysis with the top 10 LUAD driver genes



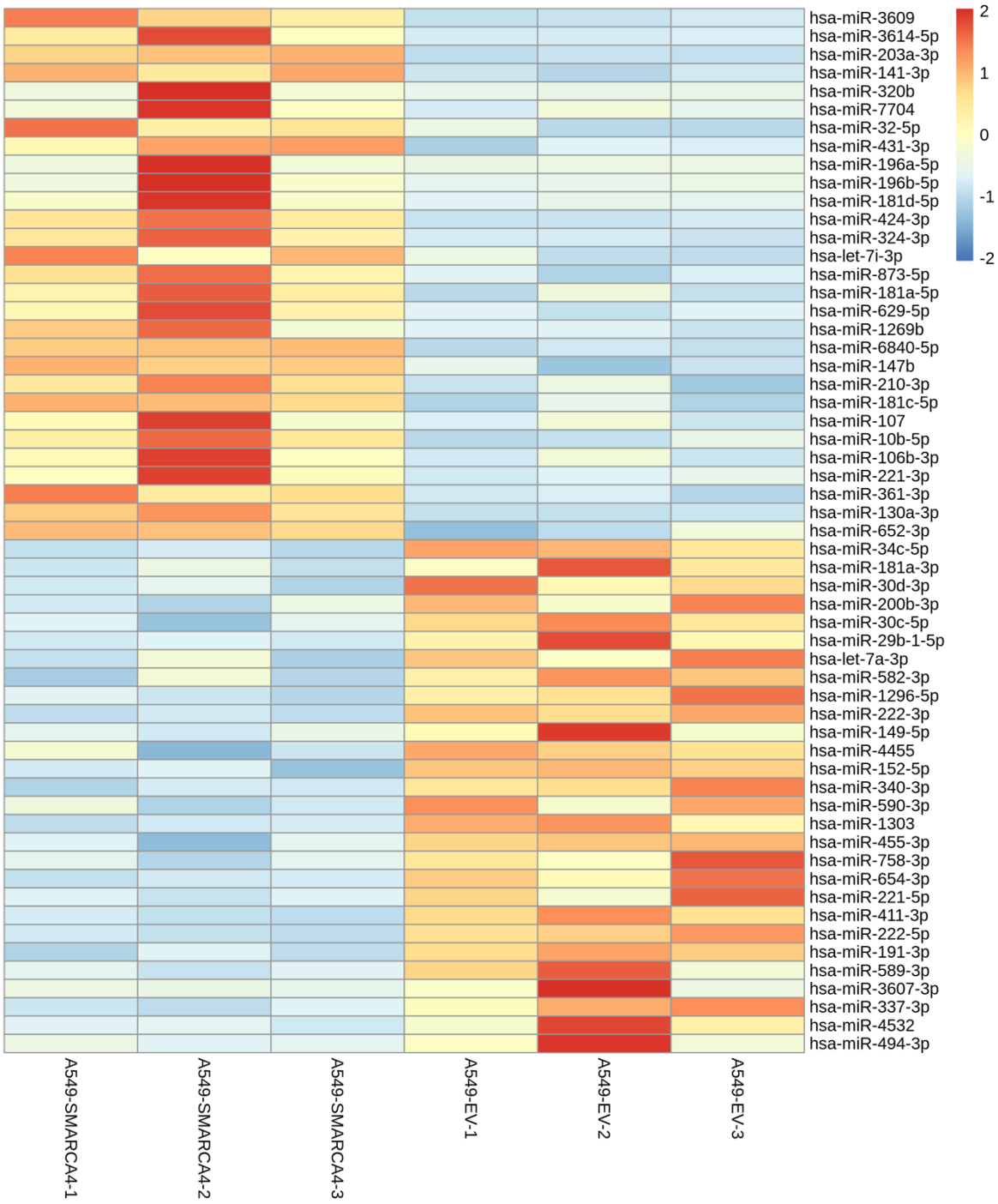
Supplementary Figure 7: Analysis of overall survival of LUAD patients with different mutational status of the top 10 LUAD driver genes. (**A-J**) Kaplan-Meier curves grouping the TCGA-LUAD cohort by the mutational status of each of the top 10 LUAD driver genes.

Annex XI: Western blot analysis of the ATPases and ARID subunits of the SWI/SNF complex in our panel of LUAD cell lines



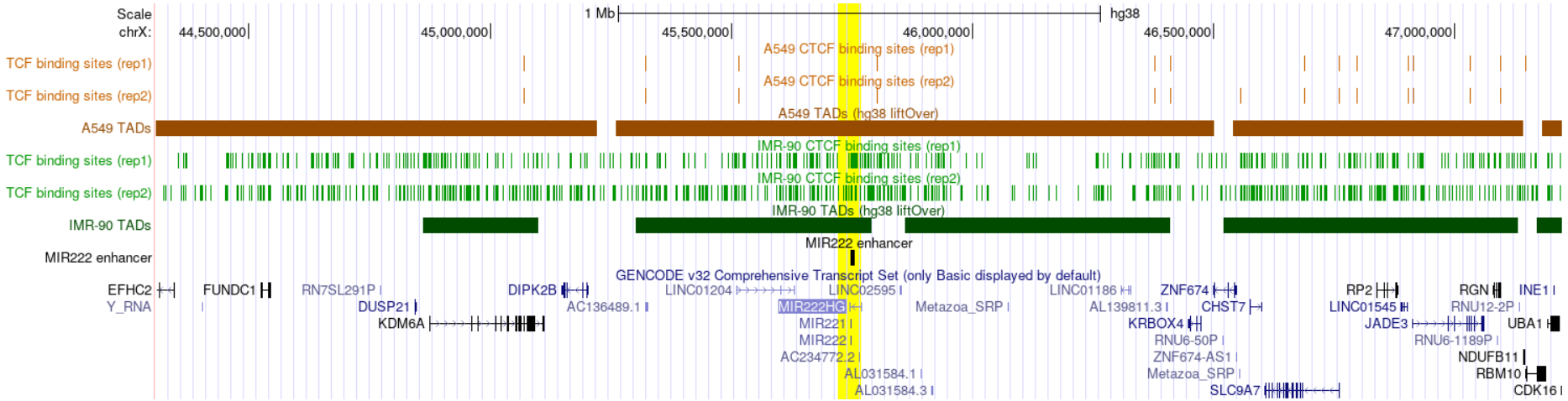
Supplementary Figure 8: Western blot of SMARCA4, SMARCA2, ARID1A, ARID1B, and ARID2 in our 38 LUAD cell lines and a normal lung cell line (NL20). The sizes depicted in the image correspond to the observed molecular weights of each of these proteins. ACTB was used as a loading control.

Annex XII: Differentially expressed miRNAs upon SMARCA4 restoration



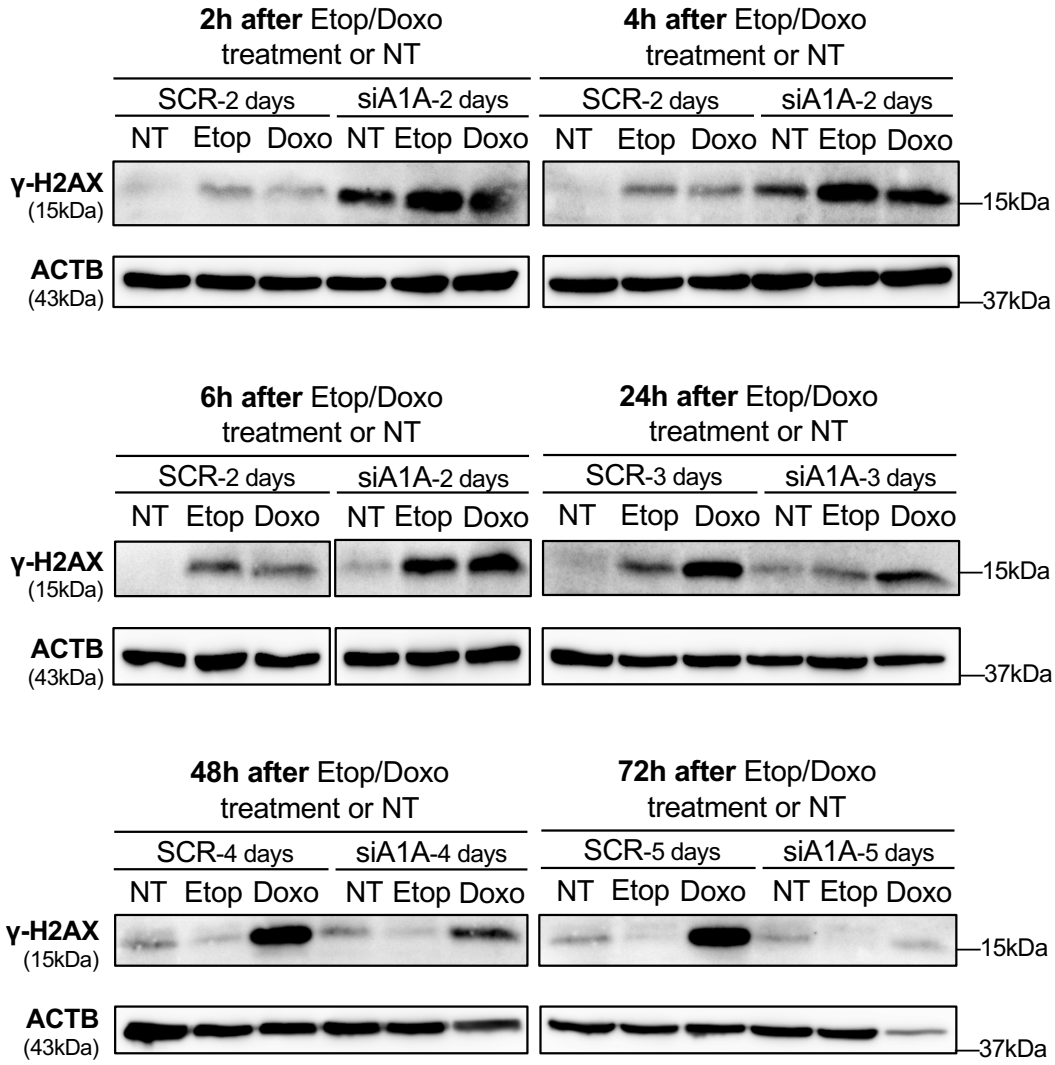
Supplementary Figure 9: Differentially expressed microRNAs upon SMARCA4 restoration in A549 cells. Heatmap of significantly dysregulated miRNAs after 48h of transfection of A549 with empty vector (EV) or SMARCA4 plasmid (SMARCA4). Only miRNAs with an FDR < 0.05 and an absolute log_2Fold-Change > 1 are represented. Z-score values are plotted.

Annex XIII: Three-dimensional organization of the genomic region around the miR-222 enhancer



Supplementary Figure 10: Topologically associating domains (TADs) and CTCF binding sites near the miR-222 enhancer. ENCODE data were analyzed for the A549 and IMR-90 cell lines to determine the three-dimensional organization of the genomic region around the miR-222 enhancer. The TADs were derived from high-throughput chromosome conformation capture (Hi-C) data with a resolution of 40 kb for A549 and 5 kb for IMR-90. The CTCF binding data had two replicates per cell line. An area around the miR-222 enhancer is highlighted in yellow.

Annex XIV: Western blot analysis of γ-H2AX after the treatment of A549 cell line with genotoxic agents



Supplementary Figure 11: Western blot analysis of the DNA damage marker γ-H2AX in the A549 cell line. After 48 h of the transfection with a nonsense scrambled siRNA (SCR) or an siRNA specific against ARID1A (siA1A), the cells were treated with either etoposide 10μM (Etop), doxorubicin 0.5μM (Doxo), or not treated (NT). The cells were collected at different time points and the γ-H2AX expression was analyzed using ACTB as a loading control.

Annex XV: Ph.D. article and other related publications

Human Molecular Genetics, 2021, Vol. 00, No. 00 1–9

https://doi.org/10.1093/hmg/ddab187
Advance Access Publication Date: 8 July 2021
General Article

GENERAL ARTICLE

The SWI/SNF complex regulates the expression of miR-222, a tumor suppressor microRNA in lung adenocarcinoma

Paola Peinado[1,2], Alvaro Andrades[1,2], Jordi Martorell-Marugán[2], Jeffrey R. Haswell[3], Frank J. Slack[3,4], Pedro Carmona-Sáez[2,5] and Pedro P. Medina[1,2,6,*]

[1]Department of Biochemistry and Molecular Biology I, University of Granada, Granada 18071, Spain, [2]GENYO, Centre for Genomics and Oncological Research, Pfizer/University of Granada/Andalusian Regional Government, Granada 18016, Spain, [3]Department of Pathology, Cancer Center, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA, [4]Harvard Medical School Initiative for RNA Medicine, Boston, MA 02215, USA, [5]Department of Statistics, University of Granada, Granada 18071, Spain and [6]Health Research Institute of Granada (ibs.Granada), Granada 18012, Spain

*To whom correspondence should be addressed at: Department of Biochemistry and Molecular Biology I, University of Granada, Campus Fuentenueva s/n 18071 Granada, Spain. Tel: +34 958243252; Fax: +34 958249945; Email: pedromedina@ugr.es

Abstract

SWitch/Sucrose Non-Fermentable (SWI/SNF) chromatin remodeling complexes are key epigenetic regulators that are recurrently mutated in cancer. Most studies of these complexes are focused on their role in regulating protein-coding genes. However, here, we show that SWI/SNF complexes control the expression of microRNAs. We used a SMARCA4-deficient model of lung adenocarcinoma (LUAD) to track changes in the miRNome upon SMARCA4 restoration. We found that SMARCA4-SWI/SNF complexes induced significant changes in the expression of cancer-related microRNAs. The most significantly dysregulated microRNA was miR-222, whose expression was promoted by SMARCA4-SWI/SNF complexes, but not by SMARCA2-SWI/SNF complexes via their direct binding to a miR-222 enhancer region. Importantly, miR-222 expression decreased cell viability, phenocopying the tumor suppressor role of SMARCA4-SWI/SNF complexes in LUAD. Finally, we showed that the miR-222 enhancer region resides in a topologically associating domain that does not contain any cancer-related protein-coding genes, suggesting that miR-222 may be involved in exerting the tumor suppressor role of SMARCA4. Overall, this study highlights the relevant role of the SWI/SNF complex in regulating the non-coding genome, opening new insights into the pathogenesis of LUAD.

Introduction

Gene expression regulation is essential for an adequate execution of all biological processes and alterations in these regulatory mechanisms can lead to tumor formation and progression. Indeed, the loss of epigenetic control of gene expression is a key hallmark of cancer (1). Among all epigenetic changes, chromatin remodeling has spurred great interest in the field. Several studies have shown that the SWitch/Sucrose Non-Fermentable (SWI/SNF) complex, a member of the family of ATP-dependent chromatin remodeling complexes, is mutated in nearly 25% of

Received: April 12, 2021. **Revised:** June 25, 2021. **Accepted:** July 5, 2021

Other articles or book chapters published during this Ph.D.

Peinado P*, Andrades A*, Cuadros M*, Rodriguez MI*, Coira IF, Garcia DJ, Benitez-Cantos MS, Cano C, Rufino-Palomares E, Zarzuela E, Muñoz J, Loidi C, Saiz M, Medina PP. Integrative multi-omic analysis of the SWI/SNF complex defines a new clinical subgroup of lung adenocarcinoma patients. *Under Review*.

Haswell JR*, Mattioli K*, Gerhardinger C, Maass PG, Foster DJ, **Peinado P**, Wang X, Medina PP, Rinn JL, Slack FJ. Genome-wide CRISPR interference screen identifies long non-coding RNA loci required for differentiation and pluripotency. *PLoS One*. 2021 Nov 3;16(11):e0252848.

Cuadros M*, García DJ*, Andrades A*, Arenas AM, Coira IF, Baliñas-Gavira C, **Peinado P**, Rodríguez MI, Álvarez-Pérez JC, Ruiz-Cabello F, Camós M, Jiménez-Velasco A, Medina PP. LncRNA-mRNA Co-Expression Analysis Identifies AL133346.1/CCN2 as Biomarkers in Pediatric B-Cell Acute Lymphoblastic Leukemia. *Cancers* (Basel). 2020 Dec 17;12(12):3803.

Peinado P*, Andrades A*, Cuadros M*, Rodriguez MI*, Coira IF, Garcia DJ, Álvarez-Perez JC, Baliñas-Gavira C, Arenas AM, Patiño-Mercau JR, Sanjuan-Hidalgo J, Romero OA, Montuenga LM, Carretero J, Sanchez-Cespedes M, Medina PP. Comprehensive Analysis of SWI/SNF Inactivation in Lung Adenocarcinoma Cell Models. *Cancers* (Basel). 2020 Dec 10;12(12):E3712.

Arenas AM*, Cuadros M*, Andrades A, García DJ, Coira IF, Rodríguez MI, Baliñas-Gavira C, **Peinado P**, Álvarez-Pérez JC, Medina PP. LncRNA DLG2-AS1 as a Novel Biomarker in Lung Adenocarcinoma. *Cancers* (Basel). 2020 Jul 28;12(8):2080.

Martin-Padron J*, Boyero L*, Rodriguez MI, Andrades A, Díaz-Cano I, **Peinado P**, Baliñas-Gavira C, Alvarez-Perez JC, Coira IF, Fárez-Vidal ME, Medina PP. Plakophilin 1 enhances MYC translation, promoting squamous cell lung cancer. *Oncogene*. 2020 Aug;39(32):5479-5493.

Cuadros M*, Andrades A*, Coira IF, Baliñas C, Rodríguez MI, Álvarez-Pérez JC, **Peinado P**, Arenas AM, García DJ, Jiménez P, Camós M, Jiménez-Velasco A, Medina PP. Expression of the long non-coding RNA TCL6 is associated with clinical outcome in pediatric B-cell acute lymphoblastic leukemia. *Blood Cancer* J. 2019 Nov 25;9(12):93.

Peinado P*, Herrera A*, Baliñas C, Martín-Padrón J, Boyero-Corral L; Cuadros Marta, Coira IF, Rodríguez-Lara MI, Reyes-Zurita F, Rufino-Palomares EE, Lupiañez JA, Medina PP: Long Non-Coding RNAs as cancer biomarkers. Chapter of the book: Cancer and Noncoding RNAs. *Elsevier*. 2018 Vol 1: 95-114. ISBN: 9780128110225.

Cuadros M*, Sánchez-Martín V*, Herrera A, Baliñas C, Martín-Padrón J, Boyero L, **Peinado P**, Medina PP. BRG1 regulation by miR-155 in human leukemia and lymphoma cell lines. *Clin Transl Oncol*. 2017 Aug;19(8):1010-1017.

** Contributed equally to this work*

Annex XVI: Copyright permissions

Figure 3

License Number	5186930474654
License date	Nov 13, 2021

Licensed Content

Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell
Licensed Content Title	Mapping the Hallmarks of Lung Adenocarcinoma with Massively Parallel Sequencing
Licensed Content Author	Marcin Imielinski,Alice H. Berger,Peter S. Hammerman,Bryan Hernandez,Trevor J. Pugh,Eran Hodis,Jeonghee Cho,James Suh,Marzia Capelletti,Andrey Sivachenko,Carrie Sougnez,Daniel Auclair,Michael S. Lawrence,Petar Stojanov,Kristian Cibulskis et al.
Licensed Content Date	Sep 14, 2012
Licensed Content Volume	150
Licensed Content Issue	6
Licensed Content Pages	14

Order Details

Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables /illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No

About Your Work

Title	Implications of the chromatin remodeling complex SWI/SNF in lung cancer
Institution name	University of Granada
Expected presentation date	Jan 2022

Additional Data

Portions	Figure 6

Figure 8