universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Identification of SMARCA4 as a synthetic lethal dependency in SMARCA2^{low} esophageal squamous cell carcinoma cell lines“

verfasst von / submitted by

Katharina Ehrenhöfer-Wölfer, MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 685 490
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Molekulare Biologie
Betreut von / Supervisor:	Dr. Wolfgang Sommergruber, Privatdoz.

Acknowledgements

Firstly, I would like to thank my supervisor Simon Wöhrle who devised the main conceptual idea. After the identification of many targets from the screen, he significantly helped me to pursue the most interesting one. Simon offered me a lot of opportunities to further deepen my scientific knowledge including an internship at CSHL as well as the attendance of global conferences. The time he spent in supervising and discussing experiments has been truly appreciated.

Another great thanks belong to my supervisor, Wolfgang Sommergruber, for the support and the feedback on my project. I would like to thank him and my additional PhD committee members, Christian Seiser and Helmut Dolznig, for discussing the project in annual progress meetings. In addition, I would like to thank Mark Pearson, Manfred Kögl, Mark Petronczki, Ralph Neumüller, Norbert Schweifer, Andreas Schlattl as well as Thomas Zichner for their scientific input. Many thanks to the lab members Janine Rippka, Ursula Strobl, Teresa Puchner, Cornelia Schwarz and Silvia Blaha-Ostermann who strongly supported me by teaching me diverse methods, helping me with experiments and keeping the lab organized. Special thanks to Teresa Puchner and Annika Oßwald for always taking time for discussions and moral support. A great thanks belongs to all the colleagues I shared an office with, making every day memorable. Last but not least, I would like to thank all the lab scientists at Boehringer Ingelheim for being extremely open and helpful by sharing all their expertise and protocols. I always felt comfortable working with them. I would like to thank Boehringer Ingelheim RCV for the financial support of my PhD project and the MFPL for giving me the opportunity in participating in the “SMICH” program. I was pleased to be given the opportunity to spend five weeks at CSHL and would like to thank Christopher Vakoc, Junwei Shi as well as Yuhan Huang for sharing their technique on CRISPR-Cas9 domain-based screening. Thanks to Alexandra Hörmann for supporting me in New York and transferring the screen to Vienna.

Furthermore, I would like to thank all of my friends, especially Lisa Nika for supporting and motivating me during the whole time as a PhD student. I would also like to thank my family, my mother and father especially for their moral and financial support throughout my whole education. A big thanks belongs to my husband Jürgen, who kept me motivated during the most stressful moments with his unconditional support.

Table of Contents

1 Introduction

1.1 Esophageal squamous cell carcinoma (ESCC)

Esophageal carcinoma (EC) is the sixth leading cause of deaths due to cancer worldwide (509.000 deaths), ranking seventh in terms of incidences (572.000 new cases) (Figure 1) [1, 2] . The 5-year survival rate is 15-25% [3]. Remarkably, approximately 70% of all cases occur in men and variations of incidence and mortality numbers are observed between regions [1]. Early diagnosis of the malignancy without progression allows the best outcome, whereas the outcome for patients with progressed tumor stages remains poor with a median survival of only six months upon chemotherapy [4, 5]. Due to the limited therapy options and the late diagnosis, EC is a malignancy with a high medical need [6]. Esophageal squamous cell carcinoma (ESCC) is the predominant form of EC, comprising 90% of all cases. The less frequently detected form is esophageal adenocarcinoma (EAC) and uncommon subtypes are represented by melanoma, leiomyosarcoma as well as small cell carcinoma [1, 7].

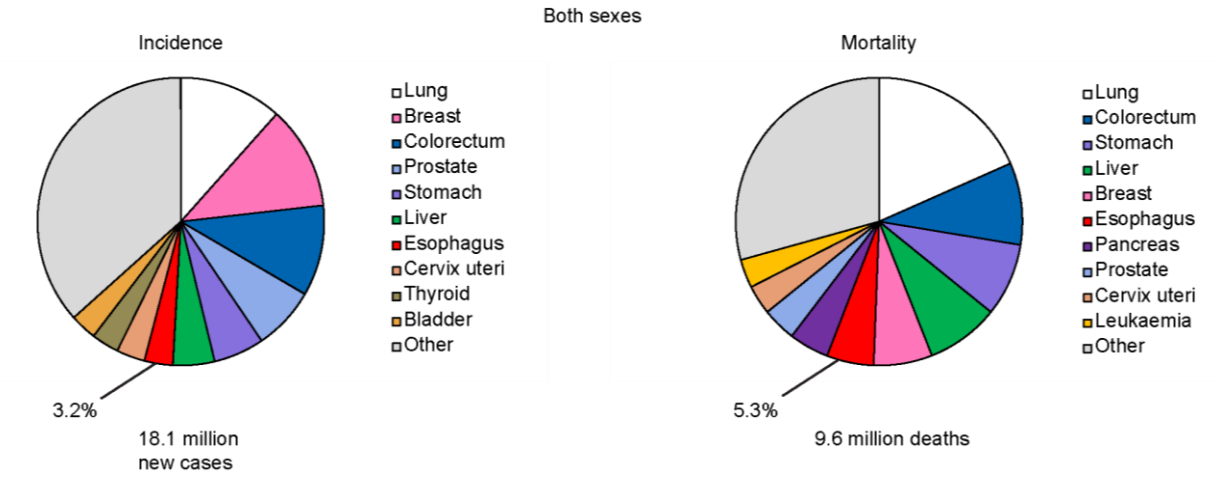


Figure 1: Pie distribution of incidence and mortality numbers of diverse cancers.

Incidence (left) and mortality (right) proportions; graphic was adopted from Bray 2018 [1].

1.1.1 ESCC – an indication with a high medical need

ESCC is one of the most aggressive squamous cell carcinomas [8] and over several decades, the outcome of ESCC remains unchanged [3]. This is due to the fact that upon diagnosis, the disease has already progressed by invading surrounding organs and developing distant metastases. Additionally, poor outcome rates are supported by the lack of biomarkers [9] and sensitive methods. Progress in early esophageal neoplastic lesions detection has been made via an image-enhanced endoscopy [10]. The identification of a new molecular biomarker would improve patient's prognosis, outcome and treatment options [11].

1.1.2 Epidemiology

Regarding the distribution over the world, a very high incidence rate of ESCC is predominantly observed in the so-called Asia belt including Northeastern Iran, Kazakhstan, Turkey, and Northern and Central China but also in Southern and Eastern Africa (Figure 2) [1, 12-14]. Remarkably, the Chinese population represents about half of all ESCC cases on earth [15]. Although, the highest incidence rates globally in both, men and women were found in Malawi [1]. EAC in contrast, is more often diagnosed in the US population and closely linked to obesity [15].

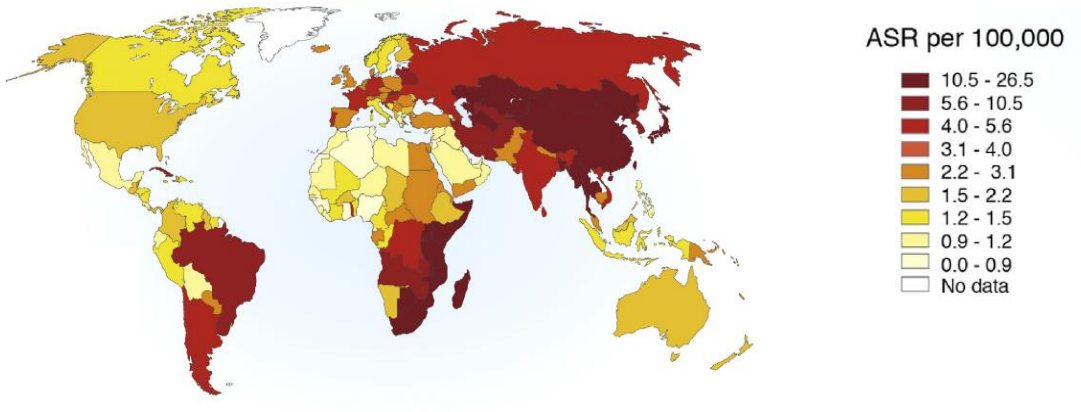


Figure 2: Global distribution in incidence rates of ESCC in men.

ASR, Age standardized rate; graphic was adopted from Abnet et al.2018 [15].

1.1.3 Development

Whereas EAC occurs in the lower esophagus and in the cells of mucus-secreting glands mainly caused by Barrett's esophagus, and gastric reflux, ESCC is more linked to the mid/upper region and the thin cells covering the surface of the esophagus [16, 17]. Common risk factors are comprised by high exposure to nitrosamine (tobacco usage), alcohol consumption, little fruit and vegetable uptake, deficiency of selenium, zinc, or vitamin E as well as poor oral hygiene [6, 7, 18]. Acetaldehyde has been described as a key carcinogen in ESCC [19]. Whereas human papillomavirus (HPV) presence is linked to head and neck cancer (HNSCC) formation, ESCC lacks HPV DNA, mRNA or p16 (INK4a) up-regulation. Therefore, HPV vaccination would be ineffective for prevention of ESCC [15]. There is evidence that some inherited disorders lead to ESCC formation including tylosis (thickening of the skin) [20] and Fanconi anemia (bone marrow failure syndrome) [21].

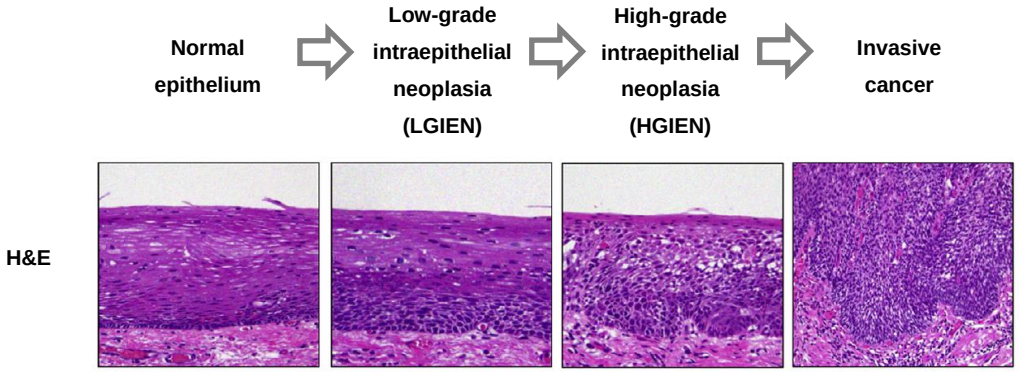


Figure 3: Scheme of different stages in ESCC development.

H&E: hematoxylin and eosin stain; graphic adopted from Bosman et al. 2010 [22].

ESCC is classified into four categories: well-differentiated, moderately differentiated, poorly differentiated, and undifferentiated SCC [22]. The development of ESCC is a multistep process starting with normal squamous epithelium progressing to low-grade intraepithelial neoplasia (LGIEN). From this alteration, a high-grade intraepithelial neoplasia (HGIEN) is formed and finally an invasive carcinoma arises (Figure 3) [22]. Even though intraepithelial neoplasia (IEN) and ESCC display differences in histology, both share genomic abnormalities constituting potential biomarkers for early detection of ESCC [23].

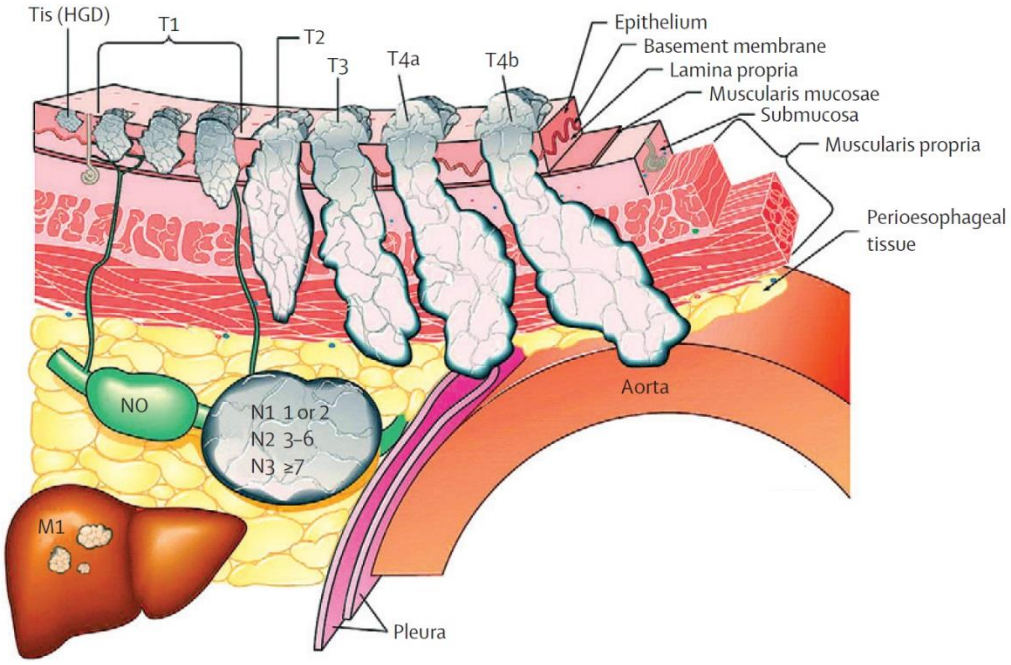


Figure 4: Tumor, node and metastasis staging system for esophageal carcinoma.

T: tumor stage; N: lymph node involvement; M: presence of metastases; Tis: intraepithelial neoplasia; HGD: high-grade dysplasia; graphic adopted from Pennathur et al. 2013 [6].

The tumor, node, metastasis (TNM) staging system is applied [11] for the characterization of the cancer progression (Figure 4) in order to estimate the patient's prognosis and to determine the most suitable line of treatment. Whereas "T" (T1-T4) describes the progression of the cell mass with an additional discrimination between "T4a" (resectable) and "T4b" (unresectable), "N" indicates the involvement of the lymph nodes. The presence of distant metastasis which are mainly found in the liver, lung, bone, and the brain [24] is indicated with "M1". By introducing an additional factor "G", information about the differentiation status is incorporated (G1-G4, well differentiated- poorly differentiated) [6, 8]. Differentiation of the squamous cell is linked to NOTCH1, SOX2, TP63, and ZNF750, serving as lineage-specific transcription factors [3]. In addition, Involucrin is only expressed in well-differentiated ESCC and therefore considered as differentiation marker [8]. In order to distinguish between early and advanced-stage EC, the number and location of involved lymph nodes are assessed [11]. The tumor evolves in a multistep process amplifying diverse genetic mutations. To allow for the final invasion of surrounding tissues,

matrix metalloproteinases (MMPs) are involved, whereby the activation of MMP2 and MMP9 display critical processes [25-27].

1.1.4 Molecular characteristics of ESCC

1.1.1.1 Genetic alterations

Various groups have investigated the genomic landscape of ESCC aiming for the identification of driving mechanisms such as mutated or amplified genes. Therefore, high-throughput technologies such as next generation sequencing have been applied. Four independent studies using whole-genome as well as exome sequencing unraveled a definition of the ESCC genomic landscape [28-31].

Generally, the mutational profile of ESCC has been shown to be more related to squamous cell carcinomas from distinct indications rather than to EAC [28]. Whereas the overall mutation pattern from ESCC is similar to head and neck squamous cell carcinoma [32, 33], it is different to lung squamous cell carcinoma [34]. The identified mutated genes are relevant in cell cycle control and apoptosis (*CCND1, CDKN2A, CHEK1, CHEK2, NFE2L2, RB1, TP53*) and differentiation (*NOTCH1, NOTCH3, ZNF750*). Additionally, mutations in the Hippo and the Notch pathway have been reported [3, 28, 30]. Somatic mutations of *TP53* were detected in 83-93% of ESCC, proposing p53 as a key factor for ESCC development [28, 30]. Even though *TP53* mutations have been described in early ESCC and are persistent to invasive cancer, solely, mutated *TP53* is insufficient for invasive cancer development [35, 36]. In contrast, *KLF5* loss is suggested to serve as critical factor leading to squamous cell transformation and invasion [35].

Beside the mutational landscape, focal amplifications as well as deletions in ESCC have been discovered, providing the cancer cells with an advantage in proliferation. Copy number alterations (CNA) have been found in the following genes: *CCND1, CDK4, CDK6, EGFR, FGFR1, KRAS, MDM2*, *MYC, NKX2-1, PIK3CA, PRKCI, SOX2*, and *TP63* [29]. In contrast, homozygous deletions have been observed for *FHIT, LRP1B*, and *PTPRD* [3, 8, 29, 37-42]. The area around location 9p21.3 has been identified as the most frequently deleted region, encoding *CDKN2A* and *CDKN2B* [3]. Epidermal growth factor receptor (*EGFR*) has been found to be overexpressed in

more than a half of ESCCs and has been associated with poor prognosis [28, 43]. Contradictory results have been observed upon interference with EGFR-mediated signaling in order to impair cell growth, which is further highlighted in section 1.1.5.

1.1.4.1 *Epigenetic modifications*

Besides the changes within the genome, various epigenetic modifications have been identified. A dysregulated epigenome is known to have a significant impact on the development of cancer [44]. In contrast to genomic mutations, epigenetic mechanisms are characterized by heritable phenotypic changes which are not linked to any DNA sequence alteration [45]. Whereas the genetic alterations have been already extensively studied over years and exploited in various ways for therapeutic purposes, the epigenome is still an unknown and an exciting territory. Particularly, studying its contribution to cancer initiation, maintenance and progression is from great interest. Especially in ESCC, interfering with the epigenome might reveal a novel strategy in order to evolve second line therapies for patients who progressed on first line treatments. In contrast to mutated drivers, epi-drivers are rather aberrantly expressed than frequently mutated [46]. Dysregulation of epigenetic mechanisms including DNA methylation, histone modifications as well as loss of genome imprinting have already been reported in the context of ESCC development [47]. Modifications have been detected in DNA/histone modifying proteins (CREBBP, EP300, MLL2), remover of histone modifications (KDM6A, TET2) as well as proteins which are responsible for modulation of the chromatin structures (ARID1A, ARID2, PBRM1, SMARCC2) [3]. In addition, overexpression of two histone methyltransferases, *EZH2* and *G9a* has been observed in ESCC [48, 49]. Analyzing the histone modification pattern, abnormal levels of H3K27me3, and H4K79me2 have been detected in primary ESCC tissue compared to non-malignant tissue [48, 50-52].

Besides histone modifiers and remodelers, additional epigenetic regulators such as DNA methyltransferases (DNMTs) contribute to genetic regulation. Methylation of promoters at CpG rich sequences (CpG islands) leads to the silencing of the respective gene [53]. Compared to nonmalignant esophageal mucosa, focal areas of hyper-methylation and widespread areas of hypo-methylation have been identified in ESCC [3]. Although hypo-methylation in ESCC is less understood, promoter hyper-

methylation, leading to gene repression, has been observed for *CDKN2A*, *CDKN2B*, *DLC1*, *LRP1B*, and *RASSF1A*, all of which are known tumor suppressor genes. Potential biomarkers might be comprised by *CDKN2A, CDKN2B,* and *TFF1*, hyper-methylation which was associated with early carcinogenesis [42, 54-57]. In addition, microRNA34a, and microRNA375, associated with anti-proliferative functions, have been down-regulated in ESCC [58, 59].

1.1.5 Treatment

Generally, the treatment options for diverse subsets of cancers have extremely evolved in terms of alternative approaches to radio as well as chemotherapy [60]. Thereby, the successful implementation of targeted cancer therapies has been achieved by exploiting cancer specific genomic alterations [61]. For instance, the molecule imatinib is targeting the prominent fusion protein BCR-ABL leading to selective inhibition in CML [62]. In ESCC however, therapy options are sparse as they are largely limited to surgery combined with chemo and/or radiotherapy. Strikingly, there is no second-line treatment for patients who progressed on first-line chemotherapy [63].

Curative treatment for patients with less progressed tumors involves surgery according to the histology type of the tumor. The surgery can be conducted either by transhiatal- or transthoracic esophagectomy together with lymph node dissection depending on the location of the tumor [6]. Unfavorably, a high mortality rate of 1-23% and a worsened quality of life is associated with esophagectomy which might lead to more minimally invasive approaches in the future, also decreasing the risk of pulmonary complications [6].

To reduce the chance of an early spread of the disease, surgery is combined with neo-adjuvant chemotherapy, showing a significant improved 3-year survival [64]. Neoadjuvant chemotherapy is composed of either cisplatin/fluorouracil or carboplatin/paclitaxel [65]. Clinical trials for chemoradiotherapy of ESCC without distant metastasis resulted in a complete response in 62.2% of the patients [66].

More advanced stages (metastatic or disseminated tumors) are treated with the intent to prolong patients' lives. The palliative methods consist of chemotherapy as well as local treatments of the dysphagia including radiotherapy or endoscopic

treatments (stent implantation) [6]. Additionally, pain killers are prescribed, psychological stress and resulting mental illness are addressed and treated [67].

In summary, chemo or chemoradiotherapy in combination with surgery still remain the most beneficial methods to treat ESCC [6]. The most efficacious treatment is considered to be chemotherapy composed of fluorouracil and cisplatin in optional combination with a third drug such as epirubicin or taxane [68]. The main challenge remain arising resistances upon treatment with chemotherapy [11]. Therefore, the expansion of second line therapies such as targeted or immune therapy is of great importance.

In the field of targeted therapy, investigations in the usage of bevacizumab (angiogenesis inhibitor), panitumumab, cetuximab and erlotinib (epidermal-growth-factor receptor (EGFR) inhibitors) are still ongoing [6]. Limited success has been observed in clinical trials targeting EGFR by failing to significantly improve overall survival of patients with ESCC [69, 70]. Nimotuzumab, a humanized EGFR monoclonal antibody, showed promising results in combination with paclitaxel and cisplatin in advanced ESCC [71]. In contrast, gefitinib, an EGFR inhibitor, did not improve overall survival as a second line treatment in ESCC [70].

Due to the fact that many modified proteins, as a result of genetic mutations are not targetable because of their lack of an enzymatic function, this restriction may be circumvented by interfering with other members of the pathway [46]. For instance, *CCND1* amplification, commonly detected in ESCC, functions through CDK activation. Therefore, targeting CDK4 and CDK6 may be beneficial instead of cyclin D1 [72]. To interfere with mutated p53, restoring of its function by delivering wild-type *TP53* is aimed to be achieved in order to target the MDM2-p53 interaction [73, 74].

An alternative approach to interfere with ESCC cell growth is comprised by applying novel strategies like immune checkpoint inhibitors. Thereby the interference with the well described programmed cell death protein 1 signaling pathway (PD-1/PD-L1), an immunological signaling pathway, [75] might be beneficial. Antibodies against PD-1 like nivolumab and pembrolizumab showed good efficiency in other squamous subtypes including metastatic squamous cell NSCLC and head and neck squamous cell carcinoma (HNSCC) [76]. Patients with an active PD-1/PD-L1 pathway have significantly poorer outcomes compared to patients with an inactive pathway [77, 78].

Therefore, patients which do not respond to chemotherapy are considered to be tested in clinical trials for inhibition of the PD-1/PD-L1 pathway [8]. Overall, the identification of a second line treatment is urgently needed which may be provided by targeted therapy or the combination with immune checkpoint inhibitors in order to prolong and improve patients' lives.

1.2 CRISPR screens for target identification

Despite the genetic landscape of ESCC being extensively studied in order to identify novel targeted therapies, many vulnerabilities are not predictable analyzing the mutation profile. Such dependencies are mainly caused by synthetic lethal interaction, described in section 1.4. In order to identify novel drug targets, the CRISPR-Cas9 screening technology has been successfully implemented assessing vulnerabilities in defined cancer subtypes of the whole genome or defined gene families such as the epigenome [79-84].

1.2.1 Adaptive immunity in bacteria

CRISPR stands for clustered regularly interspaced palindromic repeats, which have been characterized in *E.coli* showing short direct repeats, interspaced with short sequences (spacer) [85]. Initially characterized in 1987, CRISPRs have been detected in several bacteria and archaea [86, 87]. Interestingly, the spacer regions within the CRISPRs have been identified to originate from plasmid or viral sequences [88-90]. In addition to the CRISPR array (repeating elements and spacer sequences), Cas (CRISPR-associated) genes are transcribed generating a protein which possessing nuclease as well as helicase activity [88, 90-92]. Together, the CRISPR-Cas system has been proposed to comprise an adaptive immune system in bacteria, which might use antisense RNAs to memorize previous infections [86]. The confirmation of an existing bacterial immunity has been achieved in 2007 by using lytic phages for the infection of the lactic acid bacterium *Streptococcus thermophilus* [93]. In order to eliminate viral infection, a mature CRISPR RNA (crRNAs) has been identified to guide the Cas proteins to foreign sequences in *E.coli* [94] but also *Staphylococcus epidermidis* [95]. Finally, the process of the generation of an adaptive immunity has been described in three steps: insertion, transcription and cleavage [93, 96, 97]. Firstly, a short sequence of the invading DNA is inserted into the CRISPR array providing a spacer sequence. Secondly, the precursor crRNA (pre-crRNA) is transcribed and further processed into individual crRNAs and thirdly, the crRNA guides the Cas proteins to complementary sequences in order to induce cleavage of the foreign nucleic acids [93, 96, 97].

Six types of CRISPR-Cas systems can be distinguished which are furthermore grouped into two distinct classes [98-100]. The best studied systems include type I, II, and III [98, 101]. Whereas type I and type III mechanisms constitute large Cas protein complexes [94, 102-106], for RNA-guided DNA recognition and cleavage a single protein is required in the type II system which is therefore used for genome editing and screening [107, 108]. Additionally, in CRISPR-Cas systems of type II a trans-activating crRNA (tracrRNA) has been identified, being essential for crRNA maturation in *Streptococcus pyogenes* [109]. Furthermore, the CRISPR-Cas9 protein, the dual-RNA-guided DNA endonuclease, uses the tracrRNA:crRNA duplex [109] to be directed to its target sequence [107].

1.2.2 CRISPR-Cas9 provides a novel molecular tool for pooled screens

The state-of-the-art technology in terms of novel target identification is represented by the CRISPR-Cas9 methodology. In contrast to previously used target identification/validation tools like RNAi interference (RNAi), zinc finger nucleases (ZFNs) [110, 111] or transcription activator-like effector nucleases (TALENs) [112-114], CRISPR-Cas9 provides a robust and reproducible method. Compared to ZFNs and TALENs, which require protein modifications for each target site, CRISPR-Cas9 can be redirected to different targets by simply exchanging the single-guide RNA (sgRNA) sequence [115]. In addition to the precise and efficient editing feature of CRISPR-Cas9, the technology is cost-efficient, easy to use and applicable to many different organisms. Particularly, it is well suited for screening efforts such as high-throughput or multiplexed gene editing. Especially in cancer research, the usage of the CRISPR-Cas9 method allows studying the impact of specific mutations within a gene on cell proliferation and progression.

Even though RNAi has been proven to be a powerful genetic approach, the main disadvantages include the lack of robustness and reproducibility of RNAi screens. The challenges were comprised by circumventing off-target effects and the prediction of high-potent shRNAs [116]. The technology clearly revealed false positive results but also masked strong dependencies by inefficient targeting [93, 117]. In contrast, CRISPR-Cas9 provides a more efficient system using a programmable RNA sequence to guide the endonuclease to its determined region. For an eased application of the CRISPR-Cas9 technology the dual tracrRNA:crRNA has been

modified in order to retain a single-guide RNA (sgRNA), still maintaining the critical features including the 20-nucleotide sequence at the 5'end which is necessary for DNA target site recognition and the double-stranded structure at the 3'end mediating Cas9 protein binding [107].

1.2.3 Cleavage and repair of introduced double-strand breaks

The Cas9 (alternatively called: COG3513, Csx12, Cas5, or Csn1) derived from *Streptococcus pyogenes* introduces a double-strand break at the target DNA sequence which is subsequently repaired either via the non-homologous end joining pathway (NHEJ) or by homologous directed repair (HDR) [118, 119] (Figure 5). Both strategies can be used in different approaches. For the generation of gene knock-outs (KOs), the NHEJ pathway is harnessed whereas the introduction of a defined mutation is mediated by HDR by simultaneously providing the cells with a repair template. The Cas9 together with the sgRNA is directed to the target site, which must contain a protospacer-adjacent motif (PAM) consisting of a 5'-NGG sequence [107]. This sequence is specific to *Streptococcus pyogenes* whereas the usage of orthologs of Cas9 demand different PAM sequences. The PAM motif itself is critical for DNA binding, particularly the absence of a PAM site prohibits Cas9 recognition of the target site even if the sgRNA fully matches [120]. However, the need of a PAM motif constitutes no limitation of the technology due to the fact that the human genome contains respective sequences every 8-12 base pairs (bp) on average [118, 121]. Once bound to its target, the Cas9, possessing two exonuclease domains RuvC and HNH, introduces the double strand break into the DNA. The cut occurs three positions away from the PAM site. In absence of a template, the break is repaired in an error prone manner using NHEJ generating insertions or deletions (=INDELs) of different size. Furthermore, these modifications may lead to a frame-shift of the encoding sequence generating premature stop codons, leading to a KO of the respective gene (Figure 5) [122].

In terms of knock-ins the double-strand break is repaired in a precise way using HDR (Figure 5) [122]. Thereby an exact mutation can be introduced into the genome in order to study its contribution to cell survival or cancer initiation, for instance. For medicine approaches, this strategy may be used in the future to reverse a single base mutation leading to disorders such as cystic fibrosis [123].

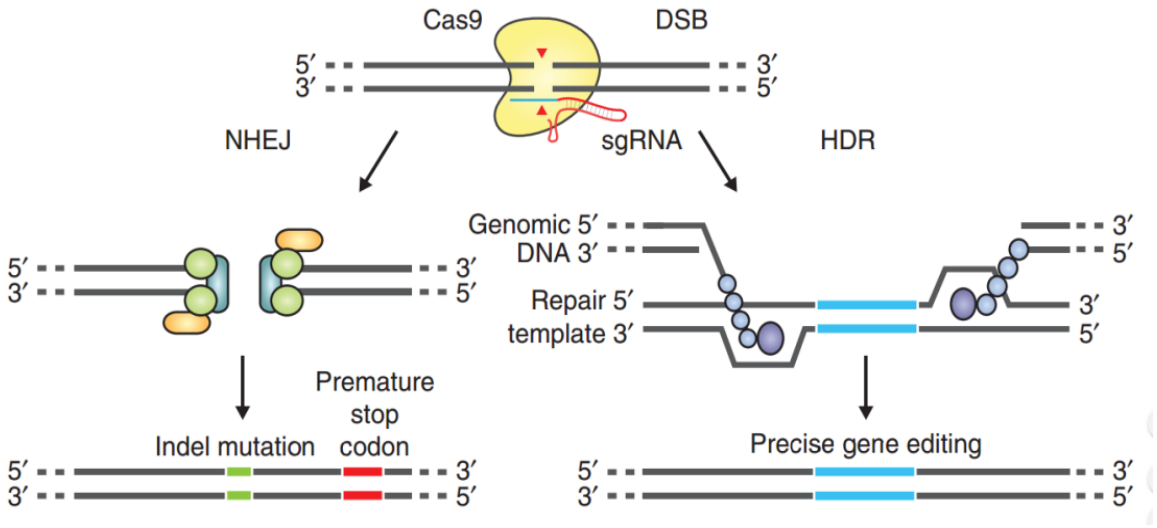


Figure 5: CRISPR-Cas9 genomic editing using NHEJ or HDR.

NHEJ: non-homologous end joining; HDR: homology directed repair; DSB: double-strand break; graphic adopted from Ran et al. [122].

1.2.4 Screening technology for identification of essential genes

The CRISPR-Cas9 method represents a break-through technology which allows precise genome modification in order to study the development and progression of diverse diseases including cancer. Furthermore, targeting genome-wide factors by a lentiviral sgRNA library allows pooled loss-of-function genetic studies by both positive as well as negative selection [82, 83]. Thereby the identification of cell essential genes in certain cancer subtypes can be elucidated [84].

Recently, two immense data-sets have been published by the groups of Sellers and Tsherniak [80, 81], targeting the whole genome in cancer cell lines. Whereas one screen is based on applying shRNAs, the other one uses sgRNAs, providing two complementary strategies to identify gene essentialities in certain cancer subtypes. Even though shRNAs are more prone to generate off-target effects, a large-scale RNAi screen targeting ~8000 genes in 398 cancer cell lines has been conducted revealing various undiscovered cancer vulnerabilities [81]. The off-target effects of shRNAs have been circumvented by designing 20 shRNAs per gene and the usage

of two different scores: RSA and ATARIS. Whereas RSA included all shRNAs, ATARIS [124] selected for shRNAs with consistent activity across the entire data set to further eliminate inert or potential off-target shRNAs [81].

Another screen has been published by the group of Tsherniak [80] including 342 cancer cells using the so called Avana sgRNA library [125]. A massive improvement of the CRISPR-Cas9 based screening technology has been enabled, optimizing the data analysis, by taking the copy number alterations into account. Amplified genes are known to be more often cut and thereby multiple double-strand breaks are introduced. Those modifications further lead to an induced G2 arrest, resulting in false positive interpretation of dependencies [126]. To overcome this barrier, the Ceres score has been introduced. The score was generated based on genome-scale CRISPR-Cas9 essentiality screening results obtained from 342 cancer cell lines [80].

1.2.5 CRISPR-Cas9 domain-based screen

Generally, sgRNAs for CRISPR-Cas9 screens are designed against the 5' exon of a respective gene [82-84, 127, 128]. This often results in in-frame editing events which still allow full gene expression. The resulting protein maintains its function and therefore obscures even strong genetic dependencies [79]. In contrast, targeting domain encoding regions within a gene ensures a higher proportion of null mutations and generates stronger effects in the following negative selection [79]. Additionally, first insights into the domain importance within a certain protein is revealed by using the domain-directed sgRNA design strategy [79]. Once a double-strand break is introduced into the targeted sequence, NHEJ repair is mediated which leads to the introduction of INDELs from different lengths. Regarding the triple amino acid code and the statistics of NHEJ repair, one out of three repairs will result in an in-frame sequence and ultimately in a fully functional protein [79]. Regarding the bi-allelic situation in a cell, a complete KO of the respective gene will only occur in four out of nine cells (Figure 6) [79]. Noteworthy, genetic dependencies are masked in five out of nine cells by leaving at least one functional allele, capable of generating a functional protein. In contrast, by introducing INDELs in functionally important domains, even in-frame repairs may result in null mutations. This might be due to inappropriate folding of the protein due to an altered amino acid sequence within the domain, resulting in a complete disruption of the function of the targeted sequence (Figure 6) [79].

In summary, designing sgRNAs against domains not only reveals the importance of the impaired region but also leads to more pronounced phenotypic effects and therefore reduces false negative results.

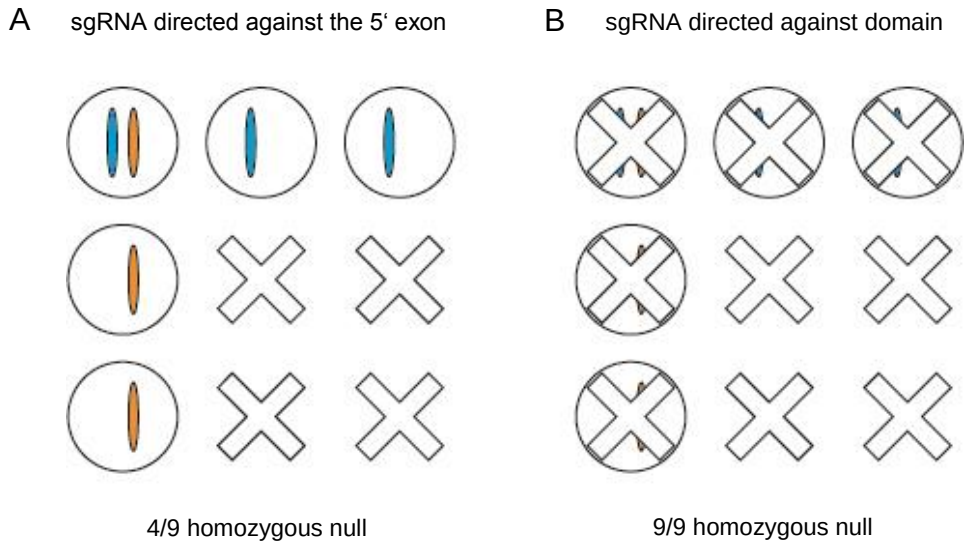


Figure 6: Targeting essential domains within a gene to generate homozygous null mutations.

A) One gene is indicated on two alleles (blue and orange). Successful out-of-frame KO is shown by loss of the gene whereas complete KO on two alleles is indicated by crosses. B) Targeting domains leads to the same probability of in-frame mutations in one of the targeted alleles but the repair results in miss-folding of the domain which leads to a non-functional protein. This results in a complete KO of the respective gene, generating a homozygous null mutation. The graphic was modified based on the publication from Shi et al. [79].

1.3 Chromatin remodeling by SWI/SNF complex

1.3.1 DNA packaging

In eukaryotes, the approximately two meters long strand of DNA is packaged into higher order fibers building up the chromatin. Hence, DNA is wrapped around histones allowing the packaging into the nucleus thereby forming the nucleosome [129]. The canonical nucleosomes are composed of two copies of the four canonical histones H3, H4, H2A, and H2B assembling an octameric disc [130, 131]. Each histone is able to wrap approximately 146 +/- 2bps of DNA [132]. The protein-DNA interaction is mediated by the differently charged molecules, which connect every ~10.4bp of the DNA [133]. Thereby, the positively charged histones interact tightly with the negatively charged phosphate backbone of the DNA. In addition to the canonical histones, all eukaryotes are able to incorporate different histone variants. The H2A.Z histone variant is linked to promoter gene activations as it flanks transcription start sites, whereas H1 or H5 contribute to packaging of the chromosome [134]. In order to loosen tightly packed sequences and initiate the expression of diverse genes, the chromatin structure has to be regulated and altered. Histone modifying enzymes as well as chromatin remodelers, capable of condensing or loosing chromatin, provide an indispensable mechanism for eukaryotic gene regulation [133].

1.3.2 Discovery of the SWI/SNF complex

The Switch/Sucrose Non-Fermenting (SWI/SNF) chromatin remodeling complex has originally been discovered in 1984 in yeast [135]. It has been described as a transcriptional regulator which activates genes necessary for mating type switching (hence the name Switch or SWI) and the change in the digestion of different energy sources on sucrose media (sucrose non-fermenting or SNF) [135-137]. Studies in yeast have revealed an adenosine triphosphate (ATP)-dependent chromatin remodeling machinery, which regulates transcription by mobilizing nucleosomes. This allows the complex to bind to the target DNA site and contribute to transcriptional activation [138-141].

The SWI/SNF complex has been independently discovered in *Drosophila melanogaster* to oppose the functions to the previously characterized Polycomb group, which is linked to transcriptional repression. Mostly, the Polycomb group mediates silencing of homeotic genes which is reversed by the SWI/SNF complex members [142]. Furthermore, the complex is essential for embryogenesis and for all development stages in *Drosophila* [143, 144]. The ATPase subunit Brahma (Brm) in flies was found to be identical to the yeast SWI2 or SNF2 [145].

Mammalian counterparts have been defined for almost all SWI/SNF complex members, initially characterized in yeast and drosophila [146]. The human orthologs of the ATP-dependent subunit SWI2/SNF2 are comprised by SMARCA2 or SMARCA4 (SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily A member 2/4). Both provide the catalytical center of the complex which mediates the mobilization of nucleosomes in an ATP-dependent manner [147, 148]. In mammals, SMARCA4 is also termed BRG1 (Brahma-related gene 1) and SMARCA2 is referred to as BRM (Brahma). The complex therefore is also named the BAF complex standing for BRG1-associated factors. From hereafter, it is referred to as the SWI/SNF complex. The contribution of the complex to transcriptional regulation is well characterized including DNA replication [137, 149, 150], DNA repair [151, 152] and decantination [153, 154]. Whereas in yeast, the SWI/SNF complex is mainly associated with transcription activation, in mammals, the complex functions either as activator or repressor [155]. Additionally, several subunits not to be found in yeast, provide the complex with chromatin interaction ability, mediated by the bromodomain (BD) which allows recognition of acetylated lysine [156-158].

1.3.3 Chromatin remodeling

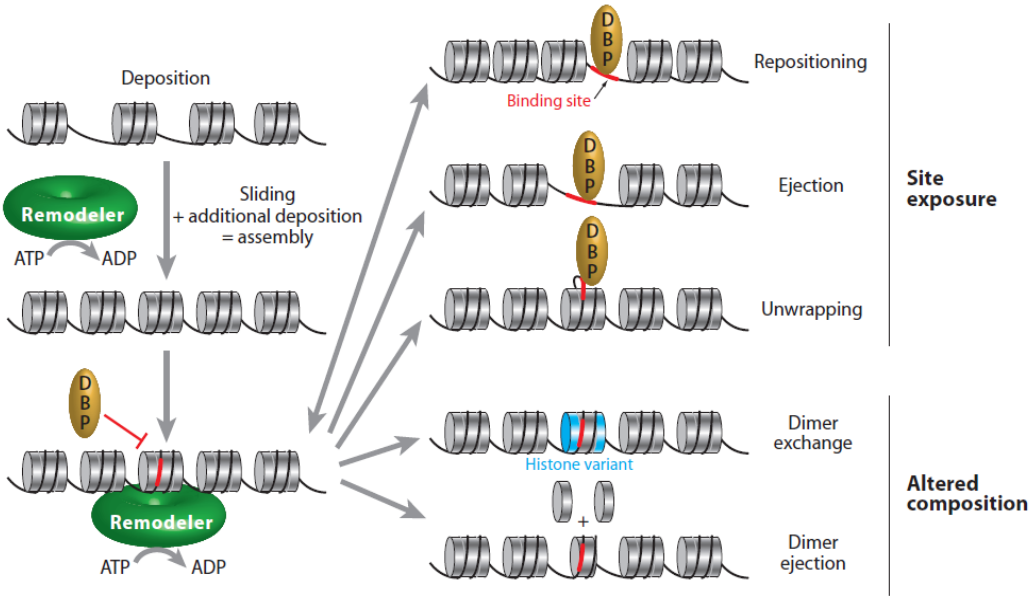


Figure 7: Chromatin remodeler function leading to site exposure or altered composition.

Accessibility of cognate DNA sequences is conducted by hydrolysis of ATP-mediated by the chromatin remodeling complex which leads to site exposure via repositioning, ejection and dimer exchange. Alternatively, the remodeling complex contributes to altered composition of the histone, by exchanging or ejecting a dimer; graphic was adopted from Clapier & Cairns [133].

Chromatin remodeling is an epigenetic mechanism which is responsible for the weakening of DNA-histone interactions and thereby revealing the cognate sites of DNA for its binding factors, necessary for activation or repression of transcription (Figure 7) [159]. Additionally, chromatin remodeling complexes are responsible for chromosome segregation, DNA replication as well as for DNA repair. For the regulation processes, ATP hydrolysis serves as a source of energy [160] allowing specific transcription factors to access and bind to DNA to finally initiate transcription. This process leads to either site exposure by repositioning, ejection or unwrapping of histones or altered composition by incorporating histone variants [133]. The dynamic alteration provides a balance between DNA access and DNA packaging.

The SWI/SNF complex is one of the five non-redundant chromatin remodelers which are present in eukaryotic cells [161]. The additional four include ISWI, NURD/Mi-2/CHD, INO80, and SWR1 [161], all of which use ATP hydrolysis to shuffle histone-DNA interfaces. Among their common properties are the conserved ATPase domain and the ability to recognize covalent histone modifications. Even though the different chromatin remodeling complexes share common features, they possess diverging additional domains, allowing specificity of each individual complex [138, 161]. The two most extensively studied chromatin remodelers are the SWI/SNF and the ISWI complex [161], both belonging to the superfamily II (SF2) of DEAD/H-box helicases and translocases [162].

1.3.4 SWI/SNF complex composition

The composition of the SWI/SNF multi-subunit mammalian complex has first been described in 1996 by the application of co-immunoprecipitation of the ATPase subunit SMARCA4 [163]. While the exact composition varies from cell type to cell type, four core subunits are common to all SWI/SNF complexes. Noteworthy, the catalytic center of the complex is exclusively composed of one of the two redundant ATPases SMARCA2/BRM or SMARCA4/BRG1 [164]. Together with three additional subunits including SMARCB1/BAF47/SNF5, SMARCC1/BAF155, and SMARCC2/BAF170 (Figure 8), the core complex is formed [165, 166]. Additional subunits contain specific domains including bromo, chromo, DNA-binding domains, ARID as well as zinc-finger, all are required for interaction with DNA or histones and essential for the remodeling function of the complex [167-169].

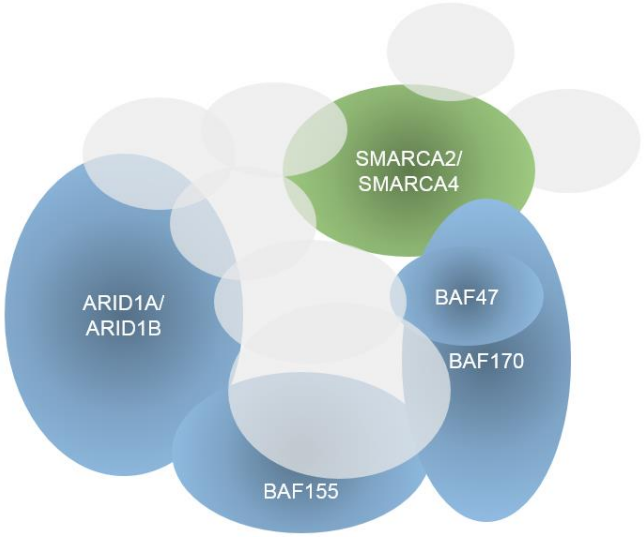


Figure 8: SWI/SNF core complex composition.

The catalytic center is comprised of SMARCA2 or SMARCA4, which are incorporated in a redundant way. The four additional core components include ARID1A/B, BAF47, BAF155, and BAF170; graphic modified from the publication of Michel et al. 2018 [170].

Furthermore, the assembly of the complex occurs in three different ways by alternatively incorporating the products of 29 genes. Recently, BCL7a/b/c, BCL11a/b, BRD9 and SS18 (SYT) have been confirmed as “new” complex members via proteomic screens [171]. Regarding the diversity in assembly of the complex members, hundreds of SWI/SNF complex versions might exist [156, 172]. Three sub-classes, which have been described so far, include canonical (cBAF/ SWI/SNF-A), polybromo-associated (PBAF/ SWI/SNF-B) and non-canonical SWI/SNF (ncBAF) complexes [170]. Networks comprising ARID1A/ARID1B (BAF250 A/B), BAF45D and SS18 are referred to the canonical SWI/SNF complex (Figure 9A), whereas the complex exclusively incorporating ARID2 (BAF200), PBRM1 (BAF180), BAF45A, BRD7 and SMARCA4 is named PBAF or SWI/SNF-B complex (Figure 9B) [156, 173, 174]. The non-canonical SWI/SNF complex is composed of GLTSCR1/GLTSCR1L, and BRD9 but lacks core subunits like SMARCB1 (BAF47) and SMARCE1 (BAF57) (Figure 9C) [157, 175-177]. Additionally, the non-canonical SWI/SNF complex selectively incorporates paralogs such as SMARCC1/BAF155 but not SMARCC2/BAF170 or SMARCD1/BAF60A instead of SMARCD2/BAF60B or SMARCD3/BAF60C [170].

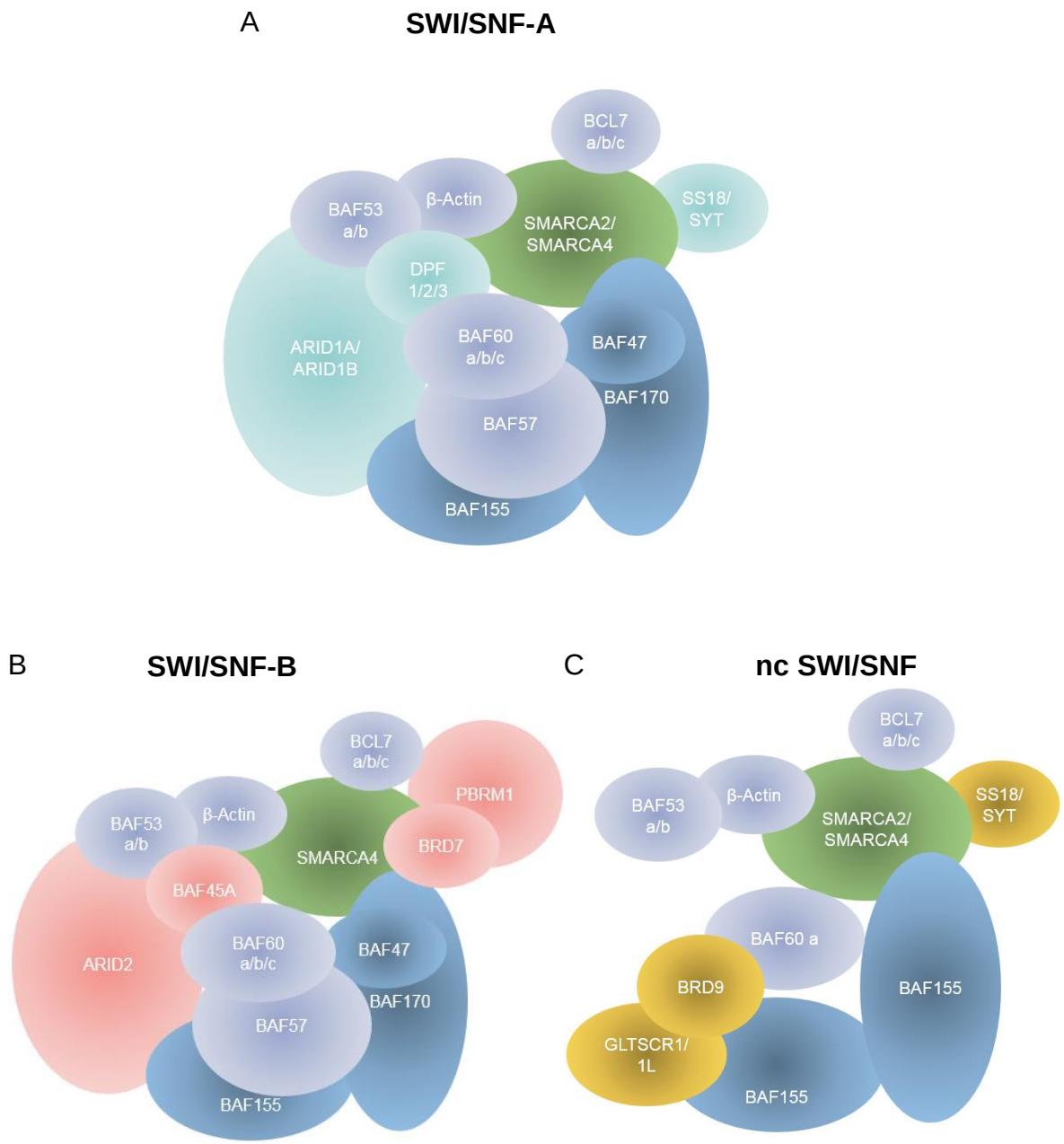


Figure 9: Diverse SWI/SNF complex composition.

A) SWI/SNF-A: cBAF, canonical BAF complex, B) SWI/SNF-B: PBAF, polybromo-associated factors, C) ncSWI/SNF: ncBAF, non-canonical BAF; graphic modified from publication of Michel et al. 2018 [170].

1.3.5 The role of the SWI/SNF complex and SMARCA4 in normal cells

The SWI/SNF complex has been reported to bind 5000 to 10.000 sites over the entire genome and thereby contributes to the expression of various genes required in diverse pathways [178]. Both, neuronal development [178] and hormone signaling [179] are dramatically influenced by the SWI/SNF complex as well as hepatic lipid metabolism [180]. Cell fate decisions are also mediated by the SWI/SNF complex including skeletal, cardiac muscle and hematopoietic differentiation [167, 181-184]. One of the first roles characterized was the ability of the SWI/SNF complex to function in the regulation of lineage-specific enhancers [185]. By directly binding to p300 histone acetyltransferase, the complex is implemented in the regulation of the H3K27ac mark on target genes [185]. Whereas the complex has been less present at super-enhancers, strong activity has been associated with typical distal enhancers enriching for genes involved in differentiation and development [185]. Furthermore, SWI/SNF antagonizes the repressive function of the Polycomb complexes. By targeting SWI/SNF to a precise promoter, rapidly removing from chromatin binding sites of both PRC1 as well as PRC2 complexes has been detected, which in turn has revealed increased chromatin accessibility for transcription factors [186].

SMARCA4 has been linked to the regulation of the transcription of diverse genes involved in disparate cellular processes, as for instance, SMARCA4 regulates basal expression of membrane glycoprotein CD44 [187]. It furthermore modulates the transcriptional activity of estrogen receptor, glucocorticoid receptor, c-Myc, BRCA1 and retinoblastoma tumor suppressor protein (RB) [188-193]. SMARCA4 is furthermore able to bind RB which leads to [194, 195] mitotic cell cycle arrest [196-200].

1.3.6 Contribution of the SWI/SNF complex to cancer

The SWI/SNF complex is the most frequently mutated tumor suppressor beside *TP53*, with a mutation rate of ~20% in all cancers assessed by exome and whole genome sequencing [171, 201]. Mutations within the approximately 15 subunits containing complex can either occur hetero or homozygous, somatic or in the germline resulting in point mutations or translocations [167]. The mutation frequency of individual complex members varies between indications and proteins, whereas some genes are more frequently mutated than others. Since members of the SWI/SNF complex have been identified to be mutated or lost at high frequency in multiple cancer types, the complex has gained considerable attention.

The first alteration identified in malignant rhabdoid tumor (MRT), an aggressive early childhood cancer, was BAF47/SMARCB1/hSNF5, presenting a loss of function in both alleles [202]. Those tumors can arise in the brain, kidney and other soft tissues. Mutations have been found in nearly all samples analyzed [202]. SNF5/SMARCB1 is one of the core subunits whereas mutations lead to gene-specific transcriptional alterations [203]. Furthermore, the mutated complex is not able to antagonize the repressive function of EZH2 (enzymatic subunit of the Polycomb complex, PRC2) mediating H3K27me3 which leads to a stem cell-associated gene expression program favoring tumor growth [204]. The de-differentiation as a hallmark of cancer is observed in multiple cellular contexts, bearing loss of SWI/SNF function [205, 206].

Dependent on the indications, different SWI/SNF subunits have been found to be mutated. Whereas mutations in *SMARCB1*/*SNF5* occur in more than 95% of MRT, *PBRM1/BAF180* mutations are detected in ~40% of renal clear cell carcinoma [207]. *ARID1A* is the most frequently mutated SWI/SNF complex gene which is often found to be mutated in hepatocellular carcinoma [208, 209], lung adenocarcinoma [210], gastric cancer [82, 211, 212], bladder cancers [213, 214] and cholangiocarcinoma [215].

With the exception of adenoid cystic carcinoma (ACC) [216], SMARCA2 mutations are rather linked to neurological disorders than to cancer [167]. However, loss of SMARCA2, has been shown in 15-20% of solid tumors including esophageal adenocarcinoma which is mainly caused by promoter polymorphisms [217]. Those polymorphisms are in turn correlated with epigenetic silencing of *SMARCA2* by

recruiting histone deacetylases (HDACs). Furthermore, increased risk of development of cancer of lung, head and neck and the upper digestive tract has been associated with *SMARCA2* loss [218-220]. Mice lacking *Smarca2* are ten-fold more prone to develop tumors after exposure to carcinogens [221]. In addition, loss of heterozygosity (LOH) at the *SMARCA2* locus 9p23-24 has been observed in diverse tumor types [222-225].

Compared to its paralog member, SMARCA4 has been identified to be mutated in a subset of cancer cell lines and is therefore considered as a tumor suppressor [226]. Mutations in *SMARCA4* have been identified in 100% of all cases in small cell cancer of the ovary hypercalcemic type (SCCOHT) [227] and SMARCA4-deficient thoracic sarcomas (SMARCA4-DTS) [228]. However, respective indications have shown concomitant loss of *SMARCA2* without any link to gene alterations [228-230]. With varying frequency *SMARCA4* mutations, mainly in the ATP-binding domain, are detected in Burkitt's lymphomas [231], lung adenocarcinomas [210, 232, 233] and esophageal adenocarcinomas [234]. Loss-of-function mutations or loss of expression of *SMARCA4* have predominantly been reported in 7-10% of NSCLC [210, 233, 235-237] but have also been detected in melanoma, liver and pancreatic cancer. Samples of NSCLC patients included 15.5% to be deficient in SMARCA*4* expression, analyzed by immunohistochemistry [238]. Importantly, mutations in *SMARCA4* occur in a mutual exclusive fashion without concomitant alterations such as therapeutically targetable proteins resulting from *EGFR* mutations, *ALK* fusions, or *FGFR1* gene amplifications [238].

Alterations in individual subunits of the SWI/SNF complex are largely loss-of-function mutations, suggesting tumor suppressive roles of the SWI/SNF complex. However, oncogenic "driving" mutations have been reported in a few indications. Those cancers include MRT (malignant rhabdoid tumors), AT/RT (atypical teratoid/rhabdoid tumor), SCCOHT, synovial sarcomas and SMARCA4-DTS [202, 229, 239-242]. In synovial sarcoma, the mutation of the *SS18* gene represents an oncogenic driver [240]. The alteration always results in a fused SS18-SSX, whereby exactly 78 amino acids of SSX are added to SS18 [243, 244]. The chimeric protein is still incorporated into the SWI/SNF complex but activates the transcription of oncogenic drivers. Whereas the incorporation of wild-type SS18 protein into the SWI/SNF complex leads to an arrest of cell proliferation and cell death, a therapeutic interference with the pre-

dominantly incorporated version SS18-SSX might be beneficial for tumor growth inhibition [240]. At the same time the incorporation of the SS18-SSX fusion protein leads to ejection of SMARCB1 from the complex, a phenomenon occurring only secondary, being non-essential for expression and proliferation of synovial sarcoma [245].

In addition, a non-functional SWI/SNF complex not necessarily demands for mutations in one of the subunits. In a few cancers, Ewing sarcoma [246], certain breast cancers [82] and prostate cancers [247], a non-functional SWI/SNF complex is detected without any mutations.

1.4 Synthetic lethality – as a novel therapeutic approach

1.4.1 Synthetic lethal interaction between members of diverse signaling pathways

Synthetic lethality (SL) constitutes a well-established concept which has been explored by genetic studies conducted in fruit fly and yeast [248-253]. Whereas concomitant defect of two (or more) partners is lethal, individual impairment of only one component is compatible with cell viability (Figure 10) [254]. In cancer cells, mutation or loss of one partner, shifts the dependency towards its synthetic lethal partner.

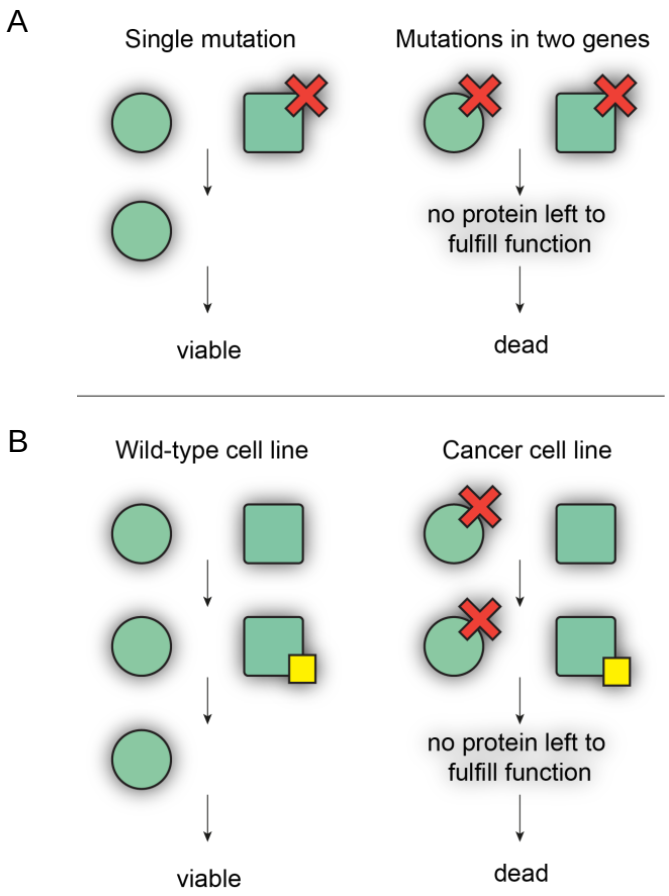


Figure 10: Principle of synthetic lethality and therapeutic approaches.

A) Mutation (red) of an individual gene does not affect cell viability, whereas mutations in two individual genes in the same cell lead to cell death, B) Pharmacological intervention, whereas inhibition of one protein (yellow) does not harm the cell in a non-mutated background, a cell which has already a mutation (red) in one partner following redundant factor inhibition is lethal.

Cancer cells are deficient in one component due to deletions, mutations or epigenetic silencing which sensitizes them towards inhibition of the redundant partner. This concept of SL arouses interest in pharmaceutic interventions. Thereby the application of a drug, specifically targeting one of the two individual proteins in cancer cells showing a defect in the SL protein, induces cell death. In contrast, the treatment spares healthy cells, still harboring a functional second partner able to compensate for the loss (Figure 10). Clinically, the concept of SL has shown promising results for the inhibition of Poly-ADP-Ribose Polymerase (PARP) in Breast Cancer 1/2 (BRCA1/BRCA2)-deficient tumors [255, 256]. Both factors are implicated in DNA repair. Upon DNA damage which is initiated by single-strand breaks, the cells retain viable when lacking only one component which is important for the repair. In contrast, the cells undergo apoptosis if they have lost both repair mechanisms: base excision repair (BER) and homologous recombination (HR) mediated by PARP and BRCA, respectively [257]. Whereas, PARP1 is involved in DNA damage recognition by binding to DNA at single-strand DNA breaks [258-262], BRCA1/2 proteins are responsible for the repair of double-strand DNA breaks (DSBs) [263]. The development of diverse cancers including breast and ovarian cancers has been associated with heterozygous deleterious mutations of the genes *BRCA1* and *BRCA2* occurring in the germline [264-266]. During tumorigenesis, the wild-type allele of *BRCA* is lost and therefore considered as tumor suppressor. Cells with a deficiency in BRCA1/2 or other pathway component circumvent the repair by using alternative mechanisms such as NHEJ [267-269]. This alternative repair frequently induces DNA deletions [267-269], participating at least partially to cancer initiation or progression. Additionally, BRCA1/2 contributes to chromatin remodeling and transcriptional regulation, possibly being relevant for pathogenesis [270]. The SL with the regard to mutations in *BRCA1* or *BRCA2* and inhibition of PARP, has been explored by two groups [271, 272] describing tumors with *BRCA* mutation to be 1000-times more sensitive towards inhibition of PARP than *BRCA*-wild type cells [271].

Harnessing the concept of SL therefore serves as a successful new strategy in inhibition of cancer growth by individualized treatment [271, 273, 274]. To date, five different PARP inhibitors are available including talazoparib, niraparib, rucaparib, olaparib and veliparib which vary in their ability to trap PARP1 on DNA, inhibiting protein poly ADP-ribosylation (PARylation), which in turn correlates with cytotoxic potency [275-278]. In high-grade serious ovarian cancer, treatment with olaparib in

combination with chemotherapy leads to significantly improved progression-free survival in *BRCA*-mutated patients [279]. Using olaparib, in a phase 1 trial including patients with *BRCA1* or *BRCA2* germline mutations, 63% of patients showed a clinical benefit [273], which has been furthermore confirmed in breast, ovarian, pancreatic and prostate cancers [255, 280]. Olaparib has been approved for patients with advanced ovarian cancer who were already on previous therapies [281]. Promising phase 3 results have been obtained from treatment of ovarian cancer with niraparib [282] but also rucaparib extended progression free survival [283]. In addition, talazoparib, has shown to reduce tumor volume in early-stage breast cancer patients with germline mutations in either *BRCA1* or *BRCA2* when treated in a neoadjuvant setting [284]. Further SL interactions highlighting novel drug targets have been validated preclinically [256, 285]. Since the discovery of the SL interactions in *BRCA1*/*BRCA2* mutated and PARP inhibited tumors, additional SL have been investigated, which are implicated in chromatin remodeling. Thereby the SWI/SNF complex members provide an interesting platform for exploring SL. In MRT, described in section 1.3.6, cells with a mutation in the SWI/SNF complex member *SNF5* induce dependency on EZH2 but also SMARCA4/BRG1 [202, 239].

Additionally, SL interactions are not only restricted to individual proteins but are also found for entire complexes such as the canonical and non-canonical SWI/SNF complex [170]. MRT and synovial sarcoma, both display a defect in the canonical complex (1.3.6), indicate dependency on BRD9, a member of the non-canonical SWI/SNF complex [170]. Whereas targeted therapies have been successfully implemented in the clinic in diverse contexts, interfering with amplified or mutated genes [286], treatment of loss-of-function mutations or deletions had limited success so far. Thus, exploiting SL dependency is a promising novel approach for targeted therapy, especially treating cancers with mutations in tumor suppressor genes.

1.4.2 Paralog dependencies

An obvious approach to study SL interactions is comprised by the investigation of dependencies of paralog family members, which are composed of at least two proteins mediating redundant cellular processes [287-289]. Paralog genes have evolved through gene duplication within the same species via acquiring mutations through evolution [290]. The SWI/SNF complex is known for the incorporation of only one member of a paralog family such as ARID1A/B or SMARCA2/4 [238, 291-293]. Therefore, members of the SWI/SNF complex arose great interest in order to study SL interactions.

Several SL interactions, constituted by other paralog proteins, have been characterized including the dependencies between ENO1 and ENO2 as well as STAG1 and STAG2. The members of each paralog family fulfill redundant but cell essential functions. In tumor cells, strong dependencies on one component upon deficiency of the paralog partner is observed [294, 295].

In glioblastoma (GBM), glycolytic gene enolase 1 (*ENO1*) deletion is tolerated because of the maintained presence of *ENO2.* Interference with ENO2 via short-hairpin-RNA-mediated (shRNA) silencing has led to growth inhibition in *ENO1* collateral deleted cells [294]. Collateral vulnerabilities are described as unintended events in which co-deletion of a tumor suppressor gene and a member from a redundant family, serving as essential housekeeping genes [294, 296], occurs. Similarly, in bladder cancer as well as in Ewing sarcoma, mutation in *STAG2*, encoding for a subunit of the cohesion complex, induces dependency on its paralog STAG1 [295]. Re-expression of STAG2 in previously *STAG2*-mutated cell lines alleviates dependency on STAG1, both implemented in supporting sister chromatin cohesion [295].

Additional SL dependencies have been confirmed for the two, core complex SWI/SNF members, namely ARID1A and its paralog ARID1B in ovarian cancer cells [291]. Also, a novel therapeutic concept has been suggested for *SMARCA4*-mutated cancers which gain dependency on the redundant ATPase SMARCA2. Loss of the helicase subunit of SMARCA4 has been detected in various cancer subtypes [210, 233, 297] whereas in NSCLC the maintenance of the SWI/SNF complex function has been accomplished by the paralogous helicase SMARCA2 [238, 292, 293]. Loss of

function mutations in *SMARCA4* renders NSCLC cell lines sensitive towards SMARCA2 inhibition, identified via an shRNA screen conducted in 58 cell lines [292]. Interference with SMARCA2 in *SMARCA4*-deficient cell lines has led to cell cycle arrest, induction of senescence as well as increased levels of global H3K9me3 have been observed [292]. In addition, conditional RNAi studies *in vivo*, have revealed impaired tumor xenograft growth of *SMARCA4*-deficient cells, upon SMARCA2 depletion [238]. Mutations in *SMARCA4* occur in ~10-15% of adenocarcinomas of the lung and targeting SMARCA2 is suggested as a novel attractive therapeutic approach in those patients [292].

1.5 Approaches to therapeutically interfere with SMARCA2 or SMARCA4

Chromatin remodelers represent exciting new targets for therapeutic intervention. In pre-clinical as well as in clinical settings, inhibitors of histone modifying enzymes as well as BD containing “readers” have shown efficacy in interfering with oncogenic transcriptional programs [79, 298, 299]. Cancer selective molecular vulnerabilities which can further be linked to a certain context represent an attractive class of drug targets [300]. The findings of SL interactions within the two redundant ATPases of the SWI/SNF complex constitutes great potential of interfering with SMARCA2 in homozygous *SMARCA4*-deleted cancers [301].

1.5.1 Domains within SMARCA2/4

Designing a selective compound however is challenging due to the fact that both ATPases share ~ 75% of their amino acid sequence [146] and exhibit six conserved domains. The functional specificity has been linked to the sequence variation close to the N-terminus, whereas SMARCA2 and SMARCA4 vary in their affinity to interact with specific transcription factors [302]. The N-terminus of SMARCA2/4 comprises two domains including the QLQ domain which mediates protein-protein interactions, followed by a proline-rich domain (Figure 11). The separation of the DNA double strands is mediated by the helicase and DEXDc domain [303]. The LxCxE allows binding to the retinoblastoma (RB) protein and the C-terminal BD mediates binding to acetylated histones but also ensures stability of the SWI/SNF complex when bound to the DNA (Figure 11) [164].

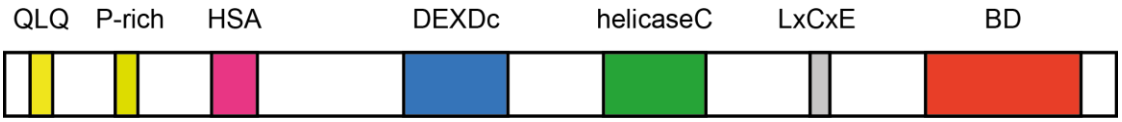


Figure 11: Scheme of SMARCA2/4 domains.

Graphic was adopted from Arnaud et al. [207]; QLQ: conserved Gln, Leu, Gln, P-rich: Proline rich, HAS: helicase-SANT-associated, DEXDc (DEAD-like helicases superfamily) and helicaseC: ATP-binding domains, LxCxE: pRb interaction domain, BD: bromodomain.

1.5.2 The ATPase domain is the functional relevant domain of SMARCA2/4

The chromatin remodeling function is mediated by the DNA-stimulated ATPase domain, which is to an extend of 92% identical between SMARCA2 and SMARCA4 [146, 156, 301, 304, 305]. Hence, selectively interfering with the ATP-binding domain is very challenging. The successfully developed dual inhibitor of the ATP-binding domain of both SMARCA2/4 has been used as a tool compound [301]. Thereby, effects of SL interactions identified using RNAi have been analyzed whether to be transformable into pharmacological inhibition of SMARCA2 in *SMARCA4*-deficient cell lines [301]. Furthermore, the functional relevant domain for sustained proliferation of *SMARCA4*-deficient NSCLC cell lines has been assessed [300]. By conducting rescue experiments, the ATP-binding domain has been shown to be the functional indispensable region within SMARCA2/4 [300]. In contrast, inactivating mutations in the BD have not revealed evidence for essentiality in SMARCA4 function in LOF (loss of function) experiments in *SMARCA4*-deficient lung cancer cells [300]. Cells have stopped proliferating upon SMARCA2 siRNA treatment. The anti-proliferative effect has been restored by SMARCA4 expression in A549 cells. Furthermore, the SL of SMARCA2 knock-down (KD) has been linked to the ATPase- and not BD- function by ectopic expression of wild-type, ATP-binding pocket deficient (K755A) [238] or BD mutant (N1482W) [306] variants of SMARCA2 [300].

1.5.3 BD inhibition

The druggability of BDs has been observed by the development of JQ1, a BET inhibitor interfering with the BD and an extra-terminal domain of BRD4, which in turn is associated to antitumor activity [298]. The aim was to extend the therapeutic approach to other BD-containing proteins. However, attempts of validating the selective SMARCA2/4 BD inhibitor (PFI-3) as a tool compound has not phenocopied growth-inhibitory effects of SMARCA2 KD in the *SMARCA4*-deficient cell lines A549, H1299 and H157 [300]. Furthermore, the BD of SMARCA2 has been explored to be dispensable for chromatin binding and for contribution of oncogenic activity in lung cancer [300]. In addition, rescue studies, ectopically expressing SMARCA2 or SMARCA4 variants, have highlighted the ATP-binding domain as functionally important in RNAi-mediated loss-of-function (LOF) experiments [300]. Furthermore,

SMARCA4 BD has been confirmed as dispensable in SMARCA4-dependent AML cell lines [300, 307].

1.5.4 PROTAC approach

In order to circumvent the limitation of developing a selective SMARCA2 or SMARCA4 ATPase-interfering inhibitor and the fact that targeting the BD does not phenocopy LOF effects, an alternative approach is comprised by using the proteolytic targeting chimera (PROTAC) system. The selective design of a bi-specific molecule against the BD of either SMARCA2 or SMARCA4 would lead to degradation of the whole protein. Comprising novel approaches for intracellular targeting, PROTACs act as heterobifunctional molecules. Therefore, two proteins are brought into close proximity upon the PROTAC application. One of the ligands binds to the target protein, the other one to the E3 ubiquitin ligase, connecting both proteins via the linker region. The ternary complex composed of the target, the PROTAC and the E3 ligase recruits the E2 ubiquitin-conjugating enzyme which subsequently transfers ubiquitin onto the target surface [308]. Following the ubiquitination, the proteasome recognizes the targeted protein and mediates its degradation (Figure 12). After the degradation process, PROTAC molecules are recycled, providing a great advantage of the method by sub-stoichiometric application of the molecule [308].

The most extensively studied E3 ligases for PROTAC applications include mouse double minute 2 homolog (MDM2), inhibitor of apoptosis proteins (IAP), von Hippel-Lindau (VHL) and cereblon (CRBN), whereas VHL and CRBN have been widely exploited for PROTAC development [308].

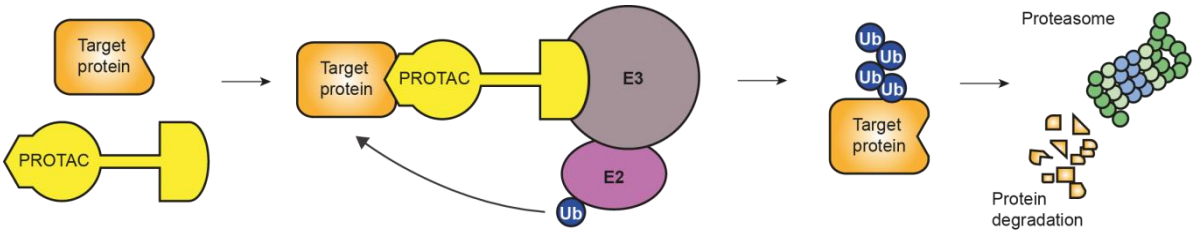


Figure 12: Scheme of proteasomal degradation using a bi-specific molecule.

Graphic was modified based on publication of Cermakova et al. [308].PROTAC binds to the targeted protein as well as the E3 ubiquitin ligase, E2 mediates the ubiquitination of the target protein leading to recognition and degradation by the proteasome.

Optimized molecules are able to degrade their targets within several hours and sustain reduced protein levels almost for 48 hours [308]. However, auto-inhibition ("hook effect") is mediated by exceeding the DC_{50} [308]. Thereby the formation of the ternary complex (target:PROTAC:E3) is impaired due to high concentrations of the binary complexes (target:PROTAC and PROTAC:E3) [308].

First steps towards PROTAC-mediated target inhibition have already been made by targeting oncoproteins including androgen receptor, estrogen receptor, ERRα, and BRD4 [309].

1.5.5 BRD9-directed PROTAC

A successful application of a cereblon-based PROTAC has been implemented for targeting BRD9 [310], a selective vulnerability in AML (acute myeloid leukemia) [311]. Interfering with the BD of BRD9 has been achieved via a small-molecule inhibitor BI-7273 [312], successfully impairing cell growth in AML cell lines [311]. Based on this compound, efforts have been made in the discovery of PROTACs targeting BRD9 (dBRD9), achieving a breakthrough in the field of novel strategies to interfere with proliferation of cancer cells. The group of Bradner [310] has been able to show degradation upon four to 24 hours of dBRD9 treatment and subsequent impairment of cell proliferation [310].

1.5.6 Domain-swap strategy

Additionally, the functional exchange of domains within certain protein families has been successfully demonstrated in AML for BRD9. By generation of a chimeric allele bearing the BD of highly homologous BD-containing members BRD7 or BRD1, initial effects of a selective inhibitor have been reversed, proofing the selectivity of the compound [311]. Within the human proteome, 61 BDs are known which are associated to 46 different proteins [313]. Their main function is to specifically recognize acetylated lysine residues and therefore mediate transcriptional activation. However, diverse BD-containing proteins have distinct binding preferences for acetyl-lysine-containing peptides [311]. Interestingly, selected BDs within different proteins can be swapped by fully retaining the function of the respective proteins [311]. Even though the BD was radically altered in AML cell lines, the protein remained functional and its expression conferred resistance towards BRD9 inhibition. Therefore, the domain swap strategy has allowed testing of the selectivity of certain compounds and has confirmed on-target activity [311].

1.6 Aim of the PhD project

The ultimate aim of the PhD project is to identify a novel vulnerability concept in ESCC and to understand the biological relevance of the target as well its genetic background. In this study, we are therefore using a CRISPR-Cas9 domain based screen assessing a pooled epigenome sgRNA library in selected ESCC cell models [79], allowing the identification of a novel dependency, SMARCA4. Secondly, the aim is to link the vulnerability to a selective molecular feature, taking the latest literature on SL-interactions between SMARCA4 and SMARCA2 into account. To test whether SMARCA2 expression may serve as potential biomarker for SMARCA4 dependency and drug developing programs later on, SMARCA*2* expression levels are analyzed in selected cell lines. Therefore, SMARCA4 dependency scores from two recently published whole-genome screens [80, 81] are collected, compared to the SMARCA4 scores of the CRISPR-Cas9 domain-based screens and the dependency correlated to the expression of SMARCA2 using the *ordino* platform [314]. Thirdly, in order to investigate the importance of certain domains within the identified factor SMARCA4, diverse assays including “CRISPR-scan” and rescue experiments are conducted. For “CRISPR-scan”, sgRNAs are designed against various domains within SMARCA4, assessing the importance of the specific domains whereas ectopic expression of SMARCA4 variants bearing a mutation in the BD or ATP-binding domain are tested on their ability on compensation upon endogenous SMARCA4-knock-out (KO). In order to proof synthetic lethal interactions, the effects upon re-expression of the paralog SMARCA2, in a deficient cell line or KO of SMARCA2 in a proficient cell line are investigated. To extend the therapeutic concept to additional indications, depletion experiments in cell lines showing similar molecular features ($SMARCA2^{low}$) to the previously analyzed ESCC cell models are performed. To circumvent the fact that selective compounds for SMARCA4 are not available, a domain swap strategy is applied. In contrast to previous studies which constructed resistant alleles, this study aims to devise specific degradation of the identified dependency by a selective PROTAC which has originally been designed for BRD9. To finally determine the translation of molecular effect into pharmacological inhibition, the PROTAC treatment is applied to cells expressing the swapped chimeric protein ($SMARCA4\text{-}BD^{BRD9}$). Degradation of the protein as well as IC_{50} values are assessed.

The aims are framed as following:

1) Setting-up the CRISPR-Cas9 domain-based technology at Boehringer Ingelheim RCV, adopted from Vakoc group at Cold Spring Harbor Laboratory (New York, USA) [79]

2) Generation of a number of sufficient editing cell lines (depletion efficacy >10 fold)

3) Identification of a potential novel drug target in ESCC using the successfully set-up screening technology. Conduction of screens with a pooled sgRNA library directed against ~178 epigenetic factors (library design by Vakoc laboratory)

4) Validation of sgRNAs by “CRISPR-scans” targeting different domain-encoding regions within the identified candidate, SMARCA4. The aim is to get first insights into the relevance of certain domains and selection of the most efficient sgRNAs

5) Corroboration of the pooled CRISPR-Cas9 screen findings by depletion studies using individual sgRNAs

6) Conducting rescue experiments to confirm on-target activity

7) Assessment of the requirement of SMARCA4 helicase or BD function in SMARCA2-deficient ESCC cell lines by ectopic expression of SMARCA4 variants

8) Interrogation of the synthetic lethal interaction of SMARCA2 and SMARCA4
   a. SMARCA2-deficient cell lines: re-expression of SMARCA2 variants, analysis of SMARCA4-dependency reversion
   b. SMARCA2-proficient cell lines: KO of SMARCA2, assessing induced SMARCA4 vulnerability

9) Extension of the studies beyond ESCC cell models by selection of non-ESCC cell lines with low SMARCA2 mRNA expression and following depletion studies of SMARCA4-KO

10) Generation of a cell line with an allele amendable for PROTAC “dBRD9”

a. Overexpression of the chimeric protein SMARCA4-BD^{BRD9}
b. KO of endogenous SMARCA4
c. Analysis of degradation levels of SMARCA4-BD^{BRD9} and clone identification of SMARCA4-KO
d. Functionality test of SMARCA4-BD^{BRD9} expressing cell line (rescue) by application of sgRNAs targeting endogenous SMARCA4

11) Investigation if vulnerabilities determined by molecular tools can be translated into pharmacological inhibition (determination of IC_{50} for cell line carrying SMARCA4-BD^{BRD9} SWAP compared to parental cell line with endogenous SMARCA4)

2 Materials and Methods

2.1 Cell lines culturing

KYSE-30 and KYSE-450 (ESCC) were cultured in 45% RPMI 1640 (Gibco) +45% Ham's F12 +10% fetal calf serum (FCS). For cultivation of KYSE-70, and KYSE-410 (ESCC) RPMI 1640 (Gibco) supplied with 10% FCS, whereas for KYSE-140, KYSE-150, KYSE-180, KYSE-510, and COLO-680N (ESCC) RPMI1640 (ATCC #30-2001) +10% FCS was used. T.T (ESCC) cell line was cultured in DMEM:F12 (ATCC: 30-2006) containing 10% FCS. KYSE-270 (ESCC) was cultured in RPMI1640 (Gibco) +HAM's F12 (Gibco, 31765-027) (1:1) including 2mM Glutamine and 2% FCS. HCT 116 (colon carcinoma) cells was cultured in McCoy's 5A medium (Gibco, 36600-021) supplemented with glutamax and 10% FCS, SK-CO-1 (colon carcinoma) was cultured in EMEM (SIGMA, M5650) with glutamax and 10% FCS supplemented with sodium-pyruvate. For culturing of OV-90 (ovarian carcinoma) a 1:1 mixture of MCDB 105 medium (Sigma, M6395) with glutamine and hepes with a final concentration of 1.5g/L sodium bicarbonate and Medium 199 (Sigma, M4530) and a final concentration of 2.2g/L sodium bicarbonate was used. HuP-T4 (pancreas carcinoma) was cultured in MEM + Earl´s Salt (Gibco, 21090-022, no glutamine) + 20% FCS and glutamax. Lentiviral particles were generated via usage of the Lenti-X Single Shot protocol (Clontech, Mountain View, CA, US). For $Cas9^{+}$ cell lines the following concentrations of puromycin were added to the standard medium: T.T, and KYSE-270: 4µg/ml; KYSE-70, KYSE-140, HCT 116, and HuP-T4: 2µg/ml; KYSE-450, KYSE-510, SK-CO-1, and OV-90: 1µg/ml; COLO-680N: 0.5µg/ml; KYSE-30, KYSE-150, and KYSE-410: 0.25µg/ml. The medium for the cell lines expressing $SMARCA4^{res}$, $SMARCA2^{ect}$, or $SMARCA4\text{-}BD^{BRD9}$ was supplied with 25µg/ml hygromycin B. All the supplements were purchased from Gibco, FCS (SH30071.03) obtained from GE Healthcare Life Sciences, puromycin from Sigma P9620 and hygromycin B from Invitrogen, 10687010. SMARCA2 and SMARCA4 expression as well as mutation status is attached (appendix 6.4). Testing the indicated cell lines on mycoplasma contamination resulted negative. The STR fingerprint, which has been analyzed for all the engineered cell lines matched the respective parental cell lines.

2.2 Cas9 cell line generation

Freshly thawed cell lines were transduced with a multiplicity of infection (MOI) >1 with lentiviral Cas9-Puro construct (codon optimized Cas9 for applications in human derived cell lines) supplemented with polybrene/hexadimethrine bromide (8µg/ml final concentration) and selected over 2 weeks until they were frozen or used for further experiments.

2.3 Virus titration in order to determine transduction efficacy

To ensure that only one sgRNA was transduced per cell, a MOI of ~0.3 was aimed to be achieved. The pooled sgRNA supernatant was titrated in order to determine the amount of virus supernatant which was sufficient for the transduction of 30% of cells in a 6-well format. Therefore 600.000 cells were seeded per 6-well in 1ml of medium supplemented with polybrene/hexadimethrine bromide (8µg/ml final concentration). Different volumes of virus were titrated straight after the cell seeding (10-320µl) and the percentage of GFP^+ cells was measured at d3 using flow cytometry analysis (BD Accuri™ C6 instrument).

2.4 CRISPR epigenome screens

$Cas9^+$ ESCC cell lines were transduced with the sgRNA library (sgRNA sequences are attached: appendix 6.5) with a MOI~0.3 and cultured for ~18 population doublings. In order to maintain the sgRNA representation (1000 fold), 6 million cells were seeded initially using transducing ten wells of a 6-well plate, containing 600.000 cells/well respectively. The amount of viral supernatant was chosen according to the results obtained from titration, ensuring 30% of transduced cells in every well. Genomic DNA was purified using QIAamp DNA MiniKit (50) (Qiagen, 51304) and sequences around the sgRNA were PCR amplified using the following primers and PCR conditions:

Primer:

LRG_F2:TCTTGTGGAAAGGACGAAACACCG
LRG_R2:TCTACTATTCTTTCCCCTGCACTGT

Master Mix

The dNTP mix (10mM) was purchased from Roche (#11814362001), Phusion Hot start Flex DNA Polymerase Kit HF buffer obtained from NEB (#M0535L) and the PCR H_2O from Sigma (#03315843001). Primers were ordered from SIGMA as oligos.

	Per reaction
Genomic DNA	50-200 ng
10µM LRG_F2	1.5 µl
10µM LRG_R2	1.5 µl
10mM dNTPs	1.5 µl
HF buffer	10 µl
HS flex enzyme (ice)	0.5 µl
PCR water	fill up to 50 µl

PCR conditions:

98°C 5min
98°C 8sec, 60°C 20sec, 72°C 10 sec, 27 cycles
72°C 5min
Final 4°C

The PCR product of 40 PCR reactions was pooled and purified using QIAquick PCR Purification Kit (250) (Qiagen, 28106) and 50ng of amplicons were applied for the library generation. Therefore, the TruSeq Nano DNA Library Prep kit for NeoPrep (Illumina) was used. For sequencing, HiSeq1500 in rapid mode with the paired end protocol was conducted with 50 cycles.

2.5 Bioinformatic analysis

Statistical analysis of depletion signals was performed with the MAGeCK tool (V 0.5.6) [315, 316]. First guide level counts were generated from sequencing data with the 'mageck count' function with parameter '—norm-method control'. The set of negative control guides was derived from genes that never show strong depletion signals in the AVANA [80] data set and that overlap with genes in our library. Next the 'mageck test' function was run with parameters '—remove-zero none –gene-lfc-method median' to derive gene-level α-RRA scores for each cell line. To improve comparability between the cell lines we scaled α-RRA scores by using three positive control genes (*CDK1, POLR2A, RPA3*) such that the mean of these control genes was -1 for all cell lines.

2.6 CRISPR singleton-gRNA depletion experiments

The sgRNA sequences were adopted from the library or designed using the MIT tool (http://crispr.mit.edu/) and cloned into GFP encoding vectors. $Cas9^+$ cells were transduced and GFP^+ cells were measured upon day 3 post-transduction. Cells were split for 21-28 days. The fold changes of the initially transduced cell population were calculated from d3 to the respective time points. For summary depletion, $\%GFP^+$ cell population on d28 was relativized to $\%GFP^+$ cells treated with positive control sgRNA, POLR2A on d28.

2.7 Sequences

sgRNAs for depletion experiments		
Name	**Target gene**	**Sequence**
Control sgRNAs		
Neg. Contr.	none	GATACACGAAGCATCACTAG
POLR2A	POLR2A	GTACAATGCAGACTTTGACG
SMARCA4 sgRNAs (N- to C- terminal order)		
SMARCA4_N-term_e2-1	SMARCA4	TGGCCGAGGAGTTCCGCCCA
SMARCA4_N-term_ e2-2	SMARCA4	CTGGCCGAGGAGTTCCGCCC
SMARCA4_N-term_ e2-3	SMARCA4	GGCCGAGGAGTTCCGCCCAG
SMARCA4_N-term_ e2-4	SMARCA4	CCGGCGAGGGACCCGGGCTA
SMARCA4_DEXDc_e16	SMARCA4	GAGGTACGTGATGAGCGCGA
SMARCA4_DEXDc_e17	SMARCA4	GTCAAACTCGTACGCCCAGT
SMARCA4_DEXDc_e18	SMARCA4	TGAACTTCCCACTCCGGAGC
SMARCA4_DEXDc_e19	SMARCA4	GAACAAGCTTCCCGAGCTCT
SMARCA4_HELIC_e24.1	SMARCA4	GTGGTTGGTTGCTCGGAGTT
SMARCA4_HELIC_e24.2	SMARCA4	GAAGATTACTTTGCGTATCG
SMARCA4_HELIC_e25.1	SMARCA4	CTGAAAACCTTCAACGAGCC
SMARCA4_HELIC_e25.2	SMARCA4	TGATCACAGTGTCTGCCGAC
SMARCA4_BD_e32_109.8	SMARCA4	CTCGGGCAGCTCCTTTCGCG
SMARCA4_BD_e32_109.9	SMARCA4	TCGGGCAGCTCCTTTCGCGA
SMARCA4_BD_e33_102.5	SMARCA4	GGTTGAAGGTCTGTGCGTTC
SMARCA4_BD_e34_59.1	SMARCA4	AGTCGGTCTTCACCAGCGTG
SMARCA4_BD_e34_59.2	SMARCA4	GAAGACCGACTGCAAGACGA
sgRNAs for gene knock-outs		
SMARCA2_N-term_e2-1	SMARCA2	TCCCATCCTATGCCGACGAT
SMARCA4_N-term_e2-4	SMARCA4	CCGGCGAGGGACCCGGGCTA

Table 1: sgRNA sequences for depletion experiments as well as for monoclonal cell line generation.

2.8 Quantitative reverse transcription PCR (qRT-PCR)

The RNA was extracted from cell pellets using RNeasy Mini Kit (Qiagen, 74106) and reverse transcribed utilizing SuperScript™ VILO™ kit (Thermo Scientific). For qPCR analysis, QuantiTect® Multiplex PCR kit (Qiagen, Hilden, Germany) and StepOne Real-Time PCR Sytem™ (Applied Biosytems) were used. The respective primers were ordered from Applied Biosystems, including house-keeping genes: 18S rRNA (VIC®/MGB, 4319413E), ACTB (VIC®/MGB, 4326315E), GAPDH (VIC®/MGB, 4326317E) and primers for SMARCA2 (Hs01030858_m1 MGB/FAM) and SMARCA4 (Hs00231324 MGB/FAM), respectively. SMARCA2/4 expression was calculated from duplicates in relation to three different house-keeping genes listed above.

2.9 Capillary Western immunoassay

Lysates were generated using MSD Tris lysis buffer (Mesoscale #R60TX-2) and supplied with Protease & Phosphatase Inhibitor Cocktail (1:100, Thermo Scientific#815-968-0747). Capillary Western immunoassay (Separation module, SMW004-1) was conducted according to manufacturer's protocol. Dilutions of protein were prepared in order to obtain a final protein concentration of 0.4µg/µl.

2.10 Antibodies

Anti-BRG1/SMARCA4 (Cell Signaling #49360, 1:20 dilution capillary Western immunoassay); anti-SMARCA2 (SIGMA #HPA029981, 1:20 dilution capillary Western immunoassay); anti-BRD9 (Bethyl Lab. A303-781A-M, 1:20 dilution capillary Western immunoassay). Respective antibodies were used in a multiplex assay together with anti-GAPDH (abcam #ab9485, 1:10000 dilution capillary Western immunoassay) to ensure equal loading in every single capillary.

2.11 Bioinformatic analysis: SMARCA4 dependency correlation with SMARCA2 expression

SMARCA2 mRNA expression values (TPM – transcripts per million) were obtained from ordino (https://ordino.caleydoapp.org/ [314]). SMARCA4 sensitivity scores were obtained from McDonald III *et al.* [81] (RSA scores) and Meyers *et al.* [80] (Ceres scores). The visualizations and statistical tests were performed using R version 3.5.0 (R Core Team (2018) R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.; available online at https://www.R-project.org/, including the R package ggplot2 version 3.0.0 (H. Wickham (2016) ggplot2: Elegant Graphics for Data Analysis; doi: 10.1007/978-3-319-24277-4).

2.12 cDNA transgene vectors

For rescue experiments (SMARCA4^{res}) or proof of the synthetic lethal (SL) interaction (SMARCA2^{ect}) the following constructs were generated by gene synthesis (GenScript, China) based on the SMARCA4 cDNA sequence NCBI NM_ 001128844.1 and SMARCA2 cDNA sequence NCBI NM_003070.5 followed by cloning into the parental pLVX vector (Clontech, Mountain View, CA, US): pLVX-empty-IRES-Hygro; pLVX-SMARCA4-wt-IRES-Hygro; pLVX-SMARCA4-ATPase-binding deficient (K785A)-IRES- Hygro; pLVX-SMARCA4-BD-dead(N1540A)-IRES-Hygro; pLVX-SMARCA2-wt-IRES-Hygro; pLVX-SMARCA2-ATPase-dead(K755A)-IRES-Hygro; pLVX-SMARCA4-BD-dead(N1482A)-IRES-Hygro. ATP-binding deficient and BD-dead sites were selected similar to known inactivating mutations in SMARCA2/4 [300, 313]. For SMARCA4, DEXDc (DEAD-like helicases superfamily) and helicaseC domains were annotated according to UniProt entry P51532, bromo domain (BD) was annotated according to NCBI entry 6597 (cd05516). For SMARCA2, domains were annotated according to UniProt entry P51531, BD was annotated according to NCBI entry 6595 (cd05516).

For the bromo swap study, pLVX-SMARCA4-BD^{BRD9}-IRES-Hygro construct was ordered from GenScript, China. The exact sequence exchange is shown below (2.15). All of the constructs were codon optimized to render them resistant towards siRNAs as well as sgRNAs.

2.13 siRNA-mediated knock-down

For siRNA-mediated knock-down (KD), cells were reverse transfected with siRNA using Lipofectamine RNAiMAX reagent (Thermo Fisher Scientific; 13778075) according to the manufacturer's protocol. The final concentration of the respective siRNAs was 25nM. The siRNA constructs were purchased from Thermo Scientific/ Dharmacon: Non-targeting pool (D-001810-10-20), SMARCA4-1 (J-010431-06), SMARCA4-2 (J-010431-07).

2.14 Knock-out and monoclonal cell line generation

All-in-one pSpCas9 BB-2A-GFP (PX458) vectors were ordered from Genscript, China. The sgRNAs directed against the N-terminus of SMARCA2/4 were used and transfected with Lipofectamine 3000 reagent (Thermo Fisher Scientific, L3000015). Successfully transfected GFP^+ cells were afterwards sorted using Sony, SH800S cell sorter. In order to generate a monoclonal cell line. Limited dilution of the bulk sorted GFP^+ cells were prepared seeding 0.3 cells per 96-well. Retrieved cell lines were analyzed according to their respective KO using capillary Western immunoassay.

2.15 SWAP cell line generation

The sequence of the SMARCA4 BD was exchanged by BRD9 encoding BD including flanking regions in order to maintain structural formation of the chimeric protein. Domains are annotated according to NCBI entry cd05516 (SMARCA4-BD) and UniProtKB entry Q9H8M2 (BRD9). Amino acid sequences selected for the swap are indicated in the table below. The construct pLVX-SWAP-IRES-Hygro was purchased from GenScript, China. After stable expression of the SMARCA4-BD^{BRD9} construct and hygromycin selection, KO of endogenous SMARCA4 using transient expression of the all in one vector pSpCas9 BB-2A-GFP (PX458) (Genscript, China) was performed. Cells transduced with the construct were separated from the pool by using Sony, SH800S cell sorter and following single cell cloning. Positive clones were confirmed according to loss of SMARCA4 protein on capillary Western immunoassay upon dBRD9 treatment (1µM).

Protein	Sequence
SMARCA4	AEKLSPNPPNLTKKMKKIVDAVIKYKDSSSGRQLSEVFIQLPSRKELPEYYELIRKPVDFKKIKE RIRNHKYRSLNDLEKDVMLLCQNAQTFNLEGSLIYEDSIVLQSVFTSVRQKIEKEDD
BRD9	AENESTPIQQLLEHFLRQLQRKDPHGFFAFPVTDAIAPGYSMIIKHPMDFGTMKDKIVANEYKS VTEFKADFKLMCDNAMTYNRPDTVYYKLAKKILHAGFKMMSKERLLALKRS

Table 2: Swap sequences of SMARCA4 and BRD9.

Substituted amino acids of SMARCA4 and BRD9 BDs in the SMARCA4^{res}-BD^{BRD9} variant (amino acids representing BDs in SMARCA4 (NCBI 6597, cd05516) and BRD9 (UniProt Q9H8M2) are indicated in red font).

2.16 Sequencing for KO confirmation

Genomic DNA was purified using QIAamp DNA MiniKit (50) (Qiagen, 51304) from respective samples and Sanger Sequencing was performed by Eurofins Genomics AT GmbH.

2.17 Cell viability assay

KYSE-30-SWAP and parental KYSE-30 cell lines were cultured for approximately 18 passages post-thawing and plated at low densities (100 cells/well) in a 96-well plate. After 24h incubation, cells were treated with dBRD9 (maxima concentration= 10.000nM) [310]. Viability was determined using CellTiter-Glo (Promega, Madison, WI, US). Ten days post-treatment, CellTiter-Glo solution was mixed to the cells and incubated for ten minutes before luminescence signal was measured.

3 Results

3.1 Generation of Cas9-stably expressing cell lines and pre-tests necessary for the pooled sgRNA library screen

For the identification of novel targets in ESCC, a pooled library screening strategy based on CRISPR-Cas9 domain-directed sgRNAs adopted from Shi *et al.* [79] was applied. To this end, Cas9-stably expressing cell lines were generated by lentiviral transduction followed by puromycin selection over a time period of two weeks. Derived cell lines were further tested according to evaluate their editing efficacy. Therefore, Cas9^{+} cells were transduced with lentiviral constructs encoding the respective sgRNA as well green fluorescent protein (GFP). Whereas cells transduced with sgRNAs targeting genes non-essential for cellular fitness remain in the population, cells transduced with sgRNAs targeting essential genes are depleted (negative selected) (Figure 13). A decrease in fraction of GFP positive (GFP^{+}) cells over time serves as readout for the determination of cell essential genes (Figure 13).

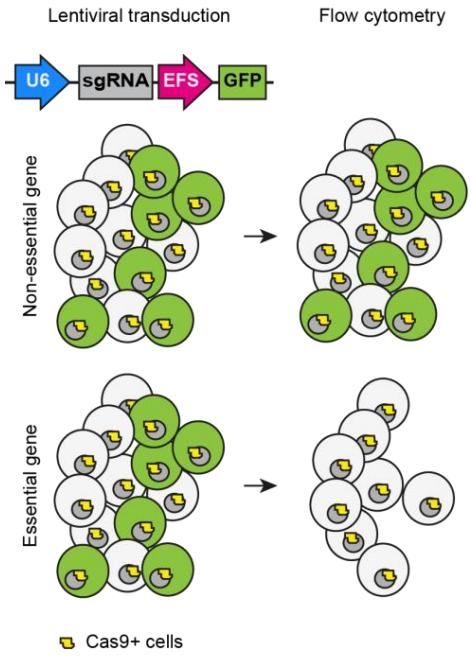


Figure 13: Scheme of CRISPR-Cas9 sgRNA depletion experiment.

Cas9^{+} cells are transduced with sgRNAs and GFP encoding vectors using a lentiviral delivery system. Cells which are transduced with the respective sgRNA also express GFP which allows monitoring via flow cytometry. Whereas cells transduced with sgRNAs targeting non-essential genes maintain in the cell population, cells infected with sgRNAs directed against essential genes are depleted, indicating a dependency on the respective gene.

To evaluate the editing efficacy, stably Cas9 expressing ESCC cell lines including COLO-680N, KYSE-30, KYSE-70, KYSE-140, KYSE-150, KYSE-270, KYSE-410, KYSE-450, KYSE-510, and T.T were profiled in sgRNA depletion assays (Figure 14 and Figure 15). The effects of sgRNA-mediated targeting of cell essential genes *CDK1* (Cyclin-dependent kinase 1), *PCNA* (Proliferating-Cell-Nuclear-Antigen), *POLR2A* (RNA Polymerase II Subunit A) and *RPA3* (Replication Protein A3) were assessed in KYSE-30, KYSE-70, KYSE-140, KYSE-150, T.T, and KYSE-410 (Figure 14 and Figure 16) compared to a negative control sgRNA (Neg. Contr.), not matching any sequence in the human genome. In order to exclude the presence of two sgRNAs in one cell, which would preclude assessment of individual sgRNA activity, the cells were transduced with a virus concentration resulting in a MOI of approximately 0.3. The fraction of GFP^+ cells relative to day three (d3) post-transduction was compared to different time points over a period of three weeks (Figure 14). Whereas the population of cells harboring a negative control sgRNA remains constant over time, the cells which are transduced with positive control sgRNAs are efficiently depleted. Very strong effects are observed already after nine days post-transduction demonstrating a sufficient Cas9 editing efficacy (Figure 14). As an exception, in KYSE-30 the best sgRNA is represented by RPA3_e1.3, whereas still approximately 20% of GFP^+ cells are maintained in the population (Figure 14). For all the positive controls, a remaining GFP^+ fraction is detected in KYSE-30, indicating semi-sufficient depletion on day 22 (Figure 14). In addition, PCNA_e2.1 depletes less efficiently compared to the other controls. For KYSE-70, KYSE-140 and KYSE-150 almost comparable depletion efficacies are measured (Figure 14). The respective cell lines depict, as expected, very little decrease in GFP^+ population using the negative control sgRNA but rapid and strong depletion is observed when assessing the positive control sgRNAs. Remarkably, the most pronounced and fastest drop-out of the GFP^+ cell population is detected after the application of sgRNA targeting *POLR2A_e10.1* (Figure 14). In the T.T cell line, the editing efficiency depends on the sgRNA used. The weakest effects are detected using RPA3_e1.3, intermediate effects are shown for PCNA_e2.1, PCNA_e3.2 and very strong depletion is measured for POLR2A_e10.1 and CDK1_e5.1 (Figure 14). In contrast to the other cell lines which show already effects on depletion upon d6, KYSE-410 starts depleting at d9. Interestingly, the percentage of GFP^+ cells increases for the

controls POLR2A_e10.1, RPA3_e1.3 and CDK1_e5.1 at d6 but sufficiently drops afterwards (Figure 14).

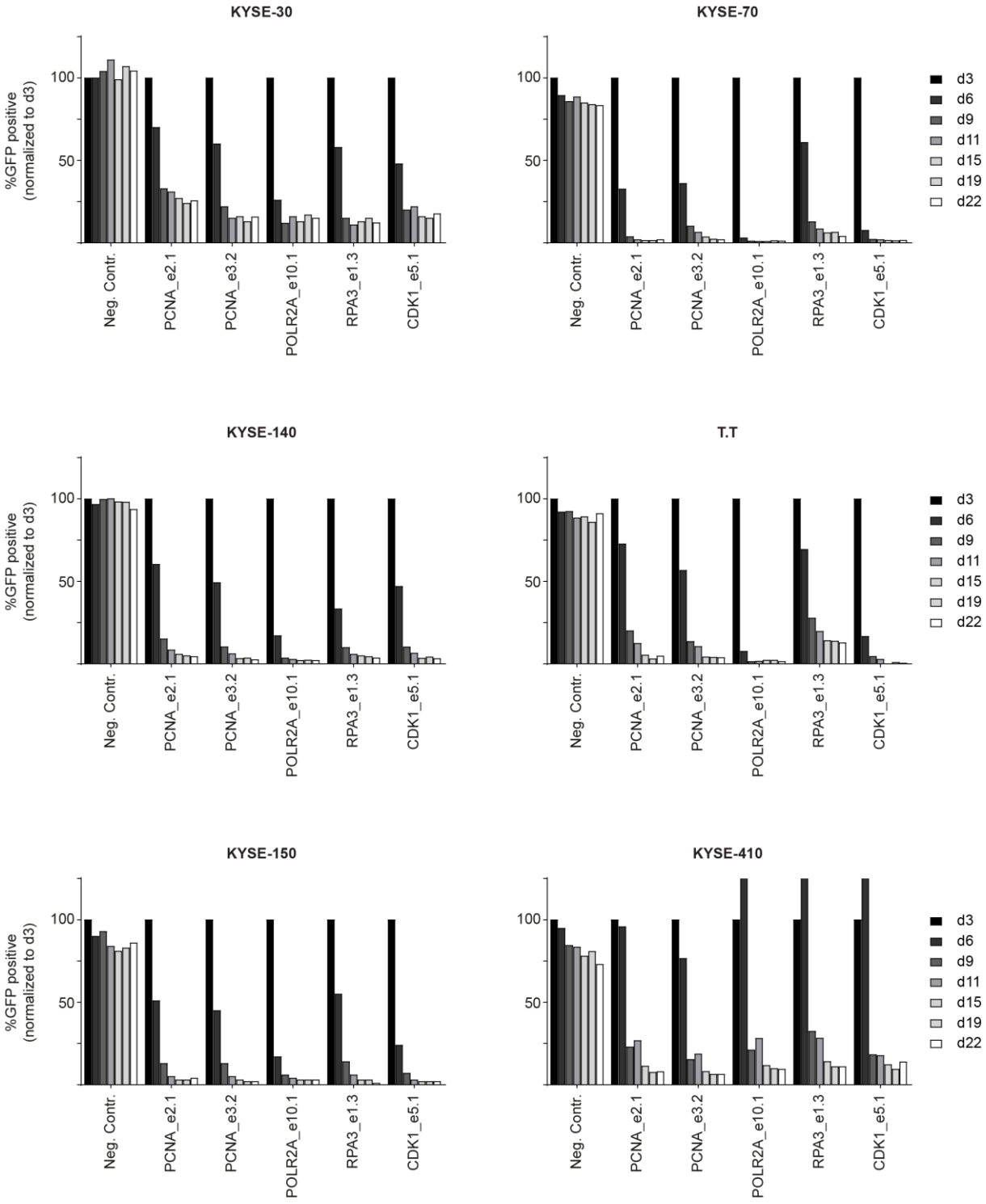


Figure 14: Depletion efficacy in ESCC cell lines using five different positive control sgRNAs.

Six ESCC cell lines were tested for depletion efficacy using the positive controls (PCNA_e2.1, PCNA_e3.2, POLR2A_e10.1, RPA3_e1.3, and CDK1_e5.1) compared to the negative control (Neg. Contr.). Every bar indicates a time point of GFP⁺ cell population measurement determined by flow cytometry from d3 until d22. The x-axis depicts the different sgRNAs used. The initial percentage of GFP⁺ cell population was set to 100% and every measurement afterwards normalized to the initial GFP⁺ percentage measured.

The most efficiently depleted control is represented by *POLR2A*, indicating pronounced effects in all of the six cell lines tested (Figure 14). Consequently, for the validation of additional ESCC cell lines, only POLR2A sgRNA was assessed. The additional Cas9-stably expressing cell lines are represented by KYSE-510, KYSE-270, COLO-680N and KYSE-450 (Figure 15). Whereas for KYSE-510, KYSE-270 and COLO-680N strong effects on depletion are observed upon d7, comparable effects are seen in KYSE-450 upon d10. The fraction of GFP$^+$ cells targeted by the negative control sgRNA remains constant over the time period of three weeks. A small decrease in the negative control targeted population is observed in KYSE-450 on d7 and d10 but reaches a plateau for the rest of the data points, excluding off-target activity (Figure 15).

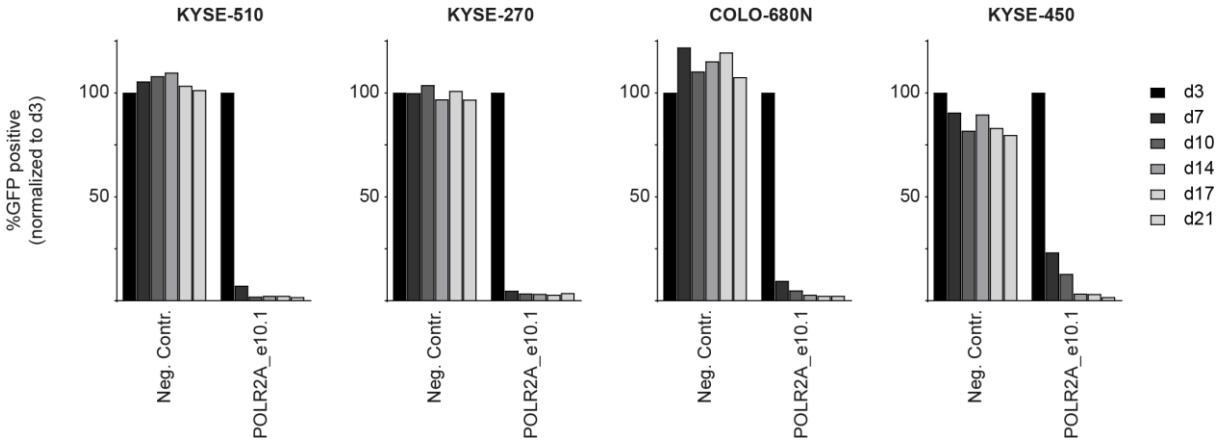


Figure 15: Testing editing efficacy via the application of POLR2A sgRNA in additional ESCC cell lines.
Four additional cell lines KYSE-510, KYSE-270, COLO-680N and KYSE-450 are tested with the selected positive control (POLR2A_e10.1) compared to negative control (Neg.Contr.) sgRNA treatment depicted on the x-axis. Every bar indicates a time point of GFP$^+$ cell population measurement determined by flow cytometry at d3 until d21. The amount of GFP$^+$ cell population is set to 100% and every measurement afterwards is normalized to the initial GFP$^+$ percentage measured.

To better visualize the editing efficacy effects, fold depletions were calculated by dividing the %GFP$^+$ of d3 by each measurement (day x) and depicted in Figure 16 and Figure 17. The pre-requisite for a screen was defined by a ten-fold editing efficacy when applying the positive control sgRNAs. Pronounced depletion efficacies allow a better separation of essential versus non-essential genes ensuring that only strong dependencies are received from a drop-out screen later on.

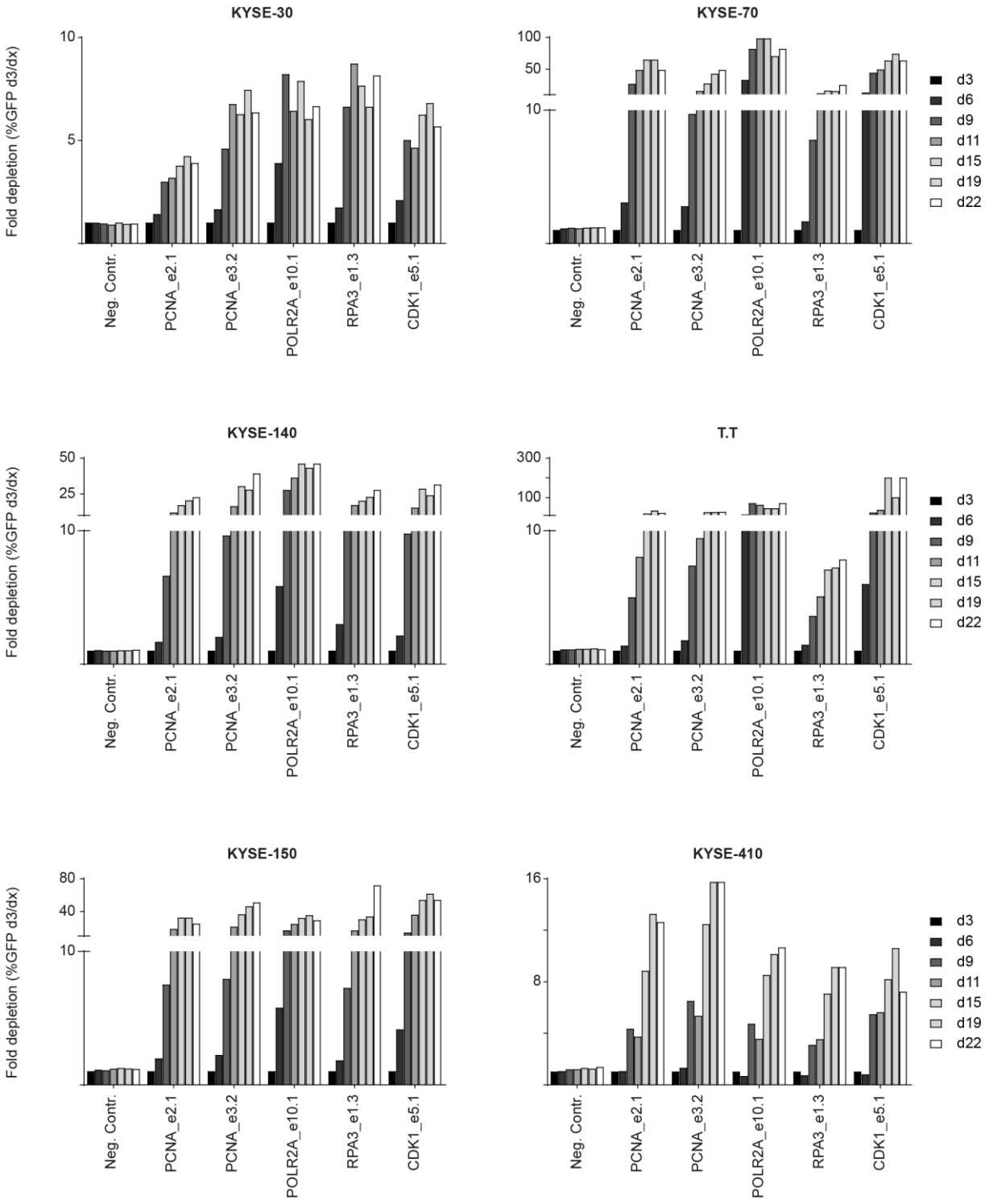


Figure 16: Fold depletion values in ESCC cell lines using five positive controls.

Depletion efficacies are re-analyzed using fold-changes to determine editing in KYSE-30, KYSE-70, KYSE-140, T.T, KYSE-150, and KYSE-410. Every bar shows a time point indicating the fold depletion of the GFP$^+$ cell population assessed by flow cytometry analysis compared to GFP$^+$ percentage at d3. Fold depletion at d3 is (per definition) set to 1. Values higher than 1, indicate dependency whereas values close to 1 are linked to no effect. The x-axis depicts the different sgRNAs used including Neg. Contr., PCNA_e2.1, PCNA_e3.2, POLR2A_e10.1, RPA3_e1.3, and CDK1_e5.1.

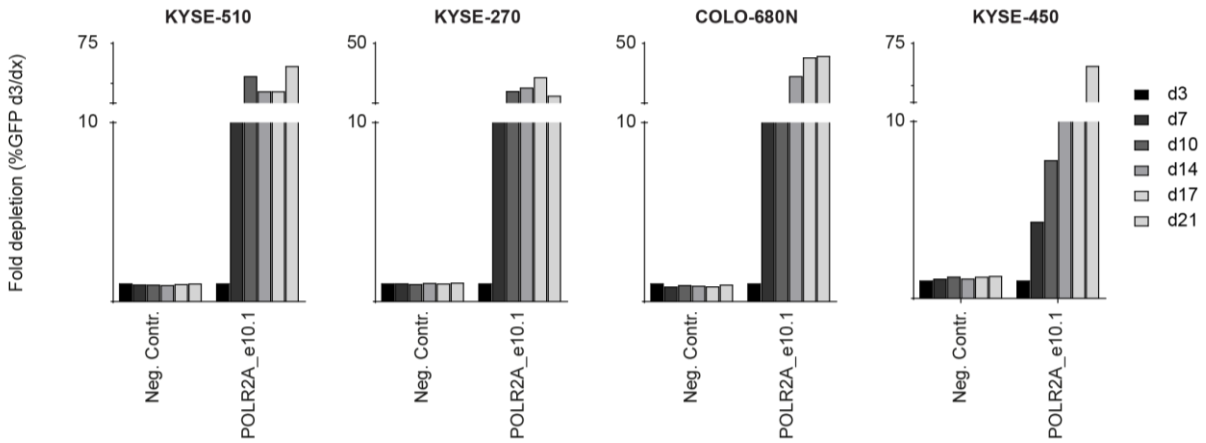


Figure 17: Fold depletions in additional ESCC cell lines using POLR2A.

Fold depletions of %GFP⁺ fraction on day x (x = individual time point) relative to d3 in KYSE-510, KYSE-270, COLO-680N, and KYSE-450 are depicted. Only two sgRNAs are applied in this setting using one negative control (Neg. Contr.) as well as POLR2A_e10.1 as positive control. Every bar represents a time point.

The depiction of fold-changes, rather than percentage of GFP⁺ cells, discriminates between weak and strong effects on negative selection. KYSE-70, KYSE-140, T.T, KYSE-150, KYSE-510, KYSE-270, COLO-680N, and KYSE-450 are confirmed to deplete more than 10-fold when testing the positive control sgRNAs (Figure 16 and Figure 17). Contradictory, none of the positive controls allow a strong depletion of the GFP⁺ cell fraction in KYSE-30 and only *PCNA* targeting sgRNAs are efficiently outcompeted in KYSE-410 cell line (Figure 16). However, depicting fold changes reveals a robust window between the applied negative and positive control sgRNAs in KYSE-30 and KYSE-410 cell lines (Figure 16 and Figure 17). In order to investigate a broader cell panel, also low editing cell lines KYSE-30 and KYSE-410 were implemented in the screens, keeping their moderate editing efficacy in mind.

3.2 CRISPR-Cas9 domain-based screen

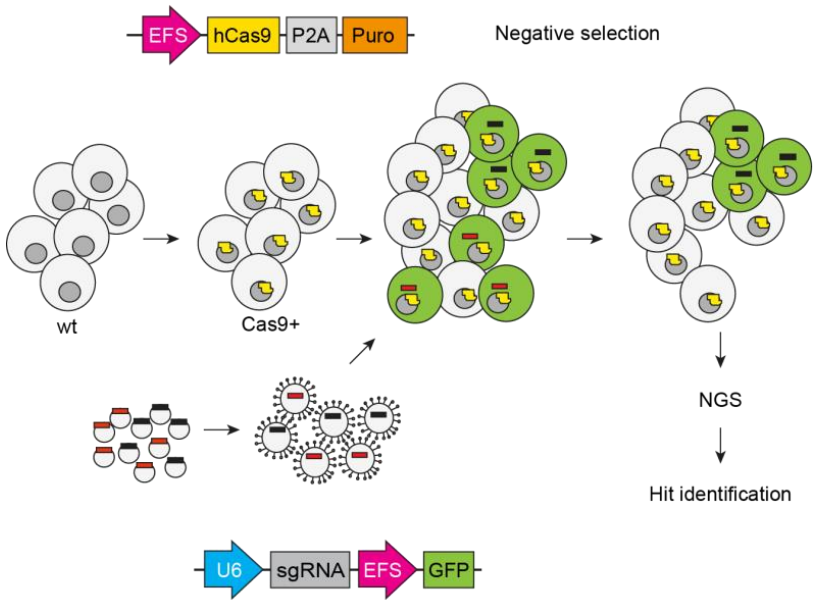


Figure 18: Schematic overview of the screening procedure.

Cas9⁺ cell lines were generated via transduction of parental cell lines using lentiviral constructs and selection of approximately two weeks. Pooled plasmids were amplified, packaged and transduced always ensuring a 1000-fold representation of the constructs. The sgRNA as well as the GFP were encoded by the same vector allowing the monitoring of positively transduced cells aiming to achieve an initial transduction efficacy of 30%. Thereby the presence of only one sgRNA per cell is ensured excluding false positive depletion. The cells were cultured for 18 population doublings. Genomic DNA was purified from plasmid library and the last passage. After the amplification of sgRNA sequence via PCR, samples were labelled with barcodes to perform next-generation sequencing.

To identify novel vulnerabilities in ESCC, a pooled CRISPR-Cas9 library (appendix 6.5) was used encompassing ~1500 gRNAs covering 179 epigenetic regulators [317]. Therefore, the validated Cas9⁺ cell lines were assessed in the drop-out screen by lentiviral transduction of the pooled library. To avoid false positive hits, coverage of a 1000-fold representation of every sgRNA in the library was aimed for and maintained over the different cell passages. Thus, for the epigenome library, encompassing ~1500 sgRNAs, a 1000-fold representation was achieved by transducing 1,5 million cells initially. The use of a low MOI (~0.3) prevents depletion of cells simultaneous infected by a sgRNA targeting an essential gene. Hence, a minimum of 5 million cells is suggested, decreasing the chance of losing the sgRNA representation during cell passaging.

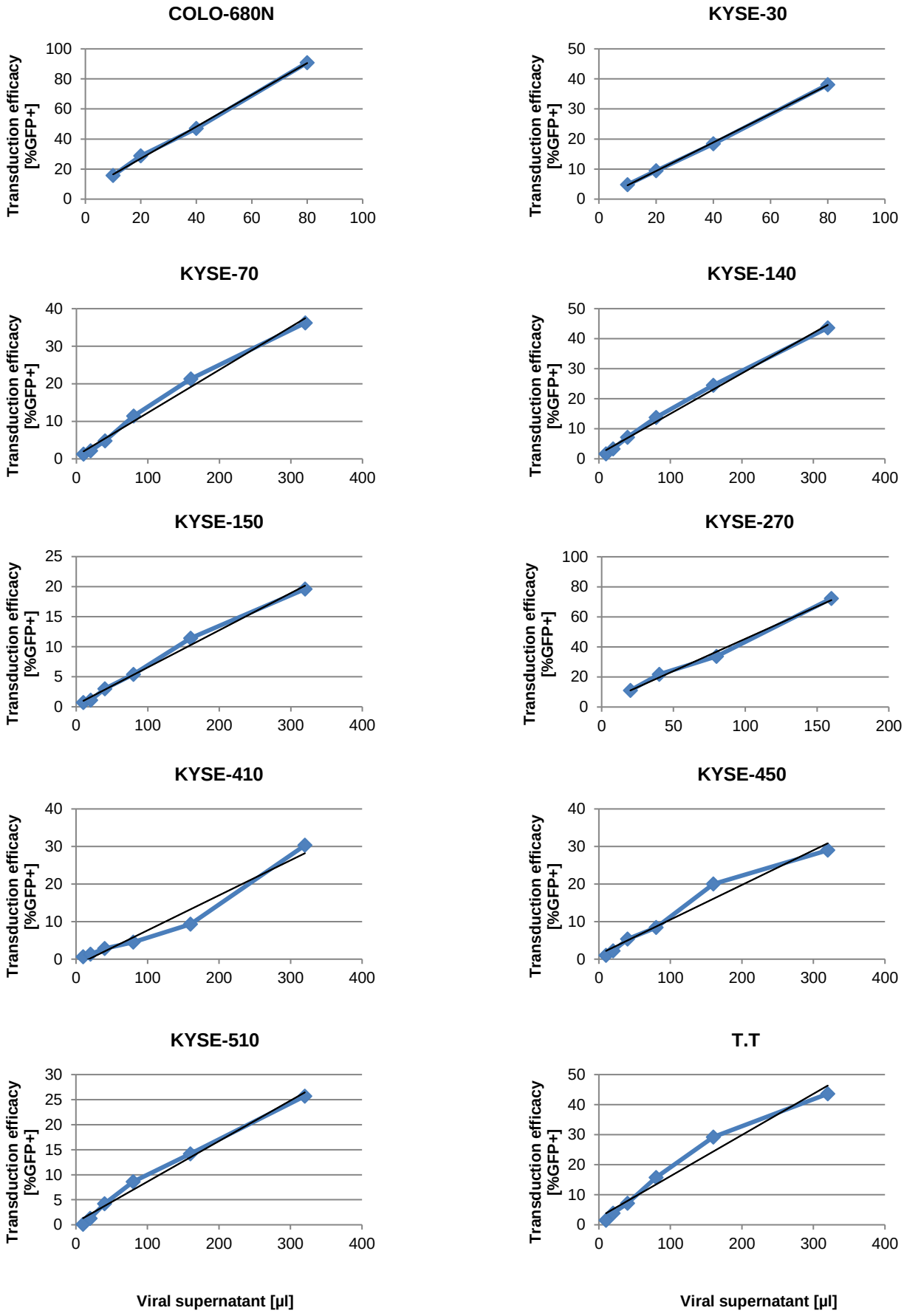


Figure 19: Titration of the pooled sgRNA-library.

Assessing the volume of viral supernatant necessary for efficient transduction of 30% of cells.

Cell line	Viral supernatant[µl]
COLO-680N	30
KYSE-30	80
KYSE-70	300
KYSE-140	250
KYSE-150	450
KYSE-270	70
KYSE-410	320
KYSE-450	350
KYSE-510	400
T.T	200

Table 3: Selected volume of virus supernatant of the pooled library for obtaining 30% transduced cells.

Similarly to the sgRNA depletion experiments, the expression of an individual sgRNA of the library was linked to GFP expression. However, the GFP^+ cell fraction measurement only allows the estimation of the MOI which was aimed to be less than 0.3. The presence of only one sgRNA per cell is relevant in order to exclude off-target effects. Additionally, the decrease of the GFP^+ expressing cell fraction ensures occurrence of deleterious events. To identify potential candidates, genomic DNA from the plasmid library as well as from the last time-point (~18 population doublings) was purified. After the amplification of the respective sequences using PCR, next-generation sequencing (NGS) was applied generating read counts for the individual sgRNAs tested. The data was further processed calculating the log2-fold changes (lfc) from the read counts of the samples from passage eight to plasmid DNA. To generate gene-based scores, individual sgRNA scores were summarized and a bioinformatics quality check was performed resulting in robust ranking aggregation scores (α-RRA) (see 2.5).

To test how much library virus was required to obtain a fraction of approximately 30% transduced cells, titration of the pooled library virus supernatant was conducted (Figure 19). Therefore, increasing volumes of the pooled library viral supernatant were added to the cells and the fraction of GFP^+ cells was analyzed via flow cytometry at d3 (Figure 19). Depending on the cell line, various amounts of virus supernatant were sufficient to infect the desired number of cells (Table 3). Whereas for COLO-680N 30µl, for KYSE-270 70µl and for KYSE-30 80µl viral supernatant were needed to obtain the respective percentage of GFP^+ cells, ~300µl of virus

supernatant were required for KYSE-70, KYSE-140, T.T, and KYSE-410. KYSE-150 and KYSE-510 depict two cell lines which were difficult to transduce with a need of 400-450µl virus supernatant.

For the pooled depletion screens, the pre-determined volume of viral supernatant was applied to the cells providing the same conditions as tested, using 600.000 cells per well in a 6-well plate. Instead of a single-well transduction, multiple wells were employed in order to ensure sgRNA representation. Therefore, ten 6-wells were assayed, seeding 600.000 cells per well with the aim of transducing a final number of 6 million cells.

As a representative sample for the screening procedure and intermediate steps, KYSE-510 was selected. All of the other ESCC cell lines assayed, showed comparable effects in terms of screening behavior such as GFP^+ depletion (Figure 20), test-PCR (Figure 21) and parallel PCR (Figure 22). After the transduction of multiple wells, the fraction of GFP^+ cells was observed over time using flow cytometry. The actual percentage of GFP^+ cells varies slightly compared to the numbers obtained by titration pre-test (Figure 20). In line with the representation of non-essential and essential sgRNAs in the library, the initially transduced GFP^+ population decreases slightly over time (Figure 20). Most notably, the GFP population measurements provide information whether cells are depleted at all. Even though the virus supernatant was titrated prior to the actual screen, for some cell lines different initial percentage of GFP^+ cells was obtained. The screens were therefore started with a cell population of GFP^+ cells between 20-40%.

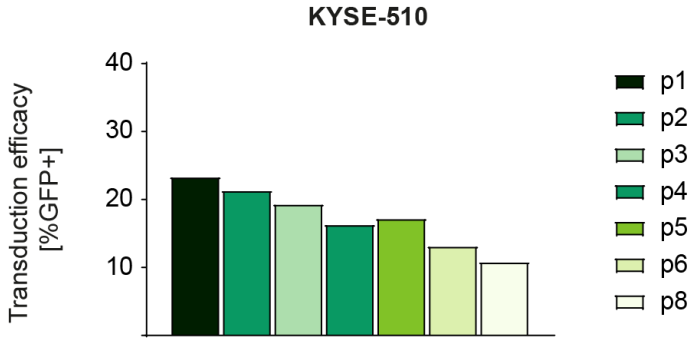


Figure 20: Depletion of pooled-sgRNA library over eight passages measuring GFP^+ cell population.

The fraction of GFP^+ cells was measured over time in KYSE-510 after transduction with the pooled library. The individual bars represent different passage; GFP^+ population was not measured at p7. Y-axis shows the % of GFP^+ cells; p= passage.

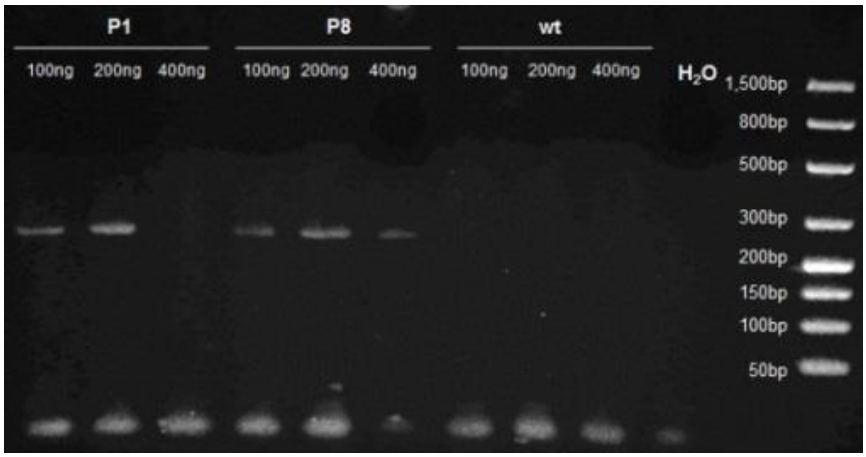


Figure 21: Determination of optimal amount of DNA via test PCR.

Different amounts including 100, 200 and 400ng of genomic DNA were tested in PCR to assess sufficient concentration for sgRNA amplification. The PCR was conducted in samples from passage 1 (P1), passage 8 (P8) and wt (=wild-type) as well as H_2O control. Expected band size is 260bp.

For setting up the screening methodology, genomic DNA (gDNA) was initially purified from an early (P1) and late passage (P8). In order to optimize PCR conditions for further sequencing, test PCRs were conducted. For amplification of the sgRNAs derived from the cell population, primers binding the flanking sequence were designed. To maintain sgRNA representation, multiple PCRs were conducted (40 parallel PCRs for the epigenetic library). The test PCR was important to determine the optimal amount of genomic DNA, which lead to a sufficient PCR product without impairing the polymerase and therefore inhibit the PCR reaction (Figure 21). For KYSE-510, taken as representative example, 200ng for both, passage one (P1) and passage eight (P8) were sufficient to achieve a clear band on a 2.2% DNA gel. The sequencing results from P1 were initially used for the calculation of fold-changes but regarding the fact that essential genes are already depleted upon day three, read-counts for plasmid DNA were pursued for further calculations. No band in DNA from non-transduced cells as well as for the H_2O control is detected, excluding contamination of different genomic DNA (Figure 21). Conduction of 40 parallel PCRs under the same conditions as the test-PCR allows maintenance of sgRNA representation. The final PCR product from all reactions was controlled on a 2.2% DNA gel (Figure 22), pooled and purified. Afterwards a sequencing library was generated, and samples were sequenced using HiSeq. The final concentration was measured on NanoDrop (KYSE510: P1= 26.72ng/µl, P8= 20.81ng/µl).

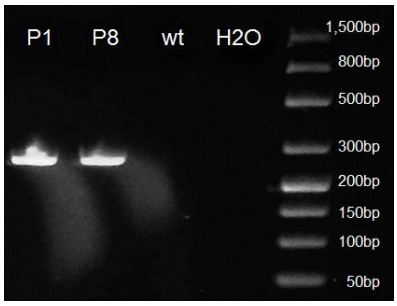


Figure 22: Confirmation of PCR product amplification for sequencing via parallel PCR.

In total, 200ng of genomic DNA was used for parallel PCR (40 reactions). Samples from passage 1 (P1) and passage 8 (P8) were pooled (band size~260bp), respectively, purified and analyzed on a 2.2% agarose gel. Controls are depicted by wt (wild-type) and H_2O.

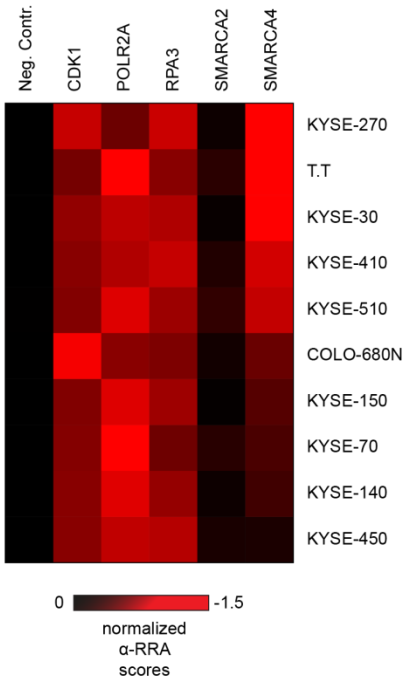


Figure 23: α-RRA (robust ranking aggregation) scores for ten ESCC cell lines obtained from CRISPR-Cas9 domain-based screen.

Selected genes targeted by sgRNAs are depicted on top of the heatmap. ESCC cell lines are shown on the right side. α-RRA scores are normalized to the mean of three positive control sgRNAs including CDK1, POLR2A, and RPA3 (mean=-1). Scores <-0.5 are regarded as dependencies whereas scores ≥-0.5 are linked to independencies. Low scores are depicted in red and high scores in black. One Neg. Contr., three positive controls CDK1, POLR2A, RPA3 as well as selected scores for sgRNA targeting SMARCA2 and SMARCA4, respectively, are depicted. Cell lines are listed according to their normalized SMARCA4 α-RRA score, showing the most dependent cell lines on top; scores from the screens were kindly provided by Andreas Schlattl.

From the sequencing robust ranking aggregation scores (α-RRA scores) were obtained. Importantly, all of the previously tested positive control sgRNAs from individual depletion experiments exhibited low α-RRA scores in the domain-based screens (Figure 23 and appendix 6.6). In addition, strong dependencies of pan-essential genes such as *BRD4*, *CDK1*, and *HDAC3* are detected (appendix 6.6). The application of the epigenome domain-based screen identifies SMARCA4, one of the two ATPases incorporated into the SWI/SNF complex, as a selective vulnerability in the ESCC cell lines used (Figure 23 and appendix 6.6). The α-RRA score of each individual gene was further normalized to the mean of the three positive controls CDK1, POLR2A, and RPA3, in order to correct for less efficient depletion in certain cell lines. All of the screened cell lines show no effect on negative selection examining the Neg. Contr. sgRNA. Noteworthy, targeting *SMARCA2*, the paralog gene of *SMARCA4,* did not score in the domain-based screen. In contrast, pronounced vulnerabilities, associated with an α-RSA scores <-0.5 are detected in a panel of ESCC cell lines (Figure 23). Whereas the positive controls show strong effects in all cell lines used for the screen, the SMARCA4 dependency is selective to individual cell lines. Six ESCC cell models are identified with a strong dependency on SMARCA4 by α-RRA scores <-0.5 including KYSE-270 (α-RRA=-2.04), T.T (α-RRA=-1.88), KYSE-30 (α-RRA=-1.52), KYSE-410 (α-RRA=-1.24), KYSE-510 (α-RRA=-1.15), and COLO-680N (α-RRA=-0.62) (Figure 23). Cell lines with an α-RRA score of -0.50 or above are regarded as SMARCA4-independent such as KYSE-150 (α-RRA= -0.50), KYSE-70 (α-RRA=-0.44), KYSE-140 (α-RRA= -0.38), and KYSE-450 (α-RRA= -0.16) (Figure 23).

3.3 SMARCA4 dependency anti-correlates with SMARCA2 expression

To further strengthen the findings obtained by CRISPR-Cas9 domain-based screens, the data was compared to published data sets from McDonald III *et al*. and Meyers *et al.* [80, 81]. Both screens have been designed to identify genome wide cancer related dependencies. RSA scores were retrieved from shRNA screens [81] and Ceres scores from CRISPR screens [80]. Both include a high number of cell lines and were therefore analyzed in this study with the aim to reveal a potential biomarker for SMARCA4-dependent cell lines, testing the SL reciprocal to the known interaction with SMARCA2 [238, 292, 293].

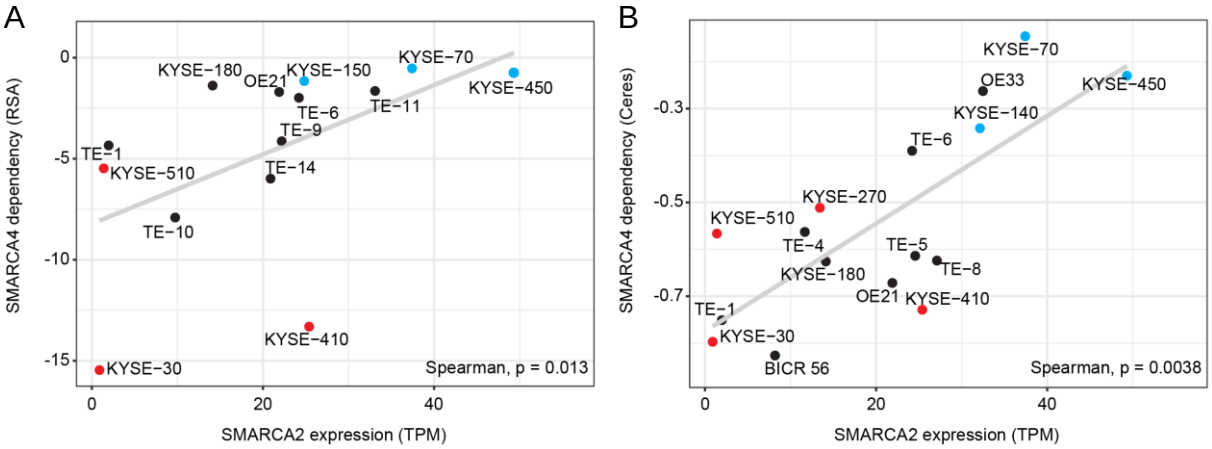


Figure 24: Spearman correlation of SMARCA4 dependency scores (RSA or Ceres) with SMARCA2 expression values (TPM).

Cell lines which were implemented in our CRISPR-Cas9 domain-based screens are colored: cell lines confirmed sensitive towards SMARCA4-KO are indicated in red, whereas identified SMARCA4 independent cell lines are depicted in blue. The black dots represent cells which were not included in our screens. The transcripts per million (TPMs) were obtained from the ordino platform (https://ordino.caleydoapp.org, [314]). A) RSA scores [81] plotted against SMARCA2 expression, B) Ceres scores [80] plotted against SMARCA2 expression. Plots were kindly provided by Thomas Zichner.

The obtained data from CRISPR-Cas9 domain-based screens is in consonance with the published data sets [80, 81]. Noteworthy, the term CRISPR-Cas9 domain-based screen is always referring to the screens conducted in this study (Figure 23). KYSE-30, KYSE-510 and KYSE-410 are among the most sensitive cell lines towards shRNA-mediated knock-down (KD) of SMARCA4 analyzing the RSA score (Figure 24A, [81]). In contrast, KYSE-70, KYSE-150, and KYSE-450 indicate high scores and are therefore considered as SMARCA4-independent (Figure 24A, [81]). In addition, the same cell lines KYSE-70 and KYSE-450 are confirmed as SMARCA4-

independent in the CRISPR screen (Figure 24B, [80]) whereas pronounced effects on SMARCA4-KO are confirmed in KYSE-30 and KYSE-410 (Figure 24B, [80]). The scores for SMARCA4 in KYSE-180 and OE21 are inconsistent among the two published data sets (Figure 24, [80, 81]). Both cell lines were not included in the CRISPR-Cas9 domain-based screen. Thus, no conclusion on SMARCA4 dependencies can be drawn. An additional cell line also implemented in the Ceres data set is represented by KYSE-270, revealing strong dependency on SMARCA4 (Figure 24B), which is in consonance with the domain-based screens. In addition, the published screens reveal additional SMARCA4-dependent (TE-4, TE-5, TE-8, TE-10, TE-14, BICR 56) as well as independent (TE-6, TE-11, OE33) cell lines (Figure 24, [80, 81]). Both published screens assayed more ESCC cell lines than used for the CRISPR-Cas9 domain-based screen and were therefore used for further statistical analysis (Figure 24).

After strong evidence of dependency on SMARCA4 in a selected panel of ESCC cell lines, the cause of leading to this vulnerability was assessed. To this end, the latest literature was taken into consideration describing paralog dependencies in the SWI/SNF complex in diverse cancer subtypes, especially direct paralog dependency between the two redundant ATPases SMARCA2 and SMARCA4 [238, 292, 293]. In non-small cell lung cancer, *SMARCA4* mutation and following loss of the respective protein renders cell lines dependent on the paralog partner SMARCA2 [238, 292, 293]. To investigate whether this paralog dependency is maintained in the ESCC cell lines tested, the RSA as well as the Ceres scores for SMARCA4 dependency from the public available data sets [80, 81] were plotted against the SMARCA2 expression values [TPM] obtained from *ordino* (https://ordino.caleydoapp.org). Indeed, a significant correlation is observed using spearman correlation test (RSA spearman=0.013; Ceres spearman=0.0038) suggesting that $SMARCA2^{low}$ sensitizes the respective cell lines toward SMARCA4 inhibition. The correlation is linear between SMARCA4 RSA or Ceres scores and SMARCA2 TPM values, indicating that SMARCA4 dependency anti-correlates with levels of SMARCA2 expression. An arbitrary cut-off (-5 for the RSA scores and -0.5 for the Ceres scores) is used, based on the results obtained from CRISPR-Cas9 domain-based screens (Figure 25 and Figure 26) in order to separate SMARCA4-dependent and independent models.

To test whether this correlation holds true in the cell lines included in the CRISPR-Cas9 domain-based screens (Figure 23), qPCR as well as capillary Western immunoassay were assessed in order to analyze the actual mRNA as well as protein levels of SMARCA2 and SMARCA4 (Figure 25 and Figure 26).

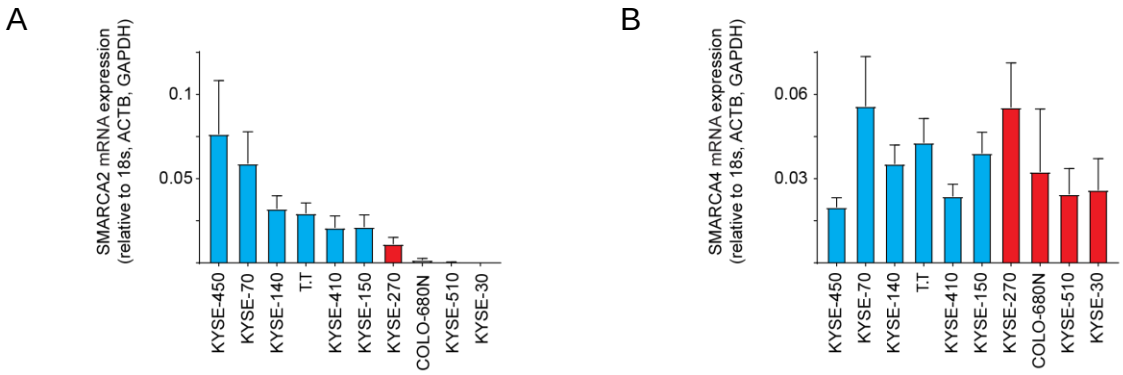


Figure 25: qRT-PCR data of SMARCA2 and SMARCA4 expression in a panel of ESCC cell lines.

Ten ESCC cell lines used for the CRISPR-Cas9 domain-based screens were analyzed according to the SMARCA2 and SMARCA4 expression. Values obtained from qPCR are normalized to three house-keeping genes 18S, ACTB, and GAPDH. Cell lines are ranked from high to low SMARCA2 mRNA expression. A: SMARCA2 expression; B: SMARCA4 expression. Characterized SMARCA4-independent cells are indicated in blue, SMARCA4-dependent cell lines are highlighted in red. qRT-PCR data is represented as mean ± SD of three independent experiments. Experiment was conducted by Ursula Strobl.

Indeed, the analysis of SMARCA2 and SMARCA4 mRNA expression confirmed the correlation with originally obtained values extracted from *ordino* [314]. Whereas SMARCA2 mRNA expression is slightly descending from cell line to cell line, mRNA of SMARCA4 is detected in all cell lines to sufficient amounts (Figure 25). In contrast, KYSE-30, KYSE-510, COLO-680N indicate low or absent values for SMARCA2 revealing a complete loss of SMARCA2 expression. KYSE-270 shows low SMARCA2 expression values suggesting that low levels might also be insufficient to compensate for the loss of the paralog partner. Although no clear cut-off between SMARCA2^{low} and SMARCA2^{high} cell lines can be introduced, COLO-680N, KYSE-510, and KYSE-30 are confirmed as SMARCA2^{low} cell lines by showing a complete loss of SMARCA2 mRNA expression (Figure 25).

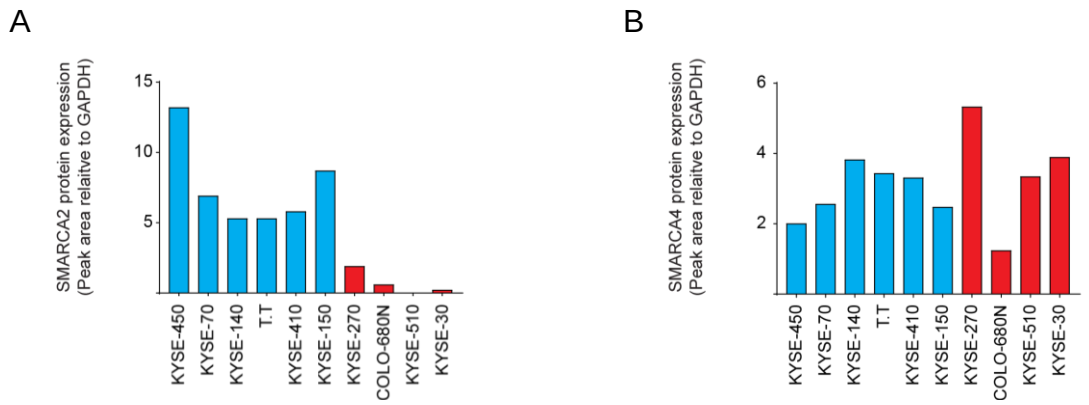


Figure 26: SMARCA2 and SMARCA4 protein analysis in a panel of ESCC cell lines using capillary Western immunoassay.

A) Analysis of SMARCA2 protein expression levels relative to GAPDH; B) Analysis of SMARCA4 protein expression values relative to GAPDH. Characterized SMARCA4-independent cells are indicated in blue, SMARCA4-dependent cell lines are shown in red.

Similarly to the expression values, the protein levels of SMARCA2 and SMARCA4 in the same panel of ESCC cell lines were determined. The protein levels of SMARCA2 are in accordance with the qPCR data, whereas a better separation between SMARCA2^{low} and SMARCA2^{high} cell lines is achieved (Figure 26). A pronounced difference between the cell lines highlighted in blue (SMARCA2^{high}) and red (SMARCA2^{low}) is observed (Figure 26A). Again, SMARCA4 protein levels are detectable in all of the analyzed ESCC cell lines (Figure 26B). However, KYSE-270 show high levels of SMARCA4 whereas in COLO-680N low levels of SMARCA4 are detected.

3.4 CRISPR-scan reveals most efficient sgRNAs and domain relevance in two different ESCC cell lines

Remarkably, only sgRNAs targeting the BD of SMARCA2 and SMARCA4 were included in the library, while the dependency has been linked to the helicase function of SMARCA2 or SMARCA4 in NSCLC [300, 313]. To test the differential requirement on ATPase versus BD function, sgRNAs targeting various parts of the gene (CRISPR-scan) including the N-terminal region, the ATPase encoding domains: DEXDc (DEAD-like helicase superfamily) and helicaseC as well as the BD, were designed (Figure 27 and Figure 28). By conducting sgRNA depletion experiments as outlined in Figure 13, different strong dependencies on the respective domains are identified. In line with the fact that a higher frequency of null mutations is observed with sgRNAs designed against sequences encoding functionally relevant protein domains [79], only minor depletion effects directing the sgRNAs against the N-terminal region of SMARCA4 are detected in two SMARCA4-dependent cell lines KYSE-30 and KYSE-510 (Figure 27 and Figure 28). In KYSE-30 cell lines, strong effects on depletion upon sgRNA targeting the BD are measured but similar or even higher fold-changes are obtained by designing the sgRNAs against the ATPase domain (Figure 27). In contrast, in KYSE-510, sgRNAs binding the ATPase-encoding sequence show a much higher fold-depletion and more pronounced effects than the sgRNAs directed against the BD (Figure 28). In line with the published data for the SMARCA4 domain-function [300, 313], the results of this study also suggest the ATPase activity as the functional requirement in the context of SMARCA4 dependency in ESCC.

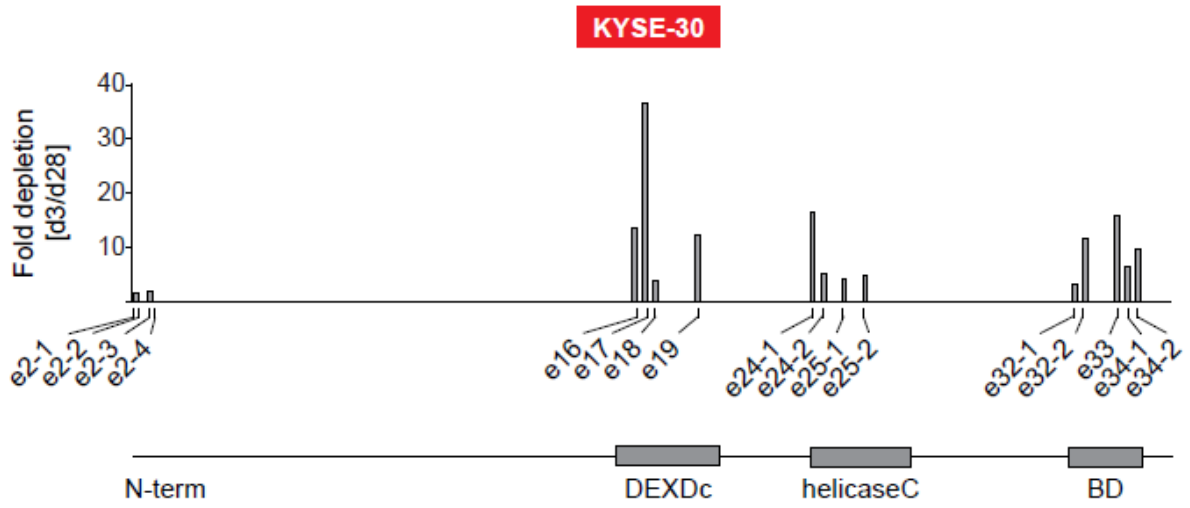


Figure 27: Domain scan by application of sgRNAs targeting different domains within SMARCA4 in ESCC cell line KYSE-30.

Fold depletion from d28 compared to d3 of different sgRNAs targeting various domains of the SMARCA4 gene. At least four sgRNAs were designed against the N-term, DEXDc (DEAD-like helicase superfamily), helicaseC, and BD (bromodomain). Respective domains are depicted on the x-axis whereas the fold depletion is shown on y-axis.

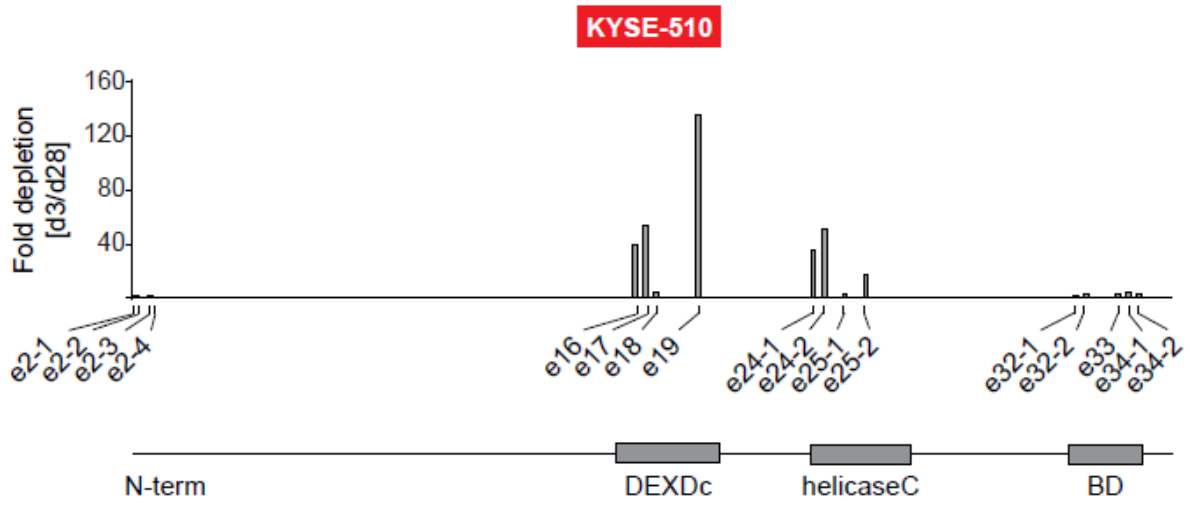


Figure 28: Domain scan by application of sgRNAs targeting different domains within SMARCA4 in ESCC cell line KYSE-510.

Fold depletion from d28 compared to d3 of different sgRNAs targeting various domains of SMARCA4 gene. At least four sgRNAs were designed against the N-term, DEXDc (DEAD-like helicase superfamily), helicaseC, and BD (bromodomain). Respective domains are depicted on the x-axis whereas the fold depletion is shown on y-axis.

3.5 Dependencies from screen confirmed by singleton gRNA depletion experiments

The CRISPR-scan allowed the identification of three efficacious sgRNAs targeting the DEXDc domain (DEXDc e16, e17, e19) and BD (BD_e33, e34-1, e34-2). To validate the findings of the pooled sgRNA screen, selected sgRNAs were used for depletion time-resolved experiments, outlined in Figure 13, in four of the SMARCA2-proficient and deficient ESCC lines, respectively (Figure 29 and Figure 30). The values are depicted in fold-changes revealing a robust distinction between SMARCA4-dependent and independent cell lines (Figure 29). To summarize the results, fold-changes were further normalized to POLR2A (Figure 30). SMARCA2-proficient as well as deficient cell lines are confirmed using capillary Western immunoassay (Figure 31). In the SMARCA2-proficient cell models composed of KYSE-450, KYSE-140, KYSE-70 and KYSE-150, no effects on depletion are detected upon KO of SMARCA4 over a time course of 28 days (Figure 29 and Figure 30). In contrast, knocking out SMARCA4 in the SMARCA2-deficient cell lines KYSE-270, KYSE-30, KYSE-510 and COLO-680N (Figure 31) results in deleterious effects similar or stronger than the effects observed upon application of the positive control sgRNA targeting *POLR2A* (Figure 29 and Figure 30). Previous studies have shown more pronounced effects on depletion via targeting the DEXDc domain compared to BD, indicating a requirement of SMARCA4 ATPase activity [79]. This finding is in consonance with the results obtained via the CRISPR-scan of SMARCA4 in ESCC, highlighting strongest fold-depletions when targeting the DEXDc domain. Targeting the BD results in deleterious effects in KYSE-30 only (Figure 29 and Figure 30). Of note, DEXDc-targeting sgRNAs were not included in the pooled sgRNA library. Therefore, the screening methodology is suggested to be more sensitive, allowing the detection of weak dependencies obtained by targeting *SMARCA4* with BD-directed sgRNAs (Figure 23). In contrast, singleton depletion experiments show a strong differential dependency on the ATP-binding domain directed sgRNAs, which have not been included in the screen compared to the BD targeting sgRNAs (Figure 23). Taken together these results highlight the SWI/SNF subunit SMARCA4 as a potential novel target in ESCC characterized by low or absent expression of its homologous gene *SMARCA2*. Furthermore, the ATP-binding domain of SMARCA4 is proposed to be the functional indispensable domain for cell survival in ESCC.

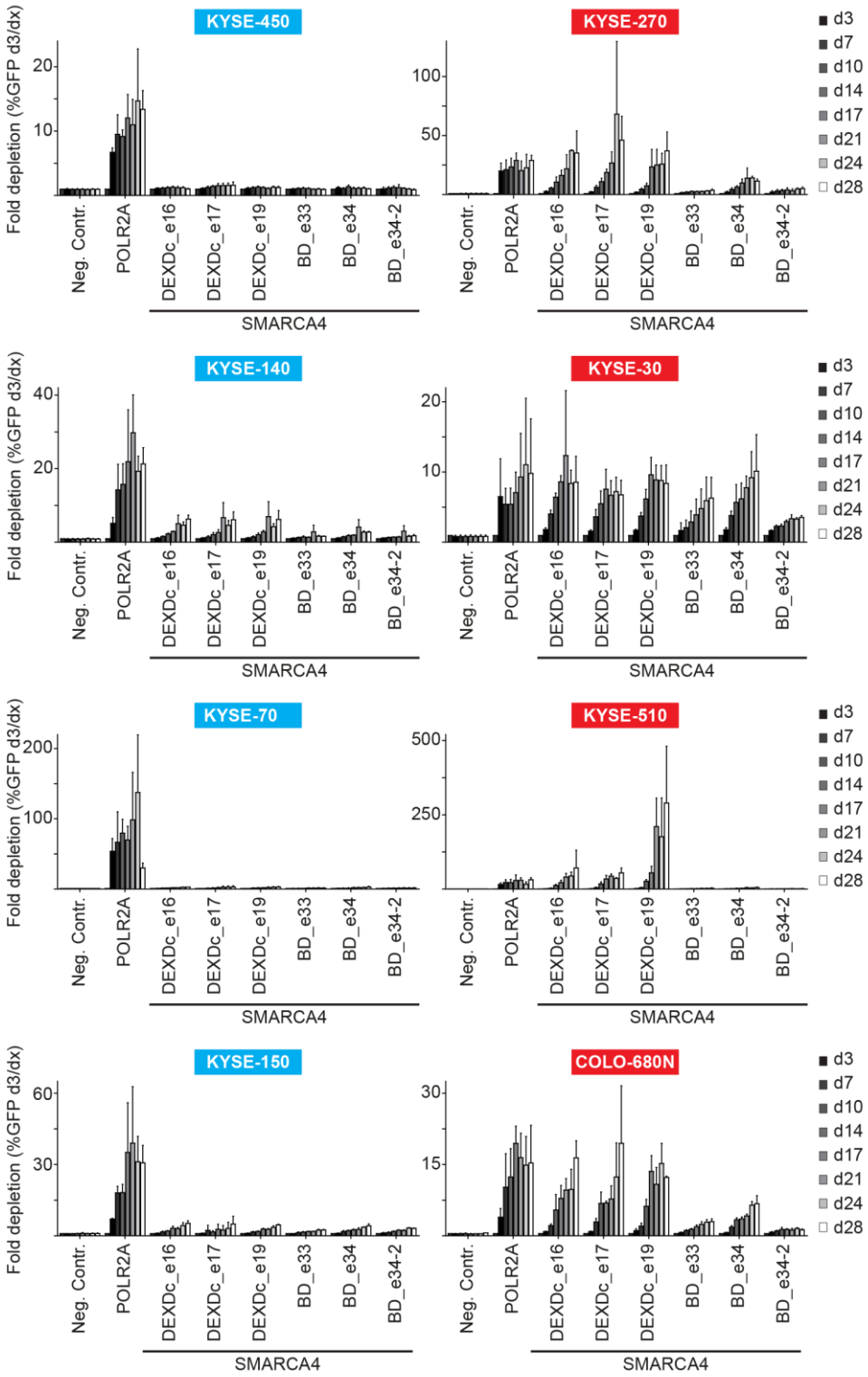


Figure 29: Validation of SMARCA4 dependency using individual sgRNAs targeting the ATPase-domain (DEXDc; DEAD-like helicase superfamily) - or bromodomain (BD) in SMARCA2-deficient (red) and proficient (blue) cell lines.

Left panel (blue) represents the SMARCA4-independent (SMARCA2^{high}) , right panel (red) shows SMARCA4-dependent (SMARCA2^{low}) cell lines. Every group of cells consists of four individual cell lines. The bars represent different time points of GFP^{+} population measurement via flow cytometry analysis starting from d3 to d28. The controls are comprised by one negative control (Neg. Contr.) and one positive control targeting sgRNA. Data is plotted from three individual experiments using the mean ± SD.

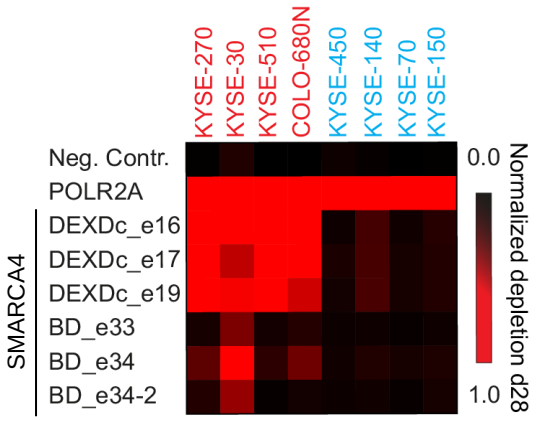


Figure 30: Heatmap of SMARCA4 dependency related to POLR2A.

Fold depletion values at d28 were normalized to POLR2A, whereas 1 (=cut-off) is the maximal depletion obtained by the positive control. SMARCA4 sgRNAs show values which are at least as strong as the positive control. The negative control (Neg Contr.) value is 0. sgRNAs are indicated on the left side, the cell lines are shown on top, summarizing the data from sgRNA depletion experiments (Figure 29). Data is plotted from three individual experiments using the mean which is normalized to the mean of POLR2A. ATPase-domain DEXDc: DEAD-like helicase superfamily; BD: bromodomain.

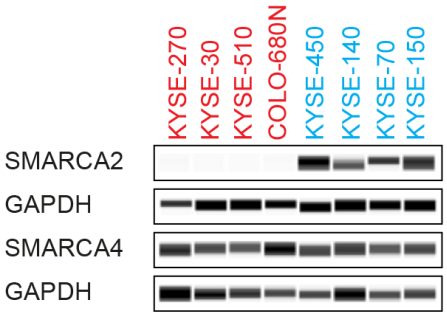


Figure 31: SMARCA2 and SMARCA4 protein level in eight selected ESCC cell lines.

Summary of capillary Western immunoassay data showing SMARCA2 and SMARCA4 protein levels in eight selected ESCC cell lines. Previously identified SMARCA4-dependent cell lines are indicated in red, SMARCA4-independent cell lines are shown in blue. Data was retrieved by using a multiplexing approach, detecting the house-keeping protein with the same capillary as the protein of interest.

3.6 SMARCA4 dependency is linked to its helicase function

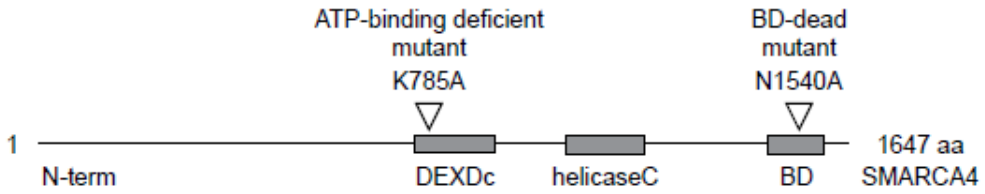


Figure 32: Schematic representation of SMARCA4 domains.

Inactivating mutations in the ATP-binding domain and the bromodomain (BD) are indicated by triangles. The sequence is codon optimized to render it resistant (SMARCA4^{res}) to siRNAs and sgRNAs. Domains are annotated according to UniProtKB entry P51532 (DEXDc+helicaseC) + NCBI entry cd05516 (BD). ATPase-domains: DEXDc (DEAD-like helicases superfamily), helicaseC; BD (bromodomain).

To robustly investigate the requirement of the ATP-binding deficient and the BD of SMARCA4 in SMARCA2-deficient ESCC cells, rescue experiments were conducted by ectopically expressing SMARCA4 siRNA/sgRNA resistant (SMARCA4^{res}) variants: SMARCA4 wild-type (wt), ATP-binding deficient (K785A), and BD-dead (N1540A), respectively (Figure 32) [300, 313]. All of the generated constructs were codon optimized. Therefore, only the triple base pair sequence is changed whereas the encoded amino acid remains the same. The sequence code is modified making the recombinant genetic material inaccessible for siRNAs or sgRNAs. Wild-type and mutant forms of SMARCA4^{res} were stably transduced in KYSE-510 and KYSE-30. The KD of SMARCA4 shows a decrease in protein levels of SMARCA4 after the application of two independent siRNAs (SMARCA4_1 and SMARCA4_2) on capillary Western immunoassay analysis for both KYSE-510 and KYSE-30 parental cell lines compared to non-targeting control (NTC) (Figure 33). In both cell lines, KYSE-510 and KYSE-30, the siRNA SMARCA4_2 leads to a more pronounced effect than siRNA SMARCA4_1 (Figure 33). Due to the fact that endogenous SMARCA4 is still targeted, a small drop in protein levels is detected upon treatment with siRNAs in ectopically expressing SMARCA4 cell lines (Figure 33) but efficient protein rescue is obtained. In summary, sufficient expression as well as resistance towards siRNA-mediated KD of SMARCA4^{res} variants was confirmed in all of the engineered cell

lines including KYSE-510 (wt, ATP-binding deficient, BD-dead) and KYSE-30 (wt, ATP-binding deficient, BD-dead).

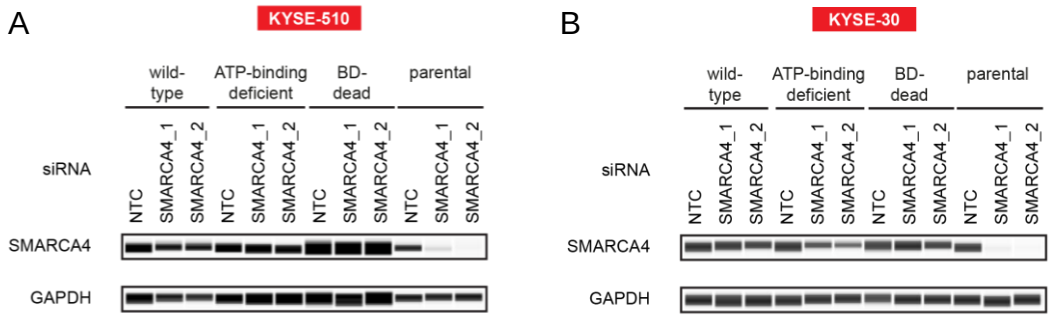


Figure 33: siRNA-mediated KD of SMARCA4 in cell lines ectopically expressing codon-optimized SMARCA4.

KYSE-510 and KYSE-30 cells were stably transduced with empty vector as a control and SMARCA4^{res} (SMARCA4 resistant to siRNA and sgRNA) variants: wild-type, ATP-binding deficient (K785A), or BD-dead (BD, bromodomain) (N1540A). Expression of the construct is confirmed by resistance towards siRNA-mediated KD of SMARCA4 using two different siRNAs (SMARCA4_1, SMARCA4_2) compared to non-targeting control (NTC) assessed by capillary Western immunoassay. Lysates were prepared 72h post-transfection and GAPDH was used as a loading control. Data is depicted from single siRNA-mediated KD experiment.

To test the generated cell lines according to their capability of rescuing effects obtained by KO of SMARCA4, singleton depletion assays were conducted. In line with the previous results (Figure 23), strong negative selection is observed assessing SMARCA4 sgRNAs in the empty vector control-transduced KYSE-510 cell line (Figure 34). In contrast, in wild-type SMARCA4^{res}-expressing KYSE-510, phenotypic effects caused by loss of endogenous SMARCA4, is reversed. Similar effects were observed upon ectopic expression of BD-mutant SMARCA4^{res} in KYSE-510 (Figure 34). However, the ATP-binding deficient variant of SMARCA4^{res} is not able to rescue the effects on depletion upon endogenous targeting of *SMARCA4* (Figure 34). Comparable results were obtained in KYSE-30 cell line (Figure 35). Ectopic expression of the wild-type or the BD-mutant SMARCA4^{res} but not the ATP-binding deficient variant rescues proliferation upon loss of endogenous SMARCA4 (Figure 35). In summary, on-activity of the sgRNAs targeting *SMARCA4* is confirmed. Furthermore, the results from the depletion screen, the CRISPR-scan and the individual sgRNA depletion experiments proof SMARCA4 as a selective dependency

in ESCC cell lines and highlight the ATP-binding domain (DEXDc) as the functional indispensable region of SMARCA4.

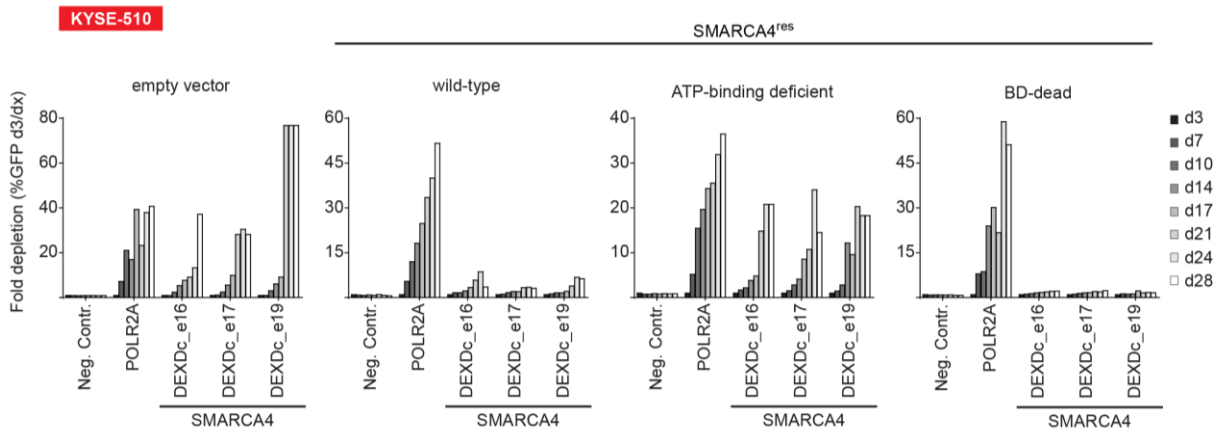


Figure 34: Rescue experiments using different SMARCA4 constructs in KYSE-510.

SMARCA4 variants (wild-type, ATP-binding deficient, and BD-dead) expressing cell lines were transduced with Cas9 for sgRNA depletion studies. Five sgRNAs were tested including Neg.Contr., POLR2A, DEXDc (e16, e17, and e19). Fold depletion is depicted on the y-axis. The observation time period was 28 days. Every bar represents the fold depletion of a respective time point relative to d3.

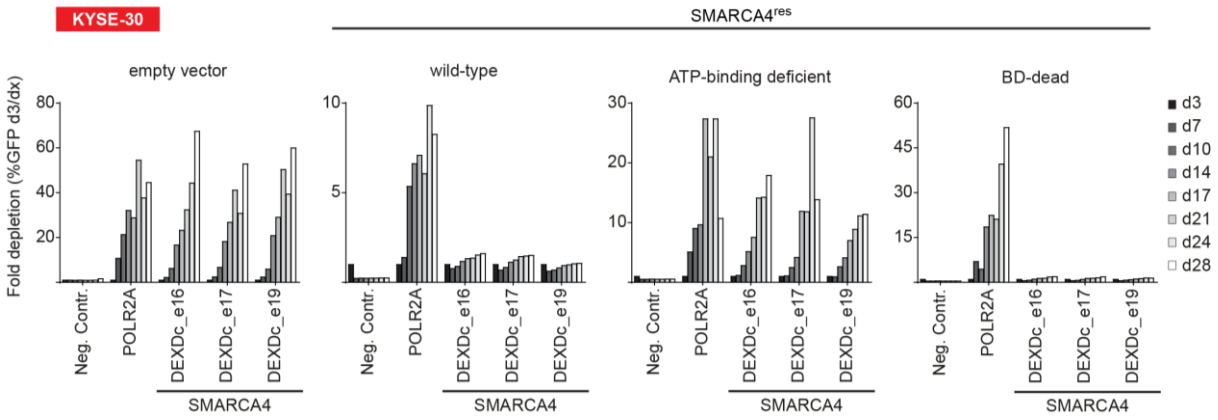


Figure 35: Rescue experiments using different SMARCA4 constructs in KYSE-30.

SMARCA4 variants (wild-type, ATP-binding deficient, and BD-dead) expressing cell lines were transduced with Cas9 for sgRNA depletion studies. Five sgRNAs were tested including Neg. Contr., POLR2A, DEXDc (e16, e17, and e19). Fold depletion is depicted on the y-axis and time period was 28 days. Every bar represents fold depletion of a respective time point to d3.

3.7 Synthetic lethal dependencies are confirmed for SMARCA2 and SMARCA4

To further evaluate the synthetic lethal (SL) interaction between SMARCA4 and SMARCA2, two reciprocal approaches were conducted. Firstly, SMARCA2 re-expression (=SMARCA2^{ect}) in SMARCA2^{low} cell lines was assessed, testing whether insensitivity towards SMARCA4-KO can be achieved. Secondly, the induction of SMARCA4 dependency was analyzed in SMARCA2-KO clones, which were derived from a SMARCA2^{high} cell line.

3.7.1 Re-expression of SMARCA2 reverses SMARCA4 dependency

For analysis of ectopic re-expression of SMARCA2^{ect}, different variants were assessed including wild-type, ATP-binding deficient (K755A) and BD-dead (N1482A) SMARCA2^{ect} (Figure 36). KYSE-510 cell line was used, showing a strong effect on depletion upon SMARCA4-KO in previous experiments (Figure 29). The expression of the variants of SMARCA2^{ect} was tested using capillary Western immunoassay. Strong expression of SMARCA2^{ect} of all the different variants including wild-type, ATP-binding deficient and BD-dead was observed compared to the parental control (Figure 37).

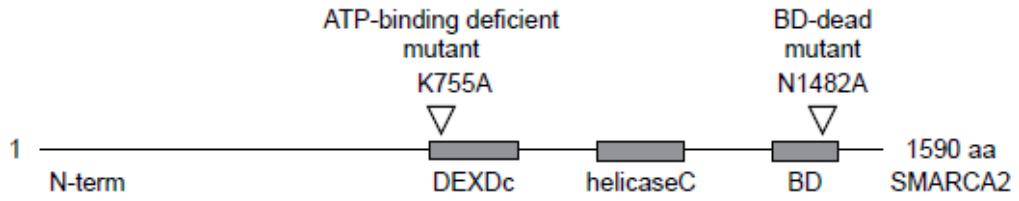


Figure 36: Schematic representation of SMARCA2 domains.

Different variants of SMARCA2 were ectopically expressed (SMARCA2^{ect}). Introduced inactivating mutations in the ATP-binding domain and the BD (bromodomain) are indicated by triangles. Domains are annotated according to UniProtKB entry P51531 (DEXDc+helicaseC) + NCBI entry cd05516 (BD). ATPbinding-domains: DEXDc (DEAD-like helicases superfamily), helicaseC; BD (bromodomain).

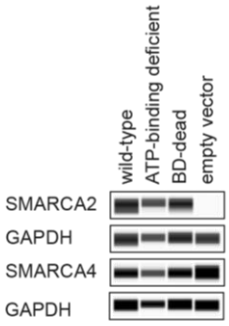


Figure 37: Detection of re-expressed SMARCA2 variants using capillary Western immunoassay.

SMARCA2^{ect} variants were re-expressed in the SMARCA2^{low} and SMARCA4-dependent cell line KYSE-510. The variants were comprised by wild-type, ATP-binding mutated as well as BD-mutated SMARCA2^{ect}. Empty vector was used as a control. SMARCA4 detection also served as a control.

Additionally, effects upon re-expression of the different variants of SMARCA2^{ect} on SMARCA4 dependency were interrogated. The engineered cell lines were tested on depletion efficacy upon KO of SMARCA4 by using three sgRNAs targeting the DEXDc domain. Whereas expression of SMARCA2^{ect} wild-type as well as BD-dead variants reverse the dependency on SMARCA4 after the application of three gRNAs targeting the DEXDc domain, no rescue is obtained upon SMARCA2 ATP-binding deficient variant re-expression in KYSE-510 (Figure 38).

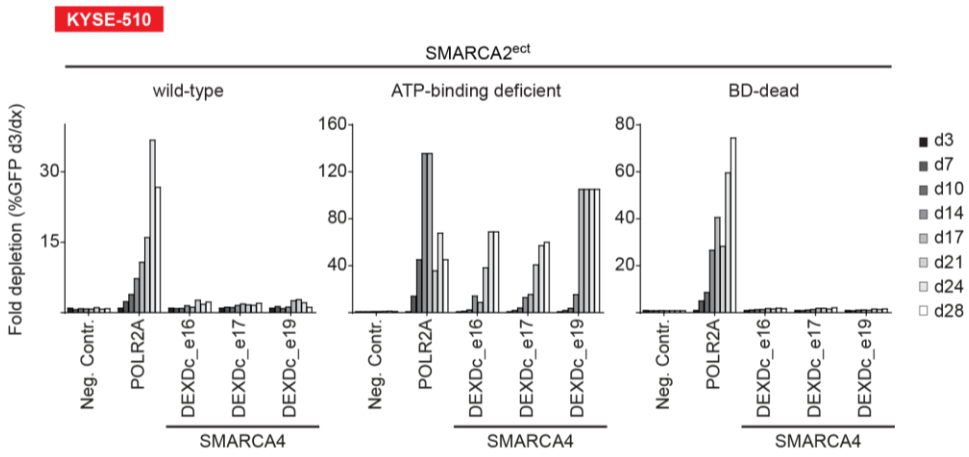


Figure 38: Singleton depletion experiments confirms SL interactions between SMARCA2 and SMARCA4 in KYSE-510.

Singleton depletion assay in SMARCA2^{ect} variant (wild-type, ATP-binding deficient, BD-dead) expressing KYSE-510 cell lines. Fold depletion shown as %GFP at d3 compared to dx (x= time points). Every single bar represents a different time point. Data is depicted from a single CRISPR experiment.

3.7.2 SMARCA2-KO induces SMARCA4 dependency

Using a reciprocal approach, SMARCA2-KO in two SMARCA2-proficient cell lines, KYSE-450 and KYSE-150 was performed. In order to derive SMARCA2-KO monoclonal cells from the parental cell line, transient transfection of an all-in-one-construct (see section 2.14) encoding the sgRNA, Cas9 as well as GFP was conducted. After efficient transfection, the cells were sorted using fluorescence activated cell sorting (FACS) and single cell clones were generated by limited dilutions ensuring that only one cell per well was seeded. In contrast to the sgRNAs used in the screen, sgRNAs targeting the N-terminal region of SMARCA2 were designed for the generation of a monoclonal cell line, minimizing the risk that a short version of the protein might be able to maintain the cellular function.

KYSE-450 was identified as the most insensitive cell line towards SMARCA4-KO via our CRISPR-Cas9 domain-based screens (Figure 23), public available screens [80, 81] as well as via singleton depletion experiments (Figure 29).

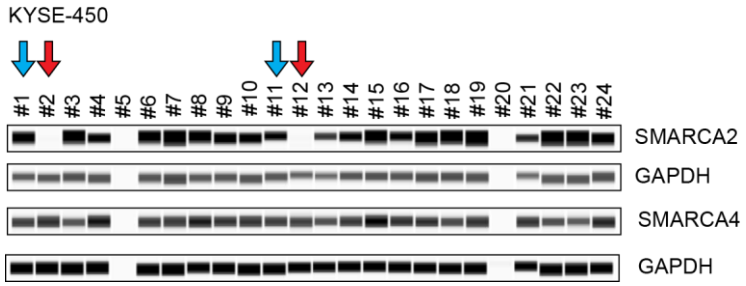


Figure 39: KO confirmation of SMARCA2 in KYSE-450 using capillary Western immunoassay.

Generated clones were analyzed according to their SMARCA2 protein levels. Clones with a complete deletion of SMARCA2 are indicated by red arrows; non-editing-clones which still express SMARCA2 are highlighted by blue arrows and are used as control cell lines in follow-up studies. SMARCA4 detection served as a control.

To further test the hypothesis that SMARCA4 dependency is caused by loss of its paralog SMARCA2 and thereby initiating a SL dependency, SMARCA2-null KYSE-450 monoclonal lines were generated using CRISPR-Cas9-mediated gene KO. Protein levels of the retrieved monoclonal KYSE-450-SMARCA2-KO cell lines were analyzed using capillary Western immunoassay (Figure 39). Out of 24 clones, only

two indicate a sufficient SMARCA2-KO, including clone#2 and clone#12 (highlighted by red arrows) (Figure 39). For further analysis, successfully edited clones (clone#2 and clone#12) as well as insufficiently edited clones (clone#1 and clone#11) (indicated by blue arrows), serving as controls, were selected (Figure 39). Low amounts of SMARCA2 protein levels are detected in clone#4, clone#9, clone#10, clone#13 and clone#21 indicating a possible editing of only one allele which might still be capable of expressing SMARCA2 (Figure 39).

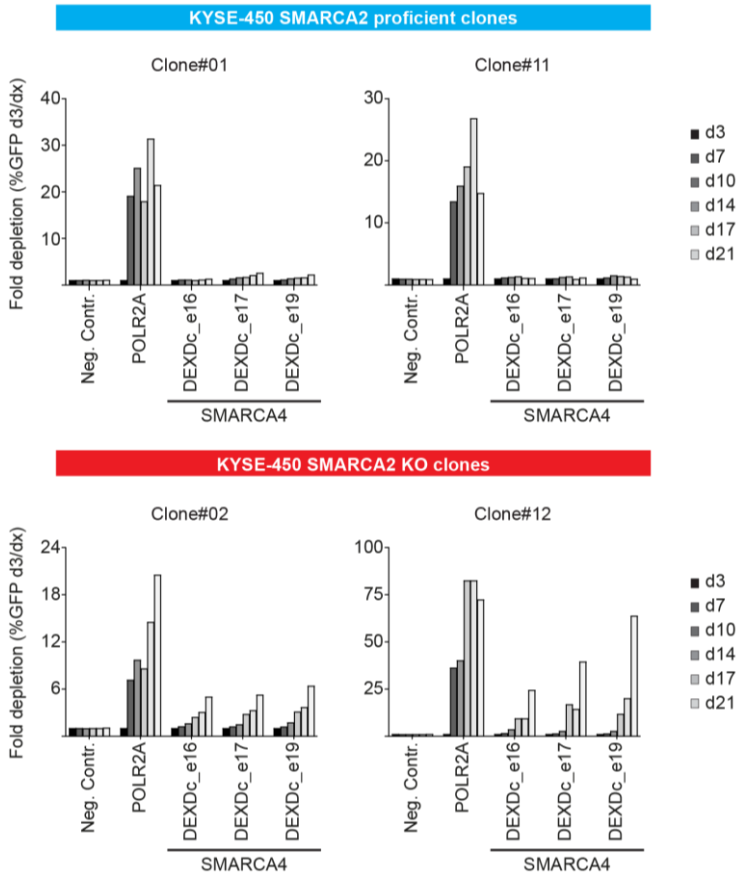


Figure 40: Depletion experiments using sgRNAs targeting SMARCA4 in SMARCA2-KO and proficient KYSE-450 clones.

Induction of SMARCA4 dependency via knock-out (KO) of SMARCA2 in KYSE-450 cell line. Depletion experiments are depicted in SMARCA2-proficienct (blue) and SMARCA2-KO (red) monoclonal cell lines. Fold depletion is shown as %GFP at d3 divided by %GFP dx (x= time points). Every bar represents a different time point. Data is depicted from a single CRISPR experiment.

To investigate the effect upon SMARCA4-KO on a monoclonal SMARCA2-KO cell line, selected clones were treated with sgRNAs in depletion experiments. Rather than comparing the effects to the parental cell line, also clones were picked for which the application of a sgRNA targeting the N-terminal region of SMARCA2 had no editing event (clone #1 and clone#11) (Figure 39). Similarly to the parental cell line, Neg. Contr. sgRNA, one positive control sgRNA targeting *POLR2A* as well as three sgRNAs targeting the DEXDc domain of SMARCA4 were used. The application of the sgRNAs against the DEXDc domain in the two SMARCA2-KO clones results in a significant increase in fold-depletion compared to the two SMARCA2-proficient clones (Figure 40). The effects on the two selected SMARCA2-proficient clones are comparable to those of the parental cell line used in the hit validation experiments (Figure 16).

The hypothesis that removing SMARCA2 from a SMARCA2-proficient cell line renders it dependent on SMARCA4-KO was further confirmed in KYSE-150 cell line. KYSE-150 was also proven to be SMARCA4-independent (Figure 29) and was therefore selected for single-cell SMARCA2-KO generation.

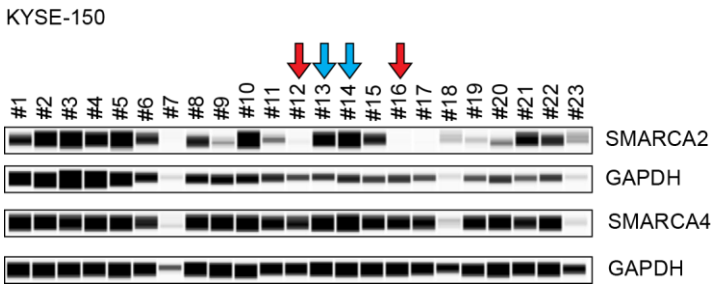


Figure 41: KO confirmation of SMARCA2 in KYSE-150 using capillary Western immunoassay.

Different clones are analyzed according to their SMARCA2 protein levels. Clones with a complete deletion of SMARCA2 are indicated by red arrows; non-edited, SMARCA2 expressing clones indicated by blue arrows are used as control cell lines in the follow-up studies. SMARCA4 detection served as a control.

From KYSE-150 cell line, 23 clones were derived and analyzed according to the SMARCA2 protein levels after transient transfection with all-in-one plasmid (sgRNA-Cas9-GFP; section 2.14), cell sorting and single cell cloning (see section, 2.14). The output was comparably high to the KYSE-450 SMARCA2-KO approach. For further analysis, clone#12 and clone#16 indicated by red arrows (SMARCA2-KO clones) as well as clone#13 and clone#14 (SMARCA2-insufficient KO clones) were selected (Figure 41).

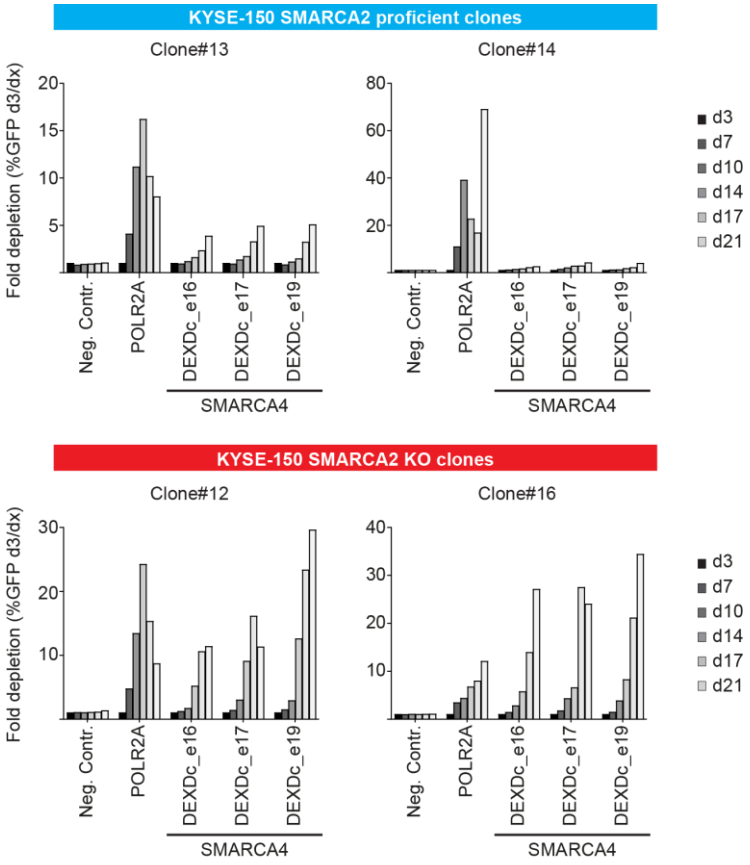


Figure 42: SL concept confirmed in KYSE-150 SMARCA2-KO cell lines.

Induction of SMARCA4 dependency via KO of SMARCA2 in KYSE-150 cell line. Depletion experiments in SMARCA2-proficient (blue) and SMARCA2-KO (red) monoclonal cell line. Fold depletion is shown as %GFP at d3 compared to %GFP at dx (x= time points). Every bar represents a single time point. Data is depicted from a single CRISPR experiment.

In line with the results obtained using KYSE-450 cell line (Figure 40), SMARCA2-KO in KYSE-150 induces SMARCA4 dependency in selected clones#12 and 16 (Figure 42). In contrast, SMARCA2-proficient monoclonal cell lines remain SMARCA4-independent (Figure 42). A small effect on fold-depletion is detected in clone#13 upon targeting SMARCA4 ATPase-domain, which might be linked to an unknown editing event affecting SMARCA2 function, which is not detectable via capillary Western immunoassay.

In summary, the hard-wired known SL interactions between SMARCA2 and SMARCA4 [238, 292, 293] was confirmed in ESCC. SMARCA2 re-introduction in a SMARCA2-deficient cell line reversed the dependency on SMARCA4, whereas SMARCA4-independent cell lines could be transformed in dependent ones by knocking out SMARCA2.

3.8 SMARCA2 might serve as a patient selection biomarker

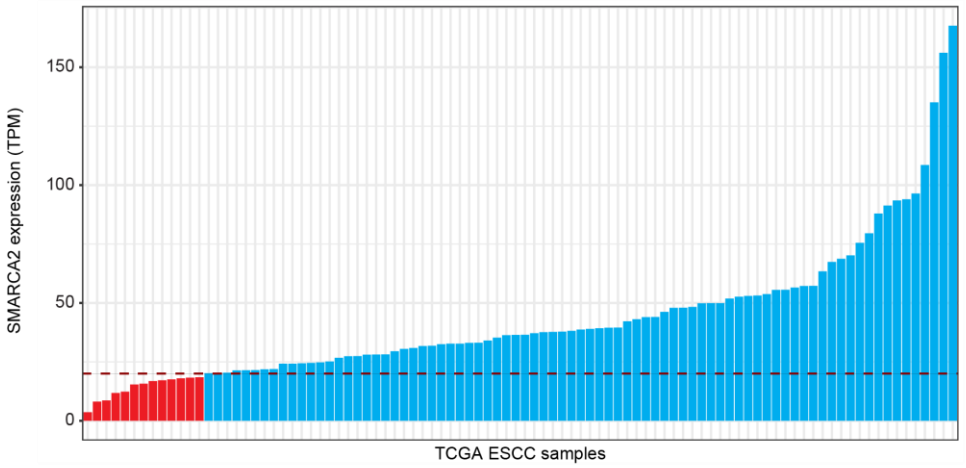


Figure 43: TCGA data of ESCC patient samples showing SMARCA2 expression (TPM).

*TCGA data was obtained from 94 ESCC samples. 13/94 (13.8%) samples show low SMARCA2 expression using an arbitrary cut-off of SMARCA2 TPM<20 (*https://ordino.caleydoapp.org*; [314]). The plot was kindly provided by Thomas Zichner.*

In order to get an estimate number of the actual SMARCA2 expression values in patient samples, the TCGA data was assessed for ESCC (https://ordino.caleydoapp.org, [314]). Similar to cancer cell lines, no straight cut-off is detected between SMARCA2 low and high patient samples (Figure 43). The same arbitrary cut-off, as previously introduced for the cell lines was implemented (Figure 25). Thereby, SMARCA2^{low} was defined by TPM<20 revealing 13.8% of ESCC patient samples as SMARCA2-deficient (Figure 43), indicating a remarkable patient population for drug discovery programs. In order to determine the real cut-off, more experiments and analyses have to be conducted.

3.9 Extension of the therapeutic concept targeting SMARCA4 in SMARCA2^{low} cell to diverse tumor types

The analysis of the public available data sets [81] reveals additional SMARCA4 dependencies in other indications apart from ESCC. When analyzing the RSA scores from all indications which have been used in the study, significantly low values were obtained for a subset of cell lines. We extracted the highest scores and – similarly to our analysis in ESCC – assessed the SMARCA2 expression values. Therefore, respective expression values were extracted from the ordino database (https://ordino.caleydoapp.org, [314]). Selected cell lines were analyzed according to their SMARCA2 protein levels and the dependency on SMARCA4 was tested via depletion experiments, again using sgRNAs against the ATP-binding domain (DEXDc) as well as the BD.

Cell line	Indication	RSA score	TPM SM2	TPM SM4
HuP-T4	Pancreas carcinoma	-9.531	1.02	62.8
SK-CO-1	Colon carcinoma	-17.278	1.24	102.5
OV-90	Ovarian carcinoma	-7.634	5.36	51.7
HCT 116	Colon carcinoma	-5.239	6.20	77.5

Table 4: SMARCA4 dependency scores (RSA) [81] and expression values for SMARCA2 and SMARCA4 (TPM) [314].

SM2= SMARCA2, SM4= SMARCA4, TPM= transcripts per million. The data was extracted from https://ordino.caleydoapp.org [314] and ranked according to SMARCA2 expression values.

Very strong SMARCA4 dependencies are observed in HuP-T4 (pancreas carcinoma), OV-90 (ovarian carcinoma), HCT 116 (colon carcinoma) and SK-CO-1 (colon carcinoma) regarding the RSA scores [81]. Indicated cell lines show low mRNA expression levels of SMARCA2 transcripts are observed. In contrast, SMARCA4 is expressed at high levels in all of the four cell lines analyzed (Table 4).

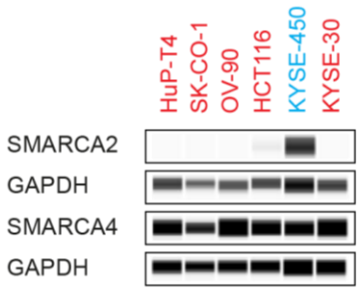


Figure 44: Capillary Western immunoassay of ESCC and additional cell lines.

Capillary Western immunoassay depicting SMARCA2, SMARCA4 protein levels and GAPDH loading control. Lysates were obtained from frozen cell pellets. Red: SMARCA2^{low} cell lines; blue: SMARCA2^{high} cell line serving as control. Experiment was performed by Teresa Puchner.

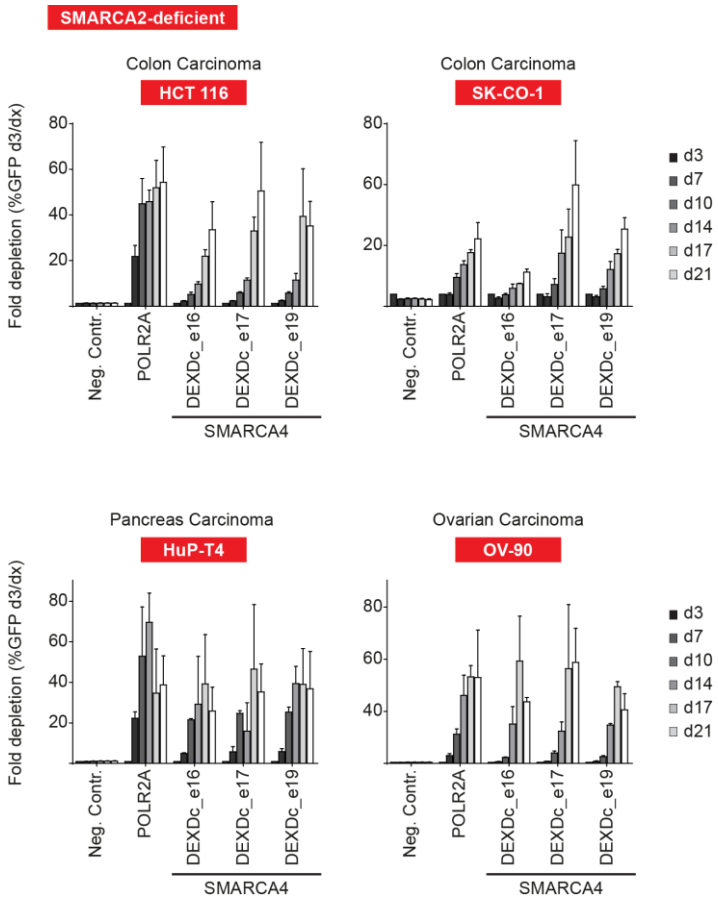


Figure 45: SMARCA4 dependency confirmation in additional indications.

Fold depletion were obtained by dividing the percentage of GFP$^+$ cells initially obtained by lentiviral transduction by the percentage of GFP$^+$ cells at respective timepoints. The bars represent the different measurements over time indicated by dark or light grey color. Negative control (Neg.Contr.), positive control (POLR2A) as well as three sgRNAs targeting the DEXDc (DEAD-like helicase superfamily) – domain were applied in single wells. The different targeted exons are indicated with an "e". Values were obtained by three independent experiments, whereby the mean ± SD is depicted. Experiments were conducted with the help of Teresa Puchner and Silvia Blaha-Ostermann.

To confirm a correlation of low mRNA with low protein expression, SMARCA2 protein levels in HuP-T4, OV-90, HCT 116, and SK-CO-1 were detected using capillary Western immunoassay. Serving as controls, the validated ESCC cell lines KYSE-30 ($SMARCA2^{low}$) and KYSE-450 ($SMARCA2^{high}$) were included in the experiment (Figure 44).

Indeed, all of the additional cell lines analyzed show absent or low expression of SMARCA2 in the capillary Western immunoassays (Figure 44). Whereas HuP-T4, SK-CO-1, and OV-90 exhibit no detectable SMARCA2 levels, in HCT 116 a thin band for SMARCA2 is detected (Figure 44). This goes in hand with the transcription values, showing a higher TPM value for SMARCA2 in HCT 116 (TPM= 6.2). As expected, the ESCC control cell line KYSE-450 shows high levels of SMARCA2, while KYSE-30 displays no detectable SMARCA2. For all samples SMARCA4 is detected to almost equal amounts (Figure 44).

Cas9-stably expressing cell lines were derived from the four selected lines including HuP-T4, OV-90, HCT 116, and SK-CO-1 and used for singleton depletion experiments. Similarly to the SMARCA2-deficient ESCC cell lines analyzed before, fold-depletion scores for SMARCA4 are comparable to those of the positive control POLR2A indicating strong dependencies in all cell lines assessed (Figure 45).

In summary, the hypothesis that $SMARCA2^{low}$ cell lines are sensitized towards impairment of its paralog partner SMARCA4 holds true for additional indications including colon, pancreas, and ovarian carcinoma. This might allow the extension of the pharmacological concept to other indications whereas our data highlights SMARCA4 as a vulnerability in SMARCA2-deficient cancer cells, irrespective of the tumor type.

3.10 Pharmacological degradation of SMARCA4 using a domain-swap strategy and a BRD9-BD-directed PROTAC

To elucidate whether the effects on proliferation, obtained from genetic loss of function experiments, knocking-out SMARCA4 can be translated into pharmacological inhibition, a domain-swap strategy was applied. To date, no selective inhibitors targeting exclusively the SMARCA4 helicase function are available [300, 301]. Hence, a PROTAC originally designed for targeting the BD of BRD9 (dBRD9), described by the group of Bradner [310] was harnessed. The derived dBRD9, based on the previously published inhibitor, BI-7273 generated by Boehringer Ingelheim [312], has been confirmed to selectively target BRD9, constituting a SWI/SNF complex member incorporated into the non-canonical SWI/SNF complex [170]. Thereby, full degradation has been achieved after four hours and maintained over 24 hours after the application of 0.05-5µM dBRD9 [310]. Hohmann *et al.* has proven the functional exchange of BDs within diverse BD-containing proteins in order to generate resistance towards domain-targeting inhibitors. Especially, the swap approach has been used for confirming the selectivity of certain compounds such as shown for the BRD9 inhibitor in AML [311].

Since there was no SMARCA4 selective PROTAC available, a SMARCA4-BD^{BRD9} allele was generated amenable to dBRD9 (Figure 46). In order to ensure proper folding of the protein, flanking regions as well as the domain encoding sequences were swapped (see 2.15). The SWAP construct (SMARCA4-BD^{BRD9}) was transduced into KYSE-30 via lentiviral delivery system and the cells were selected with hygromycin B over four weeks. Afterwards KO of endogenous SMARCA4 was performed by using an all-in-one vector (see 2.15) encoding the sgRNA against the N-terminal region of SMARCA4, Cas9 as well as GFP. Cells were analyzed using capillary Western immunoassay in presence of dBRD9 compared to vehicle (DMSO) control (Figure 48).

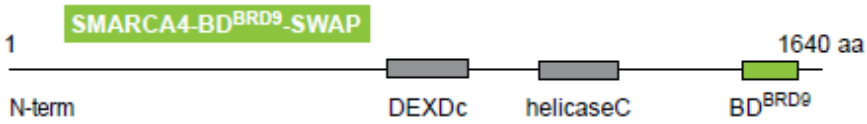


Figure 46: Domain SWAP allowing SMARCA4 degradation by usage of a BRD9-BD directed PROTAC.

Domains within SMARCA4 are indicated: N-term, DEXDc (DEAD-like helicase superfamily) and helicaseC. The swapped BD of BRD9 is highlighted in green; the detailed sequence is listed in section 462.15 (Table 2).

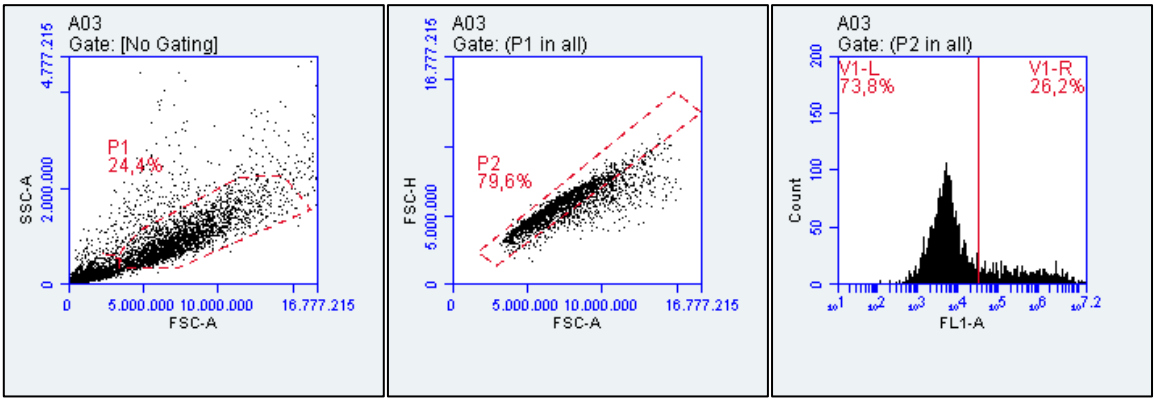


Figure 47: Transfection optimization after all-in-one plasmid transfection.

Fluorescence was measured using BD Accuri™ C6; SCC-A: side scatter area, FSC-A forward scatter area, FSC-H forward scatter height, FL1-A: filter recommended for GFP detection; A) gating for viable cells using forward and side scatter B) ensuring single cell measurements via plotting peak area against peak height C) determination of GFP⁺ cell population.

Different amounts of the plasmid and the transfection reagent were titrated in order to determine the optimal transfection conditions. Therefore, the amount of GFP expressing cells was measured. Using 8µg of plasmid DNA resulted in a successfully transfected cell population of 26.2% which represented a sufficient number of cells for subsequent cell sorting (Figure 47). The GFP$^+$ cell population was separated from the non-transfected cells via FACS. A monoclonal cell line was derived from the bulk of KYSE-30-SMARCA4-BD^{BRD9}-GFP$^+$ cells after seeding via limited dilution (0.3 cells/well) in a 96 well plate and expansion for four weeks. Engineered monoclonal cells were tested according to their endogenous SMARCA4-KO. Therefore, selected clones were treated with either 1µM of dBRD9 and compared to vehicle control (DMSO) in order to analyze SMARCA4 protein levels using capillary Western immunoassay (Figure 48).

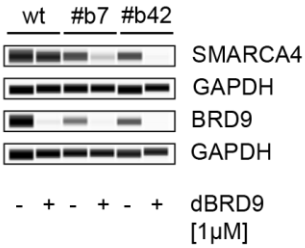


Figure 48: Capillary Western immunoassay analysis of clones after 48h treatment with dBRD9.

SMARCA4 as well as BRD9 protein levels are analyzed compared to the respective GAPDH loading control. Treatment is indicated with (+) whereas vehicle treatment (DMSO) is depicted as (-). Samples were collected from 24-well plates after 48h of incubation with dBRD9; wt: wild-type. Antibody did not distinguish between endogenous SMARCA4 and SMARCA4-BD^{BRD9}.

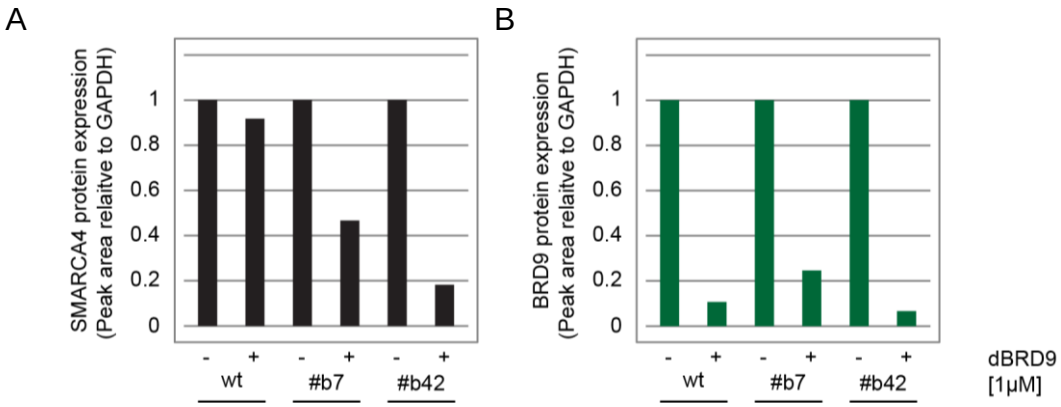


Figure 49: Quantification of degradation efficacy using capillary Western immunoassay.

Peak area of the respective analyzed protein relative to GAPDH for each individual sample. SMARCA4 degradation upon treatment with 1µM dBRD9 compared to vehicle control (DMSO). A) Degradation is only detected in the two clones (#b7 and #b42) with the SMARCA4-BD^{BRD9} allele amenable to dBRD9 treatment. B) Internal control shows efficient degradation of endogenous BRD9 in wild-type (wt) as well as in selected monoclonal cell lines.

Two out of 80 clones analyzed showed the expected effect (only selected clones are shown in Figure 48). After the application of 1µM of dBRD9, protein levels of BRD9, serving as a control for successful degradation and SMARCA4 were investigated. Indeed, both clones selected from previous analyses show a significant decrease in SMARCA4 protein levels upon the treatment with 1µM dBRD9 after 48 hours (Figure 48 and Figure 49A). In contrast, no effect on degradation of SMARCA4 is detected in the parental cell line, indicating a selectivity of the dBRD9 towards chimeric SMARCA4-BD^{BRD9} protein (Figure 49A). Treatment with 1µM dBRD9 decreases BRD9 protein levels in the parental as well as in the two generated monoclonal cell lines, confirming efficient application of the PROTAC (Figure 48 and Figure 49B). The

ability of dBRD9-mediated degradation of chimeric SMARCA4 is quantified using the internal GAPDH (Figure 49). Levels of chimeric SMARCA4-BD^{BRD9} drop only to 40% in clone#b7, whereas a pronounced degradation is observed in clone#b42 (Figure 49A). Taken together, monoclonal cell lines were successfully generated confirming the degradation of chimeric SMARCA4-BD^{BRD9} via dBRD9 and an efficient KO of endogenous SMARCA4. Mentionable, SMARCA4 detection, using capillary Western immunoassay, does not distinguish between endogenous and chimeric SMARCA4.

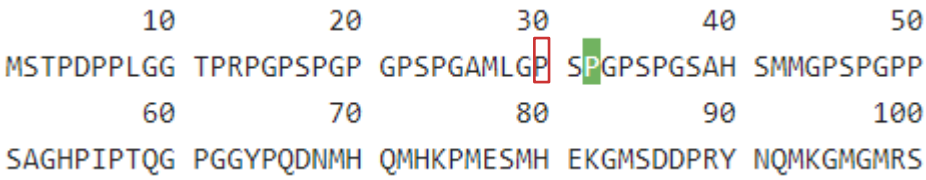


Figure 50: In-frame deletions result in deletions of proline on both alleles.

Depiction of the first 100 AA of the wild-type SMARCA4 sequence with modifications indicated; allele 1: editing event causes the deletion of proline 30 (red rectangle); allele 2: editing event causes the deletion of proline 32 (green rectangle).

To obtain insights into DNA sequence changes of endogenous SMARCA4-KO in the engineered KYSE-30-SWAP cell line, Sanger sequencing was performed. Looking at the chromatogramm, editing events are confirmed at expected cut-side, by revealing two sequences (see appendix 6.7). Different chromatograms are obtained when the KO leads to an editing event resulting in two different sequences on both alleles by the introduction of different INDELs. In both alleles, the double-strand break resulted in an in-frame deletion (sequences not shown), suggesting a fully functional protein. However, clear reduction in protein levels are detected upon dBRD9 treatment in cell lines tested for KO of endogenous SMARCA4 on protein levels (Figure 49). Since the in-frame deletion causes a loss of proline 30 (allele 1) and proline 32 (allele 2), it is speculated that the protein is not properly folded and therefore degraded, even though the double strand break was repared in-frame.

Effects upon the BRD9 PROTAC treatment are known to occur already after four hours and degradation is maintained for 24 hours in the AML cell line MOLM-13 [310]. To test the degradation efficacy the treatment on SMARCA4-BD^{BRD9} was performed

in a time resolved set-up, starting from four until 72 hours with a concentration range between 125-2000nM. SMARCA4-BD^{BRD9} protein is already sufficiently degraded after four hours and protein levels remain low until 72 hours after treatment with dBRD9 (Figure 51). Compared to the KYSE-30^{SWAP} cell line, no effect on degradation is observed on endogenous SMARCA4 in KYSE-30 parental cell line (Figure 52), suggesting a selectivity of dBRD9 for SMARCA4-BD^{BRD9}.

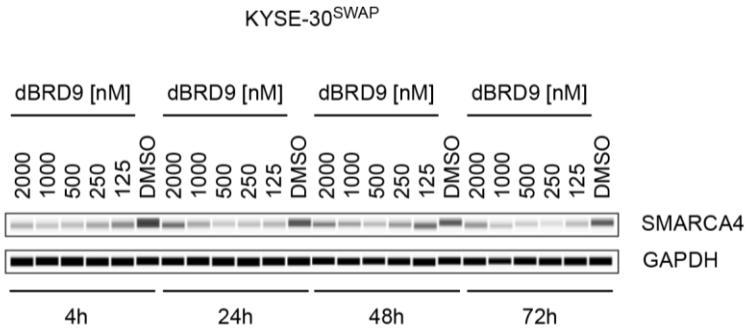


Figure 51: Capillary Western immunoassay detecting SMARCA4-BD^{BRD9} degradation in KYSE-30^{SWAP}.
Assessment of the capability of dBRD9 to degrade SMARCA4-BD^{BRD9} (125-2000nM) compared to vehicle control (DMSO) over a time frame of 4-72h.

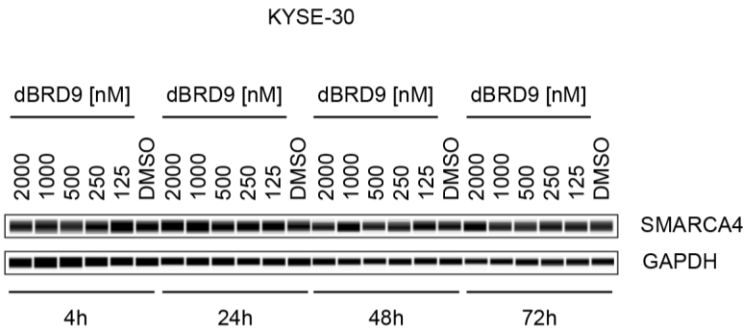


Figure 52: Capillary Western immunoassay for endogenous SMARCA4 detection in KYSE-30.
Analysis of endogenous SMARCA4 degradation upon dBRD9 treatment (125-2000nM) compared to vehicle control (DMSO) over a time period of 4-72h.

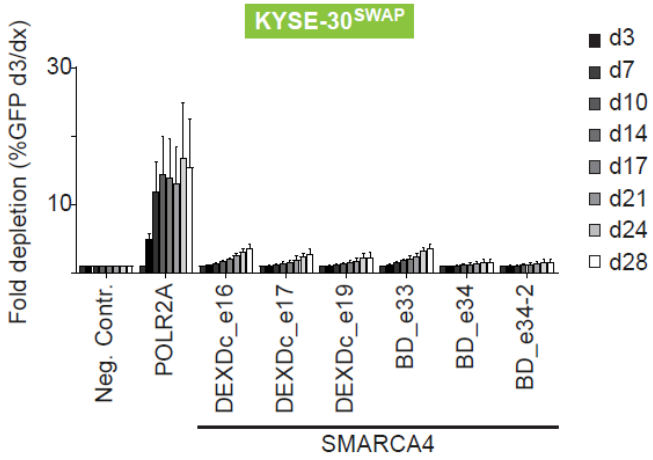
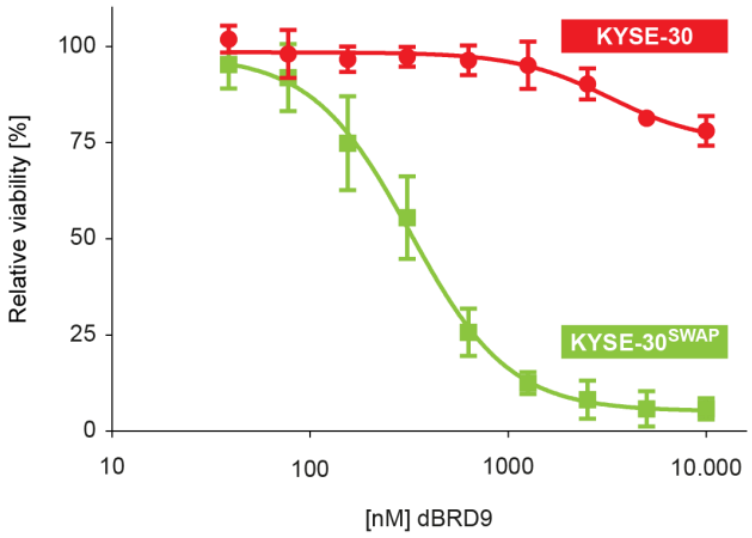


Figure 53: Depletion experiment in KYSE-30^{SWAP} cell line.

Fold depletion were obtained by dividing the percentage of GFP$^+$ cells initially obtained by lentiviral transduction by the percentage of GFP$^+$ cells at respective timepoints. The bars represent the different measurements over time indicated by dark or light grey color. Negative control (Neg.Contr.), positive control (POLR2A) as well as three sgRNAs targeting the DEXDc (DEAD-like helicase superfamily) - domain as well as the BD (bromodomain), are depicted on the x-axis. The different targeted exons are indicated with an "e". Results were obtained by three independent experiments, whereby the mean ± SD is depicted.

To proof whether the SWAP construct is functional, the generated clone KYSE-30-SWAP#b42 was stably transduced with Cas9 and depletion experiments were conducted using previously selected sgRNAs targeting the DEXDc domain and the BD of SMARCA4. KYSE-30 cell line was originally identified as SMARCA4-dependent by either targeting the ATP-binding domain (DEXDc) or the BD.

KYSE-30_Clone#b42 which lacks endogenous SMARCA4 but re-expresses the SMARCA4-BD^{BRD9} construct with the BD of BRD9 instead of the SMARCA4-BD, leads to a phenotypic rescue when applying the DEXDc as well as the BD directed SMARCA4 sgRNAs (Figure 53). The SWAP construct was codon optimized and therefore showed resistance towards sgRNA-mediated KO (Figure 53). This experiment suggests a functionality of the ectopically chimeric SMARCA4-BD^{BRD9} construct.



	IC_{50} +/- SD [nM]
KYSE-30	>10.000
KYSE-30^{SWAP}	322 +/- 112

Figure 54: CTG-viability assay in KYSE-30 vs. KYSE-30^{SWAP} upon treatment with dBRD9.

The mean + SD of four independent experiments is plotted and IC_{50} values calculated by using a four-parameter nonlinear regression model in GraphPad Prism. The experiment was performed by Janine Rippka.

Prior to the conduction of proliferation assays, a series of pre-tests was performed to determine the suitable assay conditions. The final format was chosen with 100 cells/well and incubation of ten days until Cell Titer Glo was added to determine the cell viability. The dBRD9 selectively inhibits proliferation of KYSE-30-SWAP cell line (IC_{50}= 322 ± 122nM). In contrast, no effect upon treatment is observed in KYSE-30 parental cell line (Figure 54).

To summarize these results, a functional SMARCA4-BD^{BRD9} SWAP construct is retrieved from swapping the BDs of SMARCA4 and BRD9 (Figure 53). Treatment with dBRD9 results in successful degradation of the chimeric SMARCA4-BD^{BRD9} after four hours and effects remained for 72 hours (Figure 51). After ten days, treatment with dBRD9 significantly impairs KYSE-30-SWAP proliferation whereby the parental KYSE-30 cell line is spared (Figure 54).

4 Discussion

4.1 SMARCA4 identified as a selective novel vulnerability in ESCC

ESCC is a cancer subtype with a high medical need due to its high and rapidly increasing incidence and mortality rate [1]. Whereas the two main subtypes of esophageal cancer are ESCC and EAC, ESCC comprises nearly 90% of all cases, mainly found in the so-called Asia belt and Southern and Eastern Africa [1, 6]. The mainstay of clinical therapy of ESCC is beside surgery constituted by chemo and/or radiotherapy, depending on the progression of the disease [6-8, 318]. Due to the fact that in some cases resistances occur upon chemotherapy treatment [11], a second line therapy is urgently needed to circumvent this phenomenon. Beside several approaches identifying the genomic landscape of ESCC [28-31], the development of targeted therapies, exploiting clinically selective vulnerabilities have proven unsuccessful to date [8, 318-321]. In addition, there is still a lack of diagnostic and prognostic biomarkers [6]. ESCC is diagnosed very late in development due to the absence of symptoms in early neoplasia stages. Patients which have to undergo esophagectomy are accompanied with severe side effects and often their quality of life is significantly worsened [6]. Therefore, there is an urgent need for therapeutically relevant biomarkers for ESCC diagnosis and prognosis.

This study uncovers the SWI/SNF complex ATPase SMARCA4 as a potential novel therapeutic target in ESCC. Conducting, CRISPR-Cas9 domain-based screens allowed us to robustly identify SMARCA4 dependencies in cell lines with low expression of SMARCA2 – a relationship reciprocal to the known SL interaction between SMARCA2 and SMARCA4 [238, 292, 293].

To support this concept, we confirmed strong effects in the SMARCA2^{low} cell lines KYSE-270, KYSE-510, COLO-680N, and KYSE-30 in SMARCA4-sgRNA depletion experiments. Cell lines with sufficient SMARCA2 expression, including KYSE-450, KYSE-140, KYSE-70 and KYSE-150 remained unaffected. Mentionable, two SMARCA4-dependent cell lines - T.T and KYSE-410 -, indicate no mutation and normal expression of SMARCA2. Therefore, potential SL interactions of SMARCA4 with a different deregulated SWI/SNF complex member or even novel synthetic lethal (SL) partners as known for other indications are speculated [202, 239]. Collectively,

CRISPR-Cas9 targeting of SMARCA4 revealed strong dependencies in all cell lines, indicating low expression of the paralog partner SMARCA2 in ESCC but also in additional indications including colon, pancreas and ovarian carcinoma.

4.2 The ATP-binding domain is the functional important domain of SMARCA4

The two ATPases SMARCA2 and SMARCA4 share 75% of amino acids and even a higher similarity for their ATP-binding domain [146]. Thus, targeting the ATP-binding domain with a selective inhibitor is challenging. Identification of SMARCA4 as a selective vulnerability using BD-directed sgRNAs, we initially suggested an additional dependency in ESCC on the BD. In contradiction, several other studies have linked SMARCA2/4 vulnerability on the ATPase domain [298, 300]. By conducting rescue experiments using different variants of SMARCA4, we reveal a significant stronger dependency on the ATPase domain in ESCC, which is in consonance with published data on the domain function of SMARCA4 [298, 300].

4.3 SMARCA2 and SMARCA4 indicate synthetic lethal dependencies in ESCC

We further highlight the dependency link to the hard-wired SL interaction in ESCC by conducting two complementary experiments. Our study confirms SL interactions by ectopic re-expression of SMARCA2 ($SMARCA2^{ect}$) in the two SMARCA2-deficient cell lines KYSE-510 and KYSE-30, thereby reverting the dependency on the paralog SMARCA4. Supporting this SL interaction, we show that SMARCA2-KO in SMARCA2-proficient cell lines, including KYSE-450 and KYSE-150, sensitizes them towards SMARCA4 inhibition. Similar to the rescue studies with $SMARCA4^{res}$, re-expression of ATP-binding deficient SMARCA2 did not result in phenotypic rescue upon SMARCA4-KO.

In addition, the analysis of patient samples reveals $SMARCA2^{low}$ expression in 13.8% of all samples, indicating an adequate patient population. Defining a threshold for SMARCA2 expression able to compensate loss of SMARCA4 remains difficult. Regarding the fact that homozygous loss of SMARCA4 is embryonic lethal [322], SMARCA2 might not be fully capable to adopt SMARCA4 functions. However, additional experiments are obligatory in order to determine the clear cut of $SMARCA2^{low}$ vs. $SMARCA2^{high}$ samples.

Strikingly, some cancers are able to survive with concomitant loss of SMARCA2 and SMARCA4 [238] which is particularly relevant in terms of resistance towards SWI/SNF member inhibition. For instance, in small-cell carcinoma of the ovary, hypercalcemic type (SCCOHT), bi-allelic inactivating mutations of SMARCA4 as well as loss of SMARCA2 are detected [228-230, 323], leading to the assumption of potential SWI/SNF “escape” mechanisms. Further studies are required in order to investigate alternative ATP-dependent nucleosome-remodeling complexes, such as ISWI, INO80 or CHD complexes [305] for their ability of compensating for SWI/SNF function.

4.4 Causative molecular event(s) of low SMARCA2 expression

Since *SMARCA2* is not mutated in ESCC, further investigations are necessary to determine the cause of the deficient expression. In NSCLC, epigenetic silencing of promoter regions has led to low SMARCA2 expression which has further been associated with poor prognosis [324]. Therefore, analysis of DNA methyltransferases and hyper-methylation status of the *SMARCA2* promoter would give more insights into transcriptional regulation in ESCC. Due to the fact that SMARCA2 proteins as well as transcription levels are low, the aberrant regulation is suggested to occur on transcriptional level rather than on post-translational modification. Mutations in *SMARCA2* are rarely observed [167] in cancer with the exception of adenoid cystic carcinoma (ACC), suggesting a potential tumor suppressive function [216].

4.5 SMARCA4 inhibition for therapeutic intervention in SMARCA2^{low} ESCC

SMARCA4 as well as its paralog SMARCA2 constitute the catalytic center of the SWI/SNF chromatin-remodeling complex. Into the canonical SWI/SNF complex, one of the two ATPases is incorporated in a mutual exclusive way, whereas either SMARCA4 or SMARCA2 fulfills the function of ATP hydrolysis [172] and are therefore characterized as paralog dependencies [238, 292, 293]. Those dependencies can be exploited pharmacologically following the SL concept which has been intensively investigated for targeting SMARCA2 in *SMARAC4*-mutated NSCLC cancer cell lines [238, 292, 293]. Whereas inhibition of one functional redundant partner does not affect the viability of normal cells due to compensatory effects mediated by the redundant partner, tumor cells are harmed harboring already

one inactive paralog partner. Most notably, a panel of ESCC cell lines analyzed in our study shows absent levels of SMARCA2, which are therefore rendered sensitive to SMARCA4 inhibition.

Whereas *Smarca2*$^{-/-}$ mice are viable, suggesting little toxicity when targeting SMARCA2 in patients, *Smarca4*$^{-/-}$ mice in contrast are embryonic lethal [226, 325]. Additionally, SMARCA4 has been described as a *bona fide* tumor suppressor in haplo-insufficient mice [146, 221, 226]. This is supported by spontaneous development of mammary tumors observed in *Smarca4*-null heterozygous mice [326] and enhanced tumor formation in carcinogen-induced lung cancer models ablating *Smarca4* [327]. However, *Smarca4* heterozygous tumors occur with long latency and low penetrance [226, 322], suggesting a therapeutic index for transient inhibition of SMARCA4 without risks of secondary malignancies. In addition, SMARCA4-KD in immortalized mouse embryonic fibroblasts (iMEFs) have shown no significant effects on growth impairment, suggesting limited toxic effects [307]. This is strongly supported in our study, in which a clear separation of SMARCA4-dependent and independent cell lines can be achieved. Therefore, our results suggest a clear therapeutic window in patients. More studies have to be conducted for a potential SMARCA4-selective inhibitor or PROTAC, in order to assess toxicity effects *in vivo*.

4.6 Pharmacological targeting of SMARCA4 via domain-swap and dBRD9 treatment

Most notably, we were able to proof the SMARCA4 dependency in a SMARCA2-deficient cell line using a PROTAC approach in an engineered cell model. Due to the fact that SMARCA2 and SMARCA4 are sharing a conserved ATP-binding domain [146], development of a selective inhibitor targeting the functional relevant domain is challenging. However, considering PROTAC-mediated protein degradation by targeting the BD of SMARCA4 might be beneficial. Due to the fact that no SMARCA4 selective inhibitor or PROTAC was available, we made use of a BD swap strategy. BDs have been swapped in various investigations, proofing the selectivity of a certain inhibitors targeting BRD9 in AML [311]. With this knowledge, we designed a SMARCA4 swap protein by exchanging its BD with the one of BRD9. Successfully, the SMARCA4-BD^{BRD9} construct was expressed in the SMARCA4-dependent cell line KYSE-30 and the treatment with dBRD9 [310] allowed selective degradation of

SMARCA4-BD^{BRD9}. In contrast to previously described domain-swap strategies proofing selectivity of a certain compound [311], this study highlights the degradation of a chimeric protein by the application of a PROTAC originally designed for a different protein. By conducting SMARCA4-sgRNA depletion experiments we gained evidence that the engineered SMARCA4-BD^{BRD9} is functional by phenotypically rescuing the effects obtained upon SMARCA4-KO. Finally, we showed pharmacological inhibition by using a cell viability assay. Thereby, the proliferation of KYSE-30-SWAP cell line was impaired with an IC_{50} of 322nM, sparing KYSE-30 wild-type cell line.

4.7 SMARCA4 investigations for future directions

In summary, our findings represent a novel selective vulnerability comprised by SMARCA4 in cell lines with low SMARCA2 expression. We successfully validated the concept in ESCC cell lines but also extended it to additional indications with SMARCA2 loss. By swapping the BD of SMARCA4 with the one of BRD9, we show selective degradation of SMARCA4-BD^{BRD9} and further highlighting SMARCA4 as a potential target in ESCC. To further investigate the effect of dBRD9 on ESCC proliferation, *in vivo* experiments would allow monitoring of tumor growth inhibition. To validate the functionality of the SMARCA4-BD^{BRD9} protein in terms of complex incorporation, co-immunoprecipitation with an associated SWI/SNF complex member would be necessary. Comparing the transcriptional profiles but also assay for transposase-accessible chromatin sequencing (ATAC-seq) studies [328] would reveal more insights into transcriptional regulation as well as chromatin accessibility changes between the different cell lines engineered (KYSE-30-SMARCA4-KO, KYSE-30-SWAP) compared to the parental cell line KYSE-30. Whereas conventional therapies targeting even essential proteins have been demonstrated with therapeutic window, further investigations have to be done in order to determine the toxicity effects upon interfering with members of the SWI/SNF complex [168, 329]. Our findings are particularly relevant for the treatment of non-responders to chemotherapy suggesting the interference with SMARCA4 via a PROTAC-mediated approach in order to establish a selective second-line therapy in the future.

5 References

1. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.
2. Napier, K.J., M. Scheerer, and S. Misra, *Esophageal cancer: A Review of epidemiology, pathogenesis, staging workup and treatment modalities.* World J Gastrointest Oncol, 2014. **6**(5): p. 112-20.
3. Lin, D.C., M.R. Wang, and H.P. Koeffler, *Genomic and Epigenomic Aberrations in Esophageal Squamous Cell Carcinoma and Implications for Patients.* Gastroenterology, 2018. **154**(2): p. 374-389.
4. Lorenzen, S., et al., *Cetuximab plus cisplatin-5-fluorouracil versus cisplatin-5-fluorouracil alone in first-line metastatic squamous cell carcinoma of the esophagus: a randomized phase II study of the Arbeitsgemeinschaft Internistische Onkologie.* Ann Oncol, 2009. **20**(10): p. 1667-73.
5. Bleiberg, H., et al., *Randomised phase II study of cisplatin and 5-fluorouracil (5-FU) versus cisplatin alone in advanced squamous cell oesophageal cancer.* Eur J Cancer, 1997. **33**(8): p. 1216-20.
6. Pennathur, A., et al., *Oesophageal carcinoma.* Lancet, 2013. **381**(9864): p. 400-12.
7. Enzinger, P.C. and R.J. Mayer, *Esophageal cancer.* N Engl J Med, 2003. **349**(23): p. 2241-52.
8. Ohashi, S., et al., *Recent Advances From Basic and Clinical Studies of Esophageal Squamous Cell Carcinoma.* Gastroenterology, 2015. **149**(7): p. 1700-15.
9. Tungekar, A., et al., *ESCC ATLAS: A population wide compendium of biomarkers for Esophageal Squamous Cell Carcinoma.* Sci Rep, 2018. **8**(1): p. 12715.
10. Muto, M., et al., *Improving visualization techniques by narrow band imaging and magnification endoscopy.* J Gastroenterol Hepatol, 2009. **24**(8): p. 1333-46.
11. Guo, W. and Y.G. Jiang, *Current gene expression studies in esophageal carcinoma.* Curr Genomics, 2009. **10**(8): p. 534-9.
12. Kamangar, F., G.M. Dores, and W.F. Anderson, *Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world.* J Clin Oncol, 2006. **24**(14): p. 2137-50.
13. Eslick, G.D., *Epidemiology of esophageal cancer.* Gastroenterol Clin North Am, 2009. **38**(1): p. 17-25, vii.
14. Daly, J.M., et al., *Esophageal cancer: results of an American College of Surgeons Patient Care Evaluation Study.* J Am Coll Surg, 2000. **190**(5): p. 562-72; discussion 572-3.
15. Abnet, C.C., M. Arnold, and W.Q. Wei, *Epidemiology of Esophageal Squamous Cell Carcinoma.* Gastroenterology, 2018. **154**(2): p. 360-373.
16. Steevens, J., et al., *Alcohol consumption, cigarette smoking and risk of subtypes of oesophageal and gastric cancer: a prospective cohort study.* Gut, 2010. **59**(1): p. 39-48.

17. Jemal, A., et al., *Global patterns of cancer incidence and mortality rates and trends.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(8): p. 1893-907.
18. Engel, L.S., et al., *Population attributable risks of esophageal and gastric cancers.* J Natl Cancer Inst, 2003. **95**(18): p. 1404-13.
19. Secretan, B., et al., *A review of human carcinogens--Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish.* Lancet Oncol, 2009. **10**(11): p. 1033-4.
20. Blaydon, D.C., et al., *RHBDF2 mutations are associated with tylosis, a familial esophageal cancer syndrome.* Am J Hum Genet, 2012. **90**(2): p. 340-6.
21. Rosenberg, P.S., B.P. Alter, and W. Ebell, *Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry.* Haematologica, 2008. **93**(4): p. 511-7.
22. Bosman, F.T., Carneiro, F., Hruban, R.H., Theise, N.D., *WHO Classification of Tumours of the Digestive System*. 2010, Lyon, France: IARC Press.
23. Liu, X., et al., *Genetic Alterations in Esophageal Tissues From Squamous Dysplasia to Carcinoma.* Gastroenterology, 2017. **153**(1): p. 166-177.
24. Ai, D., et al., *Patterns of distant organ metastases in esophageal cancer: a population-based study.* J Thorac Dis, 2017. **9**(9): p. 3023-3030.
25. Aghcheli, K., et al., *Prognostic factors for esophageal squamous cell carcinoma--a population-based study in Golestan Province, Iran, a high incidence area.* PLoS One, 2011. **6**(7): p. e22152.
26. Li, Y., et al., *Overexpression of MMP-2 and MMP-9 in esophageal squamous cell carcinoma.* Dis Esophagus, 2009. **22**(8): p. 664-7.
27. Liotta, L.A., et al., *Metastatic potential correlates with enzymatic degradation of basement membrane collagen.* Nature, 1980. **284**(5751): p. 67-8.
28. Gao, Y.B., et al., *Genetic landscape of esophageal squamous cell carcinoma.* Nat Genet, 2014. **46**(10): p. 1097-102.
29. Lin, D.C., et al., *Genomic and molecular characterization of esophageal squamous cell carcinoma.* Nat Genet, 2014. **46**(5): p. 467-73.
30. Song, Y., et al., *Identification of genomic alterations in oesophageal squamous cell cancer.* Nature, 2014. **509**(7498): p. 91-5.
31. Zhang, L., et al., *Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma.* Am J Hum Genet, 2015. **96**(4): p. 597-611.
32. Agrawal, N., et al., *Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1.* Science, 2011. **333**(6046): p. 1154-7.
33. Stransky, N., et al., *The mutational landscape of head and neck squamous cell carcinoma.* Science, 2011. **333**(6046): p. 1157-60.
34. Cancer Genome Atlas Research, N., et al., *The Cancer Genome Atlas Pan-Cancer analysis project.* Nat Genet, 2013. **45**(10): p. 1113-20.
35. Yang, Y., et al., *Loss of transcription factor KLF5 in the context of p53 ablation drives invasive progression of human squamous cell cancer.* Cancer Res, 2011. **71**(20): p. 6475-84.
36. Kobayashi, M., et al., *p53 Mutation analysis of low-grade dysplasia and high-grade dysplasia/carcinoma in situ of the esophagus using laser capture microdissection.* Oncology, 2006. **71**(3-4): p. 237-45.
37. Cheng, C., et al., *Whole-Genome Sequencing Reveals Diverse Models of Structural Variations in Esophageal Squamous Cell Carcinoma.* Am J Hum Genet, 2016. **98**(2): p. 256-74.

38. Takase, N., et al., *NCAM- and FGF-2-mediated FGFR1 signaling in the tumor microenvironment of esophageal cancer regulates the survival and migration of tumor-associated macrophages and cancer cells.* Cancer Lett, 2016. **380**(1): p. 47-58.
39. Yang, Y.L., et al., *Amplification of PRKCI, located in 3q26, is associated with lymph node metastasis in esophageal squamous cell carcinoma.* Genes Chromosomes Cancer, 2008. **47**(2): p. 127-36.
40. Bass, A.J., et al., *SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas.* Nat Genet, 2009. **41**(11): p. 1238-42.
41. Zheng, S., et al., *PIK3CA promotes proliferation and motility but is unassociated with lymph node metastasis or prognosis in esophageal squamous cell carcinoma.* Hum Pathol, 2016. **53**: p. 121-9.
42. Sonoda, I., et al., *Frequent silencing of low density lipoprotein receptor-related protein 1B (LRP1B) expression by genetic and epigenetic mechanisms in esophageal squamous cell carcinoma.* Cancer Res, 2004. **64**(11): p. 3741-7.
43. Zhang, W., et al., *Epidermal growth factor receptor is a prognosis predictor in patients with esophageal squamous cell carcinoma.* Ann Thorac Surg, 2014. **98**(2): p. 513-9.
44. Geutjes, E.J., P.K. Bajpe, and R. Bernards, *Targeting the epigenome for treatment of cancer.* Oncogene, 2012. **31**(34): p. 3827-44.
45. Dupont, C., D.R. Armant, and C.A. Brenner, *Epigenetics: definition, mechanisms and clinical perspective.* Semin Reprod Med, 2009. **27**(5): p. 351-7.
46. Kang, X., et al., *Personalized targeted therapy for esophageal squamous cell carcinoma.* World J Gastroenterol, 2015. **21**(25): p. 7648-58.
47. Ahrens, T.D., M. Werner, and S. Lassmann, *Epigenetics in esophageal cancers.* Cell Tissue Res, 2014. **356**(3): p. 643-55.
48. Liu, F., et al., *Aberrant overexpression of EZH2 and H3K27me3 serves as poor prognostic biomarker for esophageal squamous cell carcinoma patients.* Biomarkers, 2016. **21**(1): p. 80-90.
49. Zhong, X., et al., *Overexpression of G9a and MCM7 in oesophageal squamous cell carcinoma is associated with poor prognosis.* Histopathology, 2015. **66**(2): p. 192-200.
50. He, L.R., et al., *Prognostic impact of H3K27me3 expression on locoregional progression after chemoradiotherapy in esophageal squamous cell carcinoma.* BMC Cancer, 2009. **9**: p. 461.
51. Tzao, C., et al., *Prognostic significance of global histone modifications in resected squamous cell carcinoma of the esophagus.* Mod Pathol, 2009. **22**(2): p. 252-60.
52. Zhang, K., et al., *Comparative analysis of histone H3 and H4 post-translational modifications of esophageal squamous cell carcinoma with different invasive capabilities.* J Proteomics, 2015. **112**: p. 180-9.
53. Curradi, M., et al., *Molecular mechanisms of gene silencing mediated by DNA methylation.* Mol Cell Biol, 2002. **22**(9): p. 3157-73.
54. Maesawa, C., et al., *Inactivation of the CDKN2 gene by homozygous deletion and de novo methylation is associated with advanced stage esophageal squamous cell carcinoma.* Cancer Res, 1996. **56**(17): p. 3875-8.
55. Tokugawa, T., et al., *Modes of silencing of p16 in development of esophageal squamous cell carcinoma.* Cancer Res, 2002. **62**(17): p. 4938-44.
56. Daigo, Y., et al., *Molecular cloning of a candidate tumor suppressor gene, DLC1, from chromosome 3p21.3.* Cancer Res, 1999. **59**(8): p. 1966-72.

57. Guo, W., et al., *Aberrant hypermethylation of RASSF2 in tumors and peripheral blood DNA as a biomarker for malignant progression and poor prognosis of esophageal squamous cell carcinoma.* Clin Exp Metastasis, 2016. **33**(1): p. 73-85.
58. Cui, X., et al., *Inactivation of miR-34a by aberrant CpG methylation in Kazakh patients with esophageal carcinoma.* J Exp Clin Cancer Res, 2014. **33**: p. 20.
59. Kong, K.L., et al., *MicroRNA-375 inhibits tumour growth and metastasis in oesophageal squamous cell carcinoma through repressing insulin-like growth factor 1 receptor.* Gut, 2012. **61**(1): p. 33-42.
60. Arruebo, M., et al., *Assessment of the evolution of cancer treatment therapies.* Cancers (Basel), 2011. **3**(3): p. 3279-330.
61. Yan, L., N. Rosen, and C. Arteaga, *Targeted cancer therapies.* Chin J Cancer, 2011. **30**(1): p. 1-4.
62. Savage, D.G. and K.H. Antman, *Imatinib mesylate--a new oral targeted therapy.* N Engl J Med, 2002. **346**(9): p. 683-93.
63. Rustgi, A.K. and H.B. El-Serag, *Esophageal carcinoma.* N Engl J Med, 2014. **371**(26): p. 2499-509.
64. Medical Research Council Oesophageal Cancer Working, G., *Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial.* Lancet, 2002. **359**(9319): p. 1727-33.
65. Liu, Y., et al., *Personalized and targeted therapy of esophageal squamous cell carcinoma: an update.* Ann N Y Acad Sci, 2016. **1381**(1): p. 66-73.
66. Kato, K., et al., *Phase II study of chemoradiotherapy with 5-fluorouracil and cisplatin for Stage II-III esophageal squamous cell carcinoma: JCOG trial (JCOG 9906).* Int J Radiat Oncol Biol Phys, 2011. **81**(3): p. 684-90.
67. Reed, C.E., *Comparison of different treatments for unresectable esophageal cancer.* World J Surg, 1995. **19**(6): p. 828-35.
68. Cunningham, D., et al., *Capecitabine and oxaliplatin for advanced esophagogastric cancer.* N Engl J Med, 2008. **358**(1): p. 36-46.
69. Xu, Y., et al., *Concurrent radiotherapy with gefitinib in elderly patients with esophageal squamous cell carcinoma: Preliminary results of a phase II study.* Oncotarget, 2015. **6**(35): p. 38429-39.
70. Dutton, S.J., et al., *Gefitinib for oesophageal cancer progressing after chemotherapy (COG): a phase 3, multicentre, double-blind, placebo-controlled randomised trial.* Lancet Oncol, 2014. **15**(8): p. 894-904.
71. Lu, M., et al., *Nimotuzumab plus paclitaxel and cisplatin as the first line treatment for advanced esophageal squamous cell cancer: A single centre prospective phase II trial.* Cancer Sci, 2016. **107**(4): p. 486-90.
72. Musgrove, E.A., et al., *Cyclin D as a therapeutic target in cancer.* Nat Rev Cancer, 2011. **11**(8): p. 558-72.
73. Hong, B., et al., *Targeting tumor suppressor p53 for cancer therapy: strategies, challenges and opportunities.* Curr Drug Targets, 2014. **15**(1): p. 80-9.
74. Khoo, K.H., C.S. Verma, and D.P. Lane, *Drugging the p53 pathway: understanding the route to clinical efficacy.* Nat Rev Drug Discov, 2014. **13**(3): p. 217-36.
75. Okazaki, T. and T. Honjo, *PD-1 and PD-1 ligands: from discovery to clinical application.* Int Immunol, 2007. **19**(7): p. 813-24.
76. Pectasides, E., *Immune checkpoint blockade in esophageal squamous cell carcinoma: is it ready for prime time?* J Thorac Dis, 2018. **10**(3): p. 1276-1279.
77. Sharma, P. and J.P. Allison, *The future of immune checkpoint therapy.* Science, 2015. **348**(6230): p. 56-61.

78. Ohigashi, Y., et al., *Clinical significance of programmed death-1 ligand-1 and programmed death-1 ligand-2 expression in human esophageal cancer.* Clin Cancer Res, 2005. **11**(8): p. 2947-53.
79. Shi, J., et al., *Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains.* Nat Biotechnol, 2015. **33**(6): p. 661-7.
80. Meyers, R.M., et al., *Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells.* Nat Genet, 2017. **49**(12): p. 1779-1784.
81. McDonald, E.R., 3rd, et al., *Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships Uncovered by Large-Scale, Deep RNAi Screening.* Cell, 2017. **170**(3): p. 577-592 e10.
82. Wang, T., et al., *Genetic screens in human cells using the CRISPR-Cas9 system.* Science, 2014. **343**(6166): p. 80-4.
83. Zhou, Y., et al., *High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells.* Nature, 2014. **509**(7501): p. 487-91.
84. Shalem, O., et al., *Genome-scale CRISPR-Cas9 knockout screening in human cells.* Science, 2014. **343**(6166): p. 84-87.
85. Ishino, Y., et al., *Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product.* J Bacteriol, 1987. **169**(12): p. 5429-33.
86. Makarova, K.S., et al., *A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis.* Nucleic Acids Res, 2002. **30**(2): p. 482-96.
87. Guy, C.P., et al., *A novel nuclease-ATPase (Nar71) from archaea is part of a proposed thermophilic DNA repair system.* Nucleic Acids Res, 2004. **32**(21): p. 6176-86.
88. Bolotin, A., et al., *Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin.* Microbiology, 2005. **151**(Pt 8): p. 2551-61.
89. Mojica, F.J., et al., *Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements.* J Mol Evol, 2005. **60**(2): p. 174-82.
90. Pourcel, C., G. Salvignol, and G. Vergnaud, *CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies.* Microbiology, 2005. **151**(Pt 3): p. 653-63.
91. Jansen, R., et al., *Identification of genes that are associated with DNA repeats in prokaryotes.* Mol Microbiol, 2002. **43**(6): p. 1565-75.
92. Haft, D.H., et al., *A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes.* PLoS Comput Biol, 2005. **1**(6): p. e60.
93. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes.* Science, 2007. **315**(5819): p. 1709-12.
94. Brouns, S.J., et al., *Small CRISPR RNAs guide antiviral defense in prokaryotes.* Science, 2008. **321**(5891): p. 960-4.
95. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA.* Science, 2008. **322**(5909): p. 1843-5.
96. van der Oost, J., et al., *Unravelling the structural and mechanistic basis of CRISPR-Cas systems.* Nat Rev Microbiol, 2014. **12**(7): p. 479-92.

97. Bondy-Denomy, J. and A.R. Davidson, *To acquire or resist: the complex biological effects of CRISPR-Cas systems.* Trends Microbiol, 2014. **22**(4): p. 218-25.
98. Makarova, K.S., et al., *Evolution and classification of the CRISPR-Cas systems.* Nat Rev Microbiol, 2011. **9**(6): p. 467-77.
99. Makarova, K.S., et al., *An updated evolutionary classification of CRISPR-Cas systems.* Nat Rev Microbiol, 2015. **13**(11): p. 722-36.
100. Shmakov, S., et al., *Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems.* Mol Cell, 2015. **60**(3): p. 385-97.
101. Makarova, K.S., et al., *Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems.* Biol Direct, 2011. **6**: p. 38.
102. Nam, K.H., et al., *Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system.* Structure, 2012. **20**(9): p. 1574-84.
103. Haurwitz, R.E., et al., *Sequence- and structure-specific RNA processing by a CRISPR endonuclease.* Science, 2010. **329**(5997): p. 1355-8.
104. Hatoum-Aslan, A., I. Maniv, and L.A. Marraffini, *Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site.* Proc Natl Acad Sci U S A, 2011. **108**(52): p. 21218-22.
105. Rouillon, C., et al., *Structure of the CRISPR interference complex CSM reveals key similarities with cascade.* Mol Cell, 2013. **52**(1): p. 124-34.
106. Hale, C.R., et al., *RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex.* Cell, 2009. **139**(5): p. 945-56.
107. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
108. Gasiunas, G., et al., *Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.* Proc Natl Acad Sci U S A, 2012. **109**(39): p. E2579-86.
109. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.
110. Bibikova, M., et al., *Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases.* Genetics, 2002. **161**(3): p. 1169-75.
111. Bibikova, M., et al., *Enhancing gene targeting with designed zinc finger nucleases.* Science, 2003. **300**(5620): p. 764.
112. Boch, J., et al., *Breaking the code of DNA binding specificity of TAL-type III effectors.* Science, 2009. **326**(5959): p. 1509-12.
113. Christian, M., et al., *Targeting DNA double-strand breaks with TAL effector nucleases.* Genetics, 2010. **186**(2): p. 757-61.
114. Moscou, M.J. and A.J. Bogdanove, *A simple cipher governs DNA recognition by TAL effectors.* Science, 2009. **326**(5959): p. 1501.
115. Doudna, J.A. and E. Charpentier, *Genome editing. The new frontier of genome engineering with CRISPR-Cas9.* Science, 2014. **346**(6213): p. 1258096.
116. Sigoillot, F.D. and R.W. King, *Vigilance and validation: Keys to success in RNAi screening.* ACS Chem Biol, 2011. **6**(1): p. 47-60.
117. Kaelin, W.G., Jr., *Molecular biology. Use and abuse of RNAi to study mammalian gene function.* Science, 2012. **337**(6093): p. 421-2.
118. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.

119. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.
120. Sternberg, S.H., et al., *DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.* Nature, 2014. **507**(7490): p. 62-7.
121. Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nat Biotechnol, 2013. **31**(9): p. 827-32.
122. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.
123. Schwank, G., et al., *Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients.* Cell Stem Cell, 2013. **13**(6): p. 653-8.
124. Shao, D.D., et al., *ATARiS: computational quantification of gene suppression phenotypes from multisample RNAi screens.* Genome Res, 2013. **23**(4): p. 665-78.
125. Doench, J.G., et al., *Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9.* Nat Biotechnol, 2016. **34**(2): p. 184-191.
126. Aguirre, A.J., et al., *Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting.* Cancer Discov, 2016. **6**(8): p. 914-29.
127. Koike-Yusa, H., et al., *Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library.* Nat Biotechnol, 2014. **32**(3): p. 267-73.
128. Doench, J.G., et al., *Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation.* Nat Biotechnol, 2014. **32**(12): p. 1262-7.
129. Khorasanizadeh, S., *The nucleosome: from genomic organization to genomic regulation.* Cell, 2004. **116**(2): p. 259-72.
130. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA.* Science, 1974. **184**(4139): p. 868-71.
131. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.
132. Noll, M. and R.D. Kornberg, *Action of micrococcal nuclease on chromatin and the location of histone H1.* J Mol Biol, 1977. **109**(3): p. 393-404.
133. Clapier, C.R. and B.R. Cairns, *The biology of chromatin remodeling complexes.* Annu Rev Biochem, 2009. **78**: p. 273-304.
134. Oudet, P., M. Gross-Bellard, and P. Chambon, *Electron microscopic and biochemical evidence that chromatin structure is a repeating unit.* Cell, 1975. **4**(4): p. 281-300.
135. Neigeborn, L. and M. Carlson, *Genes affecting the regulation of SUC2 gene expression by glucose repression in Saccharomyces cerevisiae.* Genetics, 1984. **108**(4): p. 845-58.
136. Stern, M., R. Jensen, and I. Herskowitz, *Five SWI genes are required for expression of the HO gene in yeast.* J Mol Biol, 1984. **178**(4): p. 853-68.
137. Peterson, C.L. and I. Herskowitz, *Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription.* Cell, 1992. **68**(3): p. 573-83.
138. Cairns, B.R., et al., *A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast.* Proc Natl Acad Sci U S A, 1994. **91**(5): p. 1950-4.
139. Cote, J., et al., *Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex.* Science, 1994. **265**(5168): p. 53-60.

140. Hirschhorn, J.N., et al., *Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure.* Genes Dev, 1992. **6**(12A): p. 2288-98.
141. Laurent, B.C., I. Treich, and M. Carlson, *The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation.* Genes Dev, 1993. **7**(4): p. 583-91.
142. Kennison, J.A. and J.W. Tamkun, *Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila.* Proc Natl Acad Sci U S A, 1988. **85**(21): p. 8136-40.
143. Brizuela, B.J., et al., *Genetic analysis of the brahma gene of Drosophila melanogaster and polytene chromosome subdivisions 72AB.* Genetics, 1994. **137**(3): p. 803-13.
144. Elfring, L.K., et al., *Genetic analysis of brahma: the Drosophila homolog of the yeast chromatin remodeling factor SWI2/SNF2.* Genetics, 1998. **148**(1): p. 251-65.
145. Papoulas, O., et al., *The HMG-domain protein BAP111 is important for the function of the BRM chromatin-remodeling complex in vivo.* Proc Natl Acad Sci U S A, 2001. **98**(10): p. 5728-33.
146. Khavari, P.A., et al., *BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription.* Nature, 1993. **366**(6451): p. 170-4.
147. Kwon, H., et al., *Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex.* Nature, 1994. **370**(6489): p. 477-81.
148. Imbalzano, A.N., et al., *Facilitated binding of TATA-binding protein to nucleosomal DNA.* Nature, 1994. **370**(6489): p. 481-5.
149. Cohen, S.M., et al., *BRG1 co-localizes with DNA replication factors and is required for efficient replication fork progression.* Nucleic Acids Res, 2010. **38**(20): p. 6906-19.
150. Flanagan, J.F. and C.L. Peterson, *A role for the yeast SWI/SNF complex in DNA replication.* Nucleic Acids Res, 1999. **27**(9): p. 2022-8.
151. Hara, R. and A. Sancar, *The SWI/SNF chromatin-remodeling factor stimulates repair by human excision nuclease in the mononucleosome core particle.* Mol Cell Biol, 2002. **22**(19): p. 6779-87.
152. Park, J.H., et al., *Mammalian SWI/SNF complexes facilitate DNA double-strand break repair by promoting gamma-H2AX induction.* EMBO J, 2006. **25**(17): p. 3986-97.
153. Dykhuizen, E.C., et al., *BAF complexes facilitate decatenation of DNA by topoisomerase IIalpha.* Nature, 2013. **497**(7451): p. 624-7.
154. Miller, E.L., et al., *TOP2 synergizes with BAF chromatin remodeling for both resolution and formation of facultative heterochromatin.* Nat Struct Mol Biol, 2017. **24**(4): p. 344-352.
155. Chi, T.H., et al., *Reciprocal regulation of CD4/CD8 expression by SWI/SNF-like BAF complexes.* Nature, 2002. **418**(6894): p. 195-9.
156. Wang, W., et al., *Purification and biochemical heterogeneity of the mammalian SWI-SNF complex.* EMBO J, 1996. **15**(19): p. 5370-82.
157. Wang, W., et al., *Diversity and specialization of mammalian SWI/SNF complexes.* Genes Dev, 1996. **10**(17): p. 2117-30.
158. Nie, Z., et al., *A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex.* Mol Cell Biol, 2000. **20**(23): p. 8879-88.

159. Tolstorukov, M.Y., et al., *Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters.* Proc Natl Acad Sci U S A, 2013. **110**(25): p. 10165-70.
160. Becker, P.B. and W. Horz, *ATP-dependent nucleosome remodeling.* Annu Rev Biochem, 2002. **71**: p. 247-73.
161. Saha, A., J. Wittmeyer, and B.R. Cairns, *Chromatin remodelling: the industrial revolution of DNA around histones.* Nat Rev Mol Cell Biol, 2006. **7**(6): p. 437-47.
162. Singleton, M.R. and D.B. Wigley, *Modularity and specialization in superfamily 1 and 2 helicases.* J Bacteriol, 2002. **184**(7): p. 1819-26.
163. Filippakopoulos, P., et al., *Selective inhibition of BET bromodomains.* Nature, 2010. **468**(7327): p. 1067-73.
164. Winston, F. and C.D. Allis, *The bromodomain: a chromatin-targeting module?* Nat Struct Biol, 1999. **6**(7): p. 601-4.
165. Hohmann, A.F. and C.R. Vakoc, *A rationale to target the SWI/SNF complex for cancer therapy.* Trends Genet, 2014. **30**(8): p. 356-63.
166. Hay, D.A., et al., *Discovery and optimization of small-molecule ligands for the CBP/p300 bromodomains.* J Am Chem Soc, 2014. **136**(26): p. 9308-19.
167. Kadoch, C. and G.R. Crabtree, *Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics.* Sci Adv, 2015. **1**(5): p. e1500447.
168. Masliah-Planchon, J., et al., *SWI/SNF chromatin remodeling and human malignancies.* Annu Rev Pathol, 2015. **10**: p. 145-71.
169. Tang, Y., et al., *Linking long non-coding RNAs and SWI/SNF complexes to chromatin remodeling in cancer.* Mol Cancer, 2017. **16**(1): p. 42.
170. Michel, B.C., et al., *A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation.* Nat Cell Biol, 2018. **20**(12): p. 1410-1420.
171. Kadoch, C., et al., *Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy.* Nat Genet, 2013. **45**(6): p. 592-601.
172. Wu, J.I., J. Lessard, and G.R. Crabtree, *Understanding the words of chromatin regulation.* Cell, 2009. **136**(2): p. 200-6.
173. Yan, Z., et al., *PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes.* Genes Dev, 2005. **19**(14): p. 1662-7.
174. Kasten, P., et al., *Three-dimensional motion analysis of compensatory movements in patients with radioulnar synostosis performing activities of daily living.* J Orthop Sci, 2009. **14**(3): p. 307-12.
175. Pan, J., et al., *Interrogation of Mammalian Protein Complex Structure, Function, and Membership Using Genome-Scale Fitness Screens.* Cell Syst, 2018. **6**(5): p. 555-568 e7.
176. Alpsoy, A. and E.C. Dykhuizen, *Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes.* J Biol Chem, 2018. **293**(11): p. 3892-3903.
177. Kaeser, M.D., et al., *BRD7, a novel PBAF-specific SWI/SNF subunit, is required for target gene activation and repression in embryonic stem cells.* J Biol Chem, 2008. **283**(47): p. 32254-63.
178. Son, E.Y. and G.R. Crabtree, *The role of BAF (mSWI/SNF) complexes in mammalian neural development.* Am J Med Genet C Semin Med Genet, 2014. **166C**(3): p. 333-49.

179. Sarnowska, E., et al., *The Role of SWI/SNF Chromatin Remodeling Complexes in Hormone Crosstalk.* Trends Plant Sci, 2016. **21**(7): p. 594-608.
180. Zhang, P., et al., *Role of BAF60a/BAF60c in chromatin remodeling and hepatic lipid metabolism.* Nutr Metab (Lond), 2016. **13**: p. 30.
181. Takeuchi, J.K. and B.G. Bruneau, *Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors.* Nature, 2009. **459**(7247): p. 708-11.
182. Lickert, H., et al., *Baf60c is essential for function of BAF chromatin remodelling complexes in heart development.* Nature, 2004. **432**(7013): p. 107-12.
183. Albini, S., et al., *Epigenetic reprogramming of human embryonic stem cells into skeletal muscle cells and generation of contractile myospheres.* Cell Rep, 2013. **3**(3): p. 661-70.
184. Toto, P.C., P.L. Puri, and S. Albini, *SWI/SNF-directed stem cell lineage specification: dynamic composition regulates specific stages of skeletal myogenesis.* Cell Mol Life Sci, 2016. **73**(20): p. 3887-96.
185. Alver, B.H., et al., *The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers.* Nat Commun, 2017. **8**: p. 14648.
186. Kadoch, C., et al., *Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states.* Nat Genet, 2017. **49**(2): p. 213-222.
187. Strobeck, M.W., et al., *The BRG-1 subunit of the SWI/SNF complex regulates CD44 expression.* J Biol Chem, 2001. **276**(12): p. 9273-8.
188. DiRenzo, J., et al., *BRG-1 is recruited to estrogen-responsive promoters and cooperates with factors involved in histone acetylation.* Mol Cell Biol, 2000. **20**(20): p. 7541-9.
189. Fryer, C.J. and T.K. Archer, *Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex.* Nature, 1998. **393**(6680): p. 88-91.
190. Cheng, S.W., et al., *c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function.* Nat Genet, 1999. **22**(1): p. 102-5.
191. Bochar, D.A., et al., *BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer.* Cell, 2000. **102**(2): p. 257-65.
192. Strobeck, M.W., et al., *BRG-1 is required for RB-mediated cell cycle arrest.* Proc Natl Acad Sci U S A, 2000. **97**(14): p. 7748-53.
193. Zhang, H.S., et al., *Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF.* Cell, 2000. **101**(1): p. 79-89.
194. Dunaief, J.L., et al., *The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest.* Cell, 1994. **79**(1): p. 119-30.
195. Strober, B.E., et al., *Functional interactions between the hBRM/hBRG1 transcriptional activators and the pRB family of proteins.* Mol Cell Biol, 1996. **16**(4): p. 1576-83.
196. Bartkova, J., J. Lukas, and J. Bartek, *Aberrations of the G1- and G1/S-regulating genes in human cancer.* Prog Cell Cycle Res, 1997. **3**: p. 211-20.
197. Bartek, J., J. Bartkova, and J. Lukas, *The retinoblastoma protein pathway in cell cycle control and cancer.* Exp Cell Res, 1997. **237**(1): p. 1-6.
198. Dyson, N., *The regulation of E2F by pRB-family proteins.* Genes Dev, 1998. **12**(15): p. 2245-62.
199. Sherr, C.J., *Cancer cell cycles.* Science, 1996. **274**(5293): p. 1672-7.

200. Wang, J.Y., E.S. Knudsen, and P.J. Welch, *The retinoblastoma tumor suppressor protein.* Adv Cancer Res, 1994. **64**: p. 25-85.
201. Shain, A.H. and J.R. Pollack, *The spectrum of SWI/SNF mutations, ubiquitous in human cancers.* PLoS One, 2013. **8**(1): p. e55119.
202. Versteege, I., et al., *Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer.* Nature, 1998. **394**(6689): p. 203-6.
203. Doan, D.N., et al., *Loss of the INI1 tumor suppressor does not impair the expression of multiple BRG1-dependent genes or the assembly of SWI/SNF enzymes.* Oncogene, 2004. **23**(19): p. 3462-73.
204. Wilson, B.G., et al., *Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation.* Cancer Cell, 2010. **18**(4): p. 316-28.
205. Romero, O.A., et al., *The tumour suppressor and chromatin-remodelling factor BRG1 antagonizes Myc activity and promotes cell differentiation in human cancer.* EMBO Mol Med, 2012. **4**(7): p. 603-16.
206. Eroglu, E., et al., *SWI/SNF complex prevents lineage reversion and induces temporal patterning in neural stem cells.* Cell, 2014. **156**(6): p. 1259-1273.
207. Arnaud, O., F. Le Loarer, and F. Tirode, *BAFfling pathologies: Alterations of BAF complexes in cancer.* Cancer Lett, 2018. **419**: p. 266-279.
208. Guichard, C., et al., *Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma.* Nat Genet, 2012. **44**(6): p. 694-8.
209. Huang, J., et al., *Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma.* Nat Genet, 2012. **44**(10): p. 1117-21.
210. Imielinski, M., et al., *Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing.* Cell, 2012. **150**(6): p. 1107-20.
211. Zang, Z.J., et al., *Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes.* Nat Genet, 2012. **44**(5): p. 570-4.
212. Wang, K., et al., *Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer.* Nat Genet, 2011. **43**(12): p. 1219-23.
213. Gui, Y., et al., *Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder.* Nat Genet, 2011. **43**(9): p. 875-8.
214. Guo, G., et al., *Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation.* Nat Genet, 2013. **45**(12): p. 1459-63.
215. Jiao, Y., et al., *Exome sequencing identifies frequent inactivating mutations in BAP1, ARID1A and PBRM1 in intrahepatic cholangiocarcinomas.* Nat Genet, 2013. **45**(12): p. 1470-1473.
216. Ho, A.S., et al., *The mutational landscape of adenoid cystic carcinoma.* Nat Genet, 2013. **45**(7): p. 791-8.
217. Korpanty, G.J., et al., *Association of BRM promoter polymorphisms and esophageal adenocarcinoma outcome.* Oncotarget, 2017. **8**(17): p. 28093-28100.
218. Wang, J.R., et al., *Association of two BRM promoter polymorphisms with head and neck squamous cell carcinoma risk.* Carcinogenesis, 2013. **34**(5): p. 1012-7.
219. Liu, G., et al., *Two novel BRM insertion promoter sequence variants are associated with loss of BRM expression and lung cancer risk.* Oncogene, 2011. **30**(29): p. 3295-304.

220. Wong, K.M., et al., *Two BRM promoter insertion polymorphisms increase the risk of early-stage upper aerodigestive tract cancers.* Cancer Med, 2014. **3**(2): p. 426-33.
221. Glaros, S., et al., *The reversible epigenetic silencing of BRM: implications for clinical targeted therapy.* Oncogene, 2007. **26**(49): p. 7058-66.
222. An, H.X., et al., *Two regions of deletion in 9p23-24 in sporadic breast cancer.* Cancer Res, 1999. **59**(16): p. 3941-3.
223. Girard, L., et al., *Genome-wide allelotyping of lung cancer identifies new regions of allelic loss, differences between small cell lung cancer and non-small cell lung cancer, and loci clustering.* Cancer Res, 2000. **60**(17): p. 4894-906.
224. Gunduz, E., et al., *Loss of heterozygosity at the 9p21-24 region and identification of BRM as a candidate tumor suppressor gene in head and neck squamous cell carcinoma.* Cancer Invest, 2009. **27**(6): p. 661-8.
225. Sarkar, S., et al., *A novel ankyrin repeat-containing gene (Kank) located at 9p24 is a growth suppressor of renal cell carcinoma.* J Biol Chem, 2002. **277**(39): p. 36585-91.
226. Bultman, S.J., et al., *Characterization of mammary tumors from Brg1 heterozygous mice.* Oncogene, 2008. **27**(4): p. 460-8.
227. Witkowski, L., et al., *Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type.* Nat Genet, 2014. **46**(5): p. 438-43.
228. Le Loarer, F., et al., *SMARCA4 inactivation defines a group of undifferentiated thoracic malignancies transcriptionally related to BAF-deficient sarcomas.* Nat Genet, 2015. **47**(10): p. 1200-5.
229. Karnezis, A.N., et al., *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, 2016. **238**(3): p. 389-400.
230. Jelinic, P., et al., *Concomitant loss of SMARCA2 and SMARCA4 expression in small cell carcinoma of the ovary, hypercalcemic type.* Mod Pathol, 2016. **29**(1): p. 60-6.
231. Love, C., et al., *The genetic landscape of mutations in Burkitt lymphoma.* Nat Genet, 2012. **44**(12): p. 1321-5.
232. Medina, P.P., et al., *Genetic and epigenetic screening for gene alterations of the chromatin-remodeling factor, SMARCA4/BRG1, in lung tumors.* Genes Chromosomes Cancer, 2004. **41**(2): p. 170-7.
233. Medina, P.P., et al., *Frequent BRG1/SMARCA4-inactivating mutations in human lung cancer cell lines.* Hum Mutat, 2008. **29**(5): p. 617-22.
234. Dulak, A.M., et al., *Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity.* Nat Genet, 2013. **45**(5): p. 478-86.
235. Matsubara, D., et al., *Lung cancer with loss of BRG1/BRM, shows epithelial mesenchymal transition phenotype and distinct histologic and genetic features.* Cancer Sci, 2013. **104**(2): p. 266-73.
236. Reisman, D.N., et al., *Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis.* Cancer Res, 2003. **63**(3): p. 560-6.
237. Wong, A.K., et al., *BRG1, a component of the SWI-SNF complex, is mutated in multiple human tumor cell lines.* Cancer Res, 2000. **60**(21): p. 6171-7.

238. Oike, T., et al., *Inactivating mutations in SWI/SNF chromatin remodeling genes in human cancer.* Jpn J Clin Oncol, 2013. **43**(9): p. 849-55.
239. Roberts, C.W., et al., *Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice.* Proc Natl Acad Sci U S A, 2000. **97**(25): p. 13796-800.
240. Kadoch, C. and G.R. Crabtree, *Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma.* Cell, 2013. **153**(1): p. 71-85.
241. Ramos, P., et al., *Small cell carcinoma of the ovary, hypercalcemic type, displays frequent inactivating germline and somatic mutations in SMARCA4.* Nat Genet, 2014. **46**(5): p. 427-9.
242. Jelinic, P., et al., *Recurrent SMARCA4 mutations in small cell carcinoma of the ovary.* Nat Genet, 2014. **46**(5): p. 424-6.
243. Clark, J., et al., *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, 1994. **7**(4): p. 502-8.
244. Crew, A.J., et al., *Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with homology to the Kruppel-associated box in human synovial sarcoma.* EMBO J, 1995. **14**(10): p. 2333-40.
245. McBride, M.J., et al., *The SS18-SSX Fusion Oncoprotein Hijacks BAF Complex Targeting and Function to Drive Synovial Sarcoma.* Cancer Cell, 2018. **33**(6): p. 1128-1141 e7.
246. Boulay, G., et al., *Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain.* Cell, 2017. **171**(1): p. 163-178 e19.
247. Prensner, J.R., et al., *The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex.* Nat Genet, 2013. **45**(11): p. 1392-8.
248. Dobzhansky, T., *Genetics of natural populations; recombination and variability in populations of Drosophila pseudoobscura.* Genetics, 1946. **31**: p. 269-90.
249. Hennessy, K.M. and D. Botstein, *Regulation of DNA replication during the yeast cell cycle.* Cold Spring Harb Symp Quant Biol, 1991. **56**: p. 279-84.
250. Hennessy, K.M., et al., *A group of interacting yeast DNA replication genes.* Genes Dev, 1991. **5**(6): p. 958-69.
251. Kaiser, C.A. and R. Schekman, *Distinct sets of SEC genes govern transport vesicle formation and fusion early in the secretory pathway.* Cell, 1990. **61**(4): p. 723-33.
252. Lucchesi, J.C., *Synthetic lethality and semi-lethality among functionally related mutants of Drosophila melanfgaster.* Genetics, 1968. **59**(1): p. 37-44.
253. O'Neil, N.J., M.L. Bailey, and P. Hieter, *Synthetic lethality and cancer.* Nat Rev Genet, 2017. **18**(10): p. 613-623.
254. Ashworth, A., C.J. Lord, and J.S. Reis-Filho, *Genetic interactions in cancer progression and treatment.* Cell, 2011. **145**(1): p. 30-8.
255. Kaufman, B., et al., *Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation.* J Clin Oncol, 2015. **33**(3): p. 244-50.
256. Lord, C.J. and A. Ashworth, *PARP inhibitors: Synthetic lethality in the clinic.* Science, 2017. **355**(6330): p. 1152-1158.
257. Guha, M., *PARP inhibitors stumble in breast cancer.* Nat Biotechnol, 2011. **29**(5): p. 373-4.
258. Satoh, M.S. and T. Lindahl, *Role of poly(ADP-ribose) formation in DNA repair.* Nature, 1992. **356**(6367): p. 356-8.

259. Eustermann, S., et al., *Structural Basis of Detection and Signaling of DNA Single-Strand Breaks by Human PARP-1.* Mol Cell, 2015. **60**(5): p. 742-754.
260. Dawicki-McKenna, J.M., et al., *PARP-1 Activation Requires Local Unfolding of an Autoinhibitory Domain.* Mol Cell, 2015. **60**(5): p. 755-768.
261. De Vos, M., V. Schreiber, and F. Dantzer, *The diverse roles and clinical relevance of PARPs in DNA damage repair: current state of the art.* Biochem Pharmacol, 2012. **84**(2): p. 137-46.
262. Krishnakumar, R. and W.L. Kraus, *The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets.* Mol Cell, 2010. **39**(1): p. 8-24.
263. Moynahan, M.E. and M. Jasin, *Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis.* Nat Rev Mol Cell Biol, 2010. **11**(3): p. 196-207.
264. King, M.C., *"The race" to clone BRCA1.* Science, 2014. **343**(6178): p. 1462-5.
265. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
266. Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2.* Nature, 1995. **378**(6559): p. 789-92.
267. Moynahan, M.E., T.Y. Cui, and M. Jasin, *Homology-directed dna repair, mitomycin-c resistance, and chromosome stability is restored with correction of a Brca1 mutation.* Cancer Res, 2001. **61**(12): p. 4842-50.
268. Moynahan, M.E., A.J. Pierce, and M. Jasin, *BRCA2 is required for homology-directed repair of chromosomal breaks.* Mol Cell, 2001. **7**(2): p. 263-72.
269. Tutt, A., et al., *Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences.* EMBO J, 2001. **20**(17): p. 4704-16.
270. Dhillon, K.K., et al., *Synthetic lethality: the road to novel therapies for breast cancer.* Endocr Relat Cancer, 2016. **23**(10): p. T39-55.
271. Farmer, H., et al., *Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy.* Nature, 2005. **434**(7035): p. 917-21.
272. Bryant, H.E., et al., *Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase.* Nature, 2005. **434**(7035): p. 913-7.
273. Fong, P.C., et al., *Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers.* N Engl J Med, 2009. **361**(2): p. 123-34.
274. Dziadkowiec, K.N., et al., *PARP inhibitors: review of mechanisms of action and BRCA1/2 mutation targeting.* Prz Menopauzalny, 2016. **15**(4): p. 215-219.
275. Shen, Y., et al., *BMN 673, a novel and highly potent PARP1/2 inhibitor for the treatment of human cancers with DNA repair deficiency.* Clin Cancer Res, 2013. **19**(18): p. 5003-15.
276. Murai, J., et al., *Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors.* Cancer Res, 2012. **72**(21): p. 5588-99.
277. Murai, J., et al., *Stereospecific PARP trapping by BMN 673 and comparison with olaparib and rucaparib.* Mol Cancer Ther, 2014. **13**(2): p. 433-43.
278. Pommier, Y., M.J. O'Connor, and J. de Bono, *Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action.* Sci Transl Med, 2016. **8**(362): p. 362ps17.
279. Oza, A.M., et al., *Olaparib combined with chemotherapy for recurrent platinum-sensitive ovarian cancer: a randomised phase 2 trial.* Lancet Oncol, 2015. **16**(1): p. 87-97.

280. Audeh, M.W., et al., *Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial.* Lancet, 2010. **376**(9737): p. 245-51.
281. Kim, G., et al., *FDA Approval Summary: Olaparib Monotherapy in Patients with Deleterious Germline BRCA-Mutated Advanced Ovarian Cancer Treated with Three or More Lines of Chemotherapy.* Clin Cancer Res, 2015. **21**(19): p. 4257-61.
282. Mirza, M.R., et al., *Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer.* N Engl J Med, 2016. **375**(22): p. 2154-2164.
283. Swisher, E.M., et al., *Rucaparib in relapsed, platinum-sensitive high-grade ovarian carcinoma (ARIEL2 Part 1): an international, multicentre, open-label, phase 2 trial.* Lancet Oncol, 2017. **18**(1): p. 75-87.
284. Litton, J.K., et al., *A feasibility study of neoadjuvant talazoparib for operable breast cancer patients with a germline BRCA mutation demonstrates marked activity.* NPJ Breast Cancer, 2017. **3**: p. 49.
285. McLornan, D.P., A. List, and G.J. Mufti, *Applying synthetic lethality for the selective targeting of cancer.* N Engl J Med, 2014. **371**(18): p. 1725-35.
286. Druker, B.J., *Translation of the Philadelphia chromosome into therapy for CML.* Blood, 2008. **112**(13): p. 4808-17.
287. Diss, G., et al., *Molecular mechanisms of paralogous compensation and the robustness of cellular networks.* J Exp Zool B Mol Dev Evol, 2014. **322**(7): p. 488-99.
288. Brookfield, J.F., *Genetic redundancy.* Adv Genet, 1997. **36**: p. 137-55.
289. Pickett, F.B. and D.R. Meeks-Wagner, *Seeing double: appreciating genetic redundancy.* Plant Cell, 1995. **7**(9): p. 1347-56.
290. Fitch, W.M., *Homology a personal view on some of the problems.* Trends Genet, 2000. **16**(5): p. 227-31.
291. Kelso, T.W.R., et al., *Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers.* Elife, 2017. **6**.
292. Hoffman, G.R., et al., *Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers.* Proc Natl Acad Sci U S A, 2014. **111**(8): p. 3128-33.
293. Wilson, B.G., et al., *Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation.* Mol Cell Biol, 2014. **34**(6): p. 1136-44.
294. Muller, F.L., et al., *Passenger deletions generate therapeutic vulnerabilities in cancer.* Nature, 2012. **488**(7411): p. 337-42.
295. van der Lelij, P., et al., *Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts.* Elife, 2017. **6**.
296. Muller, F.L., E.A. Aquilanti, and R.A. DePinho, *Collateral Lethality: A new therapeutic strategy in oncology.* Trends Cancer, 2015. **1**(3): p. 161-173.
297. Hodges, C., J.G. Kirkland, and G.R. Crabtree, *The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer.* Cold Spring Harb Perspect Med, 2016. **6**(8).
298. Filippakopoulos, P. and S. Knapp, *Targeting bromodomains: epigenetic readers of lysine acetylation.* Nat Rev Drug Discov, 2014. **13**(5): p. 337-56.
299. Wee, S., et al., *Targeting epigenetic regulators for cancer therapy.* Ann N Y Acad Sci, 2014. **1309**: p. 30-6.
300. Vangamudi, B., et al., *The SMARCA2/4 ATPase Domain Surpasses the Bromodomain as a Drug Target in SWI/SNF-Mutant Cancers: Insights from*

cDNA Rescue and PFI-3 Inhibitor Studies. Cancer Res, 2015. **75**(18): p. 3865-3878.
301. Papillon, J.P.N., et al., *Discovery of Orally Active Inhibitors of Brahma Homolog (BRM)/SMARCA2 ATPase Activity for the Treatment of Brahma Related Gene 1 (BRG1)/SMARCA4-Mutant Cancers.* J Med Chem, 2018. **61**(22): p. 10155-10172.
302. Kadam, S. and B.M. Emerson, *Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes.* Mol Cell, 2003. **11**(2): p. 377-89.
303. Muchardt, C. and M. Yaniv, *ATP-dependent chromatin remodelling: SWI/SNF and Co. are on the job.* J Mol Biol, 1999. **293**(2): p. 187-98.
304. Phelan, M.L., et al., *Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits.* Mol Cell, 1999. **3**(2): p. 247-53.
305. Clapier, C.R., et al., *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes.* Nat Rev Mol Cell Biol, 2017. **18**(7): p. 407-422.
306. Shen, W., et al., *Solution structure of human Brg1 bromodomain and its specific binding to acetylated histone tails.* Biochemistry, 2007. **46**(8): p. 2100-10.
307. Shi, J., et al., *Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation.* Genes Dev, 2013. **27**(24): p. 2648-62.
308. Cermakova, K. and H.C. Hodges, *Next-Generation Drugs and Probes for Chromatin Biology: From Targeted Protein Degradation to Phase Separation.* Molecules, 2018. **23**(8).
309. Sakamoto, K.M., et al., *Development of Protacs to target cancer-promoting proteins for ubiquitination and degradation.* Mol Cell Proteomics, 2003. **2**(12): p. 1350-8.
310. Remillard, D., et al., *Degradation of the BAF Complex Factor BRD9 by Heterobifunctional Ligands.* Angew Chem Int Ed Engl, 2017. **56**(21): p. 5738-5743.
311. Hohmann, A.F., et al., *Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition.* Nat Chem Biol, 2016. **12**(9): p. 672-9.
312. Martin, L.J., et al., *Structure-Based Design of an in Vivo Active Selective BRD9 Inhibitor.* J Med Chem, 2016. **59**(10): p. 4462-75.
313. Filippakopoulos, P., et al., *Histone recognition and large-scale structural analysis of the human bromodomain family.* Cell, 2012. **149**(1): p. 214-31.
314. Streit, M., et al., *Ordino: a visual cancer analysis tool for ranking and exploring genes, cell lines, and tissue samples.* Bioinformatics, 2019.
315. Li, W., et al., *Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR.* Genome Biol, 2015. **16**: p. 281.
316. Li, W., et al., *MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens.* Genome Biol, 2014. **15**(12): p. 554.
317. Hormann, A., et al., *RIOK1 kinase activity is required for cell survival irrespective of MTAP status.* Oncotarget, 2018. **9**(47): p. 28625-28637.
318. Smyth, E.C., et al., *Oesophageal cancer.* Nat Rev Dis Primers, 2017. **3**: p. 17048.
319. Zhao, C., et al., *A phase II study of concurrent chemoradiotherapy and erlotinib for inoperable esophageal squamous cell carcinoma.* Oncotarget, 2016. **7**(35): p. 57310-57316.

320. Jia, J., et al., *The relation of EGFR expression by immunohistochemical staining and clinical response of combination treatment of nimotuzumab and chemotherapy in esophageal squamous cell carcinoma.* Clin Transl Oncol, 2016. **18**(6): p. 592-8.
321. Huang, J., et al., *Icotinib in Patients with Pretreated Advanced Esophageal Squamous Cell Carcinoma with EGFR Overexpression or EGFR Gene Amplification: A Single-Arm, Multicenter Phase 2 Study.* J Thorac Oncol, 2016. **11**(6): p. 910-7.
322. Bultman, S., et al., *A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes.* Mol Cell, 2000. **6**(6): p. 1287-95.
323. Chan-Penebre, E., et al., *Selective Killing of SMARCA2- and SMARCA4-deficient Small Cell Carcinoma of the Ovary, Hypercalcemic Type Cells by Inhibition of EZH2: In Vitro and In Vivo Preclinical Models.* Mol Cancer Ther, 2017. **16**(5): p. 850-860.
324. Wu, J., et al., *Inactivation of SMARCA2 by promoter hypermethylation drives lung cancer development.* Gene, 2019. **687**: p. 193-199.
325. Reyes, J.C., et al., *Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha).* EMBO J, 1998. **17**(23): p. 6979-91.
326. Schneppenheim, R., et al., *Germline nonsense mutation and somatic inactivation of SMARCA4/BRG1 in a family with rhabdoid tumor predisposition syndrome.* Am J Hum Genet, 2010. **86**(2): p. 279-84.
327. Glaros, S., et al., *Targeted knockout of BRG1 potentiates lung cancer development.* Cancer Res, 2008. **68**(10): p. 3689-96.
328. Buenrostro, J.D., et al., *Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position.* Nat Methods, 2013. **10**(12): p. 1213-8.
329. De Raedt, T., et al., *PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies.* Nature, 2014. **514**(7521): p. 247-51.

6 Appendix

6.1 List of abbreviations

abbreviation	explanation
18S	ribosomal RNA
ABL	v-abl Abelson murine leukemia viral oncogene homolog (tyrosine kinase)
ACC	adenoid cystic carcinoma
ACTB	actin beta
ALK	anaplastic lymphoma receptor tyrosine kinase
AML	acute myelogenous leukemia
ARID	AT-rich interaction domain
ASR	age standardized rate
AT/RT	atypical teratoid/rhabdoid tumor
ATAC-Seq	assay for transposase accessible chromatin
ATP	adenosine triphosphate
BAF	BRG1-associated factor complex
BCL11	BAF chromatin remodeling complex subunit 11
BCL7	BAF chromatin remodeling complex subunit 7
BCR	breakpoint cluster region (Ser/Thr kinase)
BD	bromodomain
BER	base excision repair
BET	bromodomain, extra-terminal domain
BRCA1/2	Breast Cancer 1/2
BRD4	bromodomain containing 4
BRD9	bromodomain containing 9
BRG1	Brahma-related gene 1
BRM	Brahma
Cas	CRISPR-associated
cBAF	canonical BAF (SWI/SNF)
CCND1	cyclin D1 (G1/S specific)
CDK	cyclin-dependent kinase 1
CDKN2A	cyclin dependent kinase inhibitor 2A
CDKN2B	cyclin dependent kinase inhibitor 2B
Ceres	CRISPR scores
CHD	chromatin remodeler, chromodomain-helicase-DNA binding
CHEK1	checkpoint kinase 1
CHEK2	checkpoint kinase 1
CML	chronic myeloid leukemia
CNA	copy number alteration
CRBN	cereblon
CREBBP	CREB binding protein
CRISPR	clustered regularly interspaced palindromic repeats
crRNA	CRISPR RNA

DBP	DNA binding protein
dBRD9	degrader of BRD9
DEXDc	DEAD-like helicases superfamily
DLC1	DLC1 Rho GTPase activating protein
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferases
dNTP	deoxynucleotide triphosphate
DSB	double-strand break
E.coli	Escherichia coli
EAC	esophageal adenocarcinoma
EC	esophageal carcinoma
EFS	EF-1 alpha short (promoter)
EGFR	epidermal growth factor receptor
EIN	intraepithelial neoplasia
EMEM	Eagle's minimum essential medium
ENO1	enolase 1
ENO2	enolase 2
EP300	E1A binding protein p300; histone acetyltransferase
ERRα	estrogen related receptor alpha
ESCC	esophageal squamous cell carcinoma
EZH2	enzymatic subunit of the Polycomb complex, PRC2
F12	ham's 12 nutrient mix
FACS	fluorescence activated cell sorting
FCS	fetal calf serum
FGFR1	fibroblast growth factor receptor
FHIT	fragile histidine triad
FSC-A	forward scatter area
FSC-H	forward scatter height
G9a	also EHMT2, euchromatic histone-lysine-N-methyltransferase 2
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GBM	glioblastoma
GFP	green fluorescence protein
GLTSCR1	glioma tumor suppressor candidate region gene 1
GLTSCR1L	glioma tumor suppressor candidate region gene 1 like protein
H&E	hematoxylin and eosin stain
H2A	canoncical histone
H2B	canoncical histone
H3	canonical histone
H3K27ac	histone 3 lysine 27 acetylation
H3K27me3	histone 3 lysine 27 trimethylation
H4	canonical histone
H4K79me2	histone 4 lysine 79 dimethylation
HAS	helicase-SANT-associated
HDAC	histone deacetylases
HDR	homologous directed repair

HGD	high-grade dysplasia
HGIEN	high-grade intraepithelial neoplasia
HNH	endocuclease domain of Cas9
HNSCC	head and neck squamous cell carcinoma
HPV	human papillomavirus
HR	homologous recombination
IAP	inhibitor of apoptosis proteins
IC50	half maximal inhibitory concentration
IEN	intraepithelial neoplasia
iMEF	immortalized mouse embryonic fibroblasts
INDEL	Insertion Deletion
INO80	chromatin remodeler
ISWI	chromatin remodeler
KD	knock-down
KDM6A	lysine demethylase 6a
KLF5	kruppel like factor 5
KO	knock-out
KRAS	kirsten rat sarcoma viral oncogene homolog, proto oncogene, GTPase
lfc	log2 fold changes
LGIEN	low-grade intraepithelial neoplasia
LOF	loss of function
LOH	loss of heterozygosity
LRG	leucine rich alpha-2-glycoprotein
LRP1B	LDL receptor related protein 1B
MDM2	mouse double minute 2 homolog
MLL2	KMT2D; lysine methyltransferase 2D
MMP	matrix metalloproteinases
MOI	multiplicity of infection
MRT	malignant rhabdoid tumor
MYC	protooncogene transcription factor
ncBAF	non-canonical BAF (SWI/SNF)
Neg. Contr.	negative control
NFE2L2	nuclear factor, erythroid 2 like 2
NHEJ	non-homologous end joining pathway
NKX2-1	NK2 homeobox 1
NOTCH1	notch receptor 1
NOTCH3	notch receptor 3
NSCLC	non small cell lung cancer
NTC	non-targeting control
N-term	Amino terminus
P1	passage one
p16 INK4a	CDKN2, MTS1, cyclin dependent kinase inhibitor
P8	passage eight
PAM	protospacer-adjacent motif
PARP	Poly-ADP-Ribose Polymerase
PARylation	protein poly ADP-ribosylation

PBAF	polybromo and BRG1 associated factor complex
PBRM1	polybromo1
PCNA	Proliferating Cell Nuclear Antigen
PCR	polymerase chain reaction
PD-1	programmed cell death protein 1
PD-L1	programmed cell death ligand 1
PIK3CA	phosphoinositide-3-kinase
POLR2A	RNA Polymerase II Subunit A
PRC1	polycomb repressive complex 1
PRC2	polycomb repressive complex 2
pre-crRNA	precursur crRNA
PRKCI	protein kinase C iota
PROTAC	proteolysis targeting chimera
PTPRD	protein tyrosine phosphatase receptor type D
QLQ	conserved Gln Leu Gln (glutamine, leucine, glutamine)
RASSF1A	ras association domain family member 1A
RB	retinoblastoma tumor suppressor protein
RNA	ribonucleic acid
RNAi	RNA interference
RPA3	Replication Protein A3
RSA	shRNA screen scores
RuvC	endocuclease domain of Cas9
SCC	squamous cell carcinoma
SCC-A	side scatter area
SCCOHT	small cell cancer of the ovary hypercalcemic type
SD	standard deviation
SF2	super family 2
sgRNA	single guide RNA
shRNA	short-hairpin-RNA-mediated
siRNA	small interfering RNA
SL	synthetic lethality
SMARCA2/4	SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily A member 2/4
SMARCA4-DTS	SMARCA4-deficient thoracic sarcomas
SMARCB	SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily B
SMARCC	SWI/SNF related matrix associated actin dependent regulator of chromatin subfamily C
SNF	Sucrose non-fermenting
SOX2	SRY-box 2
SS18	synovial sarcoma, member of the SWI/SNF complex
SSX	synovial sarcoma, x breakpoint
STAG1/2	stromal antigen 1/2
STR	short tandem repeats
SWI	swich
SWI/SNF	switch/ sucrose non-fermenting

SWR1	chromatin remodeling complex
TALEN	transcription activator-like effector nucleases
TCGA	the cancer genome atlas
TET2	tet methylcytosine dioxygenase 2
TFF1	trefoil factor 1
Tis	intraepithelial neoplasia
TNM	tumor, node, metastasis
TP53	tumor suppressor gene 53 (p53)
TP63	tumor protein p63
TPM	transcripts per million
tracrRNA	trans-activating crRNA
VHL	von Hippel-Lindau
ZFN	zinc finger nucleases
ZNF750	zinc finger protein 750
α-RRA	robust ranking aggregation

6.2 List of tables

6.3 List of figures

6.4 SMARCA2/4 transcription values and mutation status in selected cell lines

Cell name	Indication	Gene		Gene	
		SMARCA2		SMARCA4	
		[TPM]	mutation	[TPM]	mutation
COLO-680N_Cas9	ESCC	4,5	n.a.	46,6	p.?
KYSE-30_Cas9	ESCC	1,9	n.a.	64,5	p.D1175G
KYSE-70_Cas9	ESCC	38,4	n.a.	57,7	wt
KYSE-140_Cas9	ESCC	33,1	n.a.	108,3	p.D1235Y
KYSE-150_Cas9	ESCC	25,8	n.a.	67,2	p.N944K
KYSE-270_Cas9	ESCC	14,4	n.a.	116,6	p.P75fs;p.T494R
KYSE-410_Cas9	ESCC	26,4	n.a.	79,1	wt
KYSE-450_Cas9	ESCC	50,3	n.a.	74,8	wt
KYSE-510_Cas9	ESCC	2,4	n.a.	68,6	wt
T.T_Cas9	ESCC	39,4	n.a.	58,7	p.N1486H
HCT 116_Cas9	CRC	9,3	p.R855Q;p.?	77,5	p.L1163P;p.Y120fs
SK-CO-1_Cas9	CRC	1,2	wt	102,5	wt
OV-90_Cas9	OC	4,9	wt	51,7	wt
HuP-T4_Cas9	PC	1,1	n.a.	62,8	wt

Table 5: Cell lines and the respective SMARCA2 and SMARCA4 expression/ mutation status.

Expression and mutation status was obtained from the ordino platform (https://ordino.caleydoapp.org, [314]).

6.5 Epigenome library sgRNA sequences

Epigenome library sgRNAs	
Name	**Sequence**
ACAT1_e7_151.0	GAATATTGCACGAAATGAAC
ACAT1_e7_151.4	GTAAAGCAGCATGGGAAGCT
ACAT1_e7_151.6	TTCCCATGCTGCTTTACTTC
ACAT1_e8_96.0	TGGAAAACTGTCTTCAGCTT
ACAT1_e8_96.1	TCTTCTTTCACCACTACATC
ACTR_e18_255.23	TGCTTGGGGGTGTTTGTCCG
ACTR_e18_255.26	GTCTCATGATGTTGGCTCGT
ACTR_e19_105.4	CTGGTGGTGCTGCGGTGATG
ACTR_e20_258.1	GTCATCACTTCCGACAACAG
ACTR_e20_258.22	TTAGCAGCTCTCTGCTGCGT
ASH1L_BD_e18_55.0	TTTGGGGGAAGGTTCAAAAG
ASH1L_BD_e18_55.1	GTGGAGCTGCCAGTGCTTGC
ASH1L_BD_e19_133.3	CTGTCTTATAGTAACCAGTG
ASH1L_BD_e19_133.7	TCTATGGTGATAAGATCTAG
ASH1L_BD_e20_79.1	ATGGGCGTAAATCCCCAGTT
ASH1L_BD_e20_79.2	AGATGTTTGTCGTCTACGAA
ASH1L_e11_118.0	TCTAGAACGATTTCGAGCTG
ASH1L_e11_118.11	AGGGGAGGTCGTCAGTGAAC
ASH1L_e12_147.13	CAATCACCATCCCACTATCC
ASH1L_e12_147.5	GTGATTGACAGTTACCGCAT
ASH1L_e13_109.0	GTTAATGGAGTATACCGGAT
ASH1L_e13_109.1	GCTCTTAAAGACATGCCAGC
ATAD2_BD_e21_95.1	TAAGCCTGTTGACCCTGATG
ATAD2_BD_e21_95.2	CAACAGGCTTAGTAAACACT
ATAD2_BD_e22_197.4	GCTGTCAGTGATCTTGGCTC
ATAD2_BD_e22_197.8	GCTCAGGTGGTGGTGCTACT
ATAD2_BD_e23_86.1	TTCTTCTTTAATTATGGCAT
ATAD2_BD_e23_86.2	GGCAGTATCTCTTAAAGCAC
ATAD2B_BD_e21_95.5	TAACATCTTCAGCAAACCGG
ATAD2B_BD_e21_95.6	CAAACCGGTGGATATTGAAG
ATAD2B_BD_e22_160.4	TATAATCCAGATAAGGACCC
ATAD2B_BD_e22_160.8	ATTACTGTTGATAAGTCCAT
ATAD2B_BD_e23_86.1	GCACAGGGCTTGTACCCTGA
ATAD2B_BD_e23_86.2	TAGCATGTGCAGTGTCCTTC
ATAT1_e4_50.0	ATTATTGGTTTCATCAAAGT
ATAT1_e5_114.0	TGATCGTGAGGCTCATAATG
ATAT1_e5_114.2	CATGAGTCTGTGCAACGCCA
ATAT1_e5_114.5	ACTGGAAGAGTTCTCGCCCA
ATAT1_e6_102.0	GCGAGTGGAACCGCACCAAC
ATAT1_e6_102.1	CCTGAATAAGCACTACAATC
ATAT1_e6_102.6	AGGGTCGGTCAATTGCCAGT
ATF2_e11_150.10	ACCTTTGACAGTATCACCAT
ATF2_e11_150.4	CAAAGGTCATGGTAGCGGAT
ATF2_e11_150.5	GTCATGGTAGCGGATTGGTT
ATF2_e11_150.7	CGGAGTTTCTGTAGTGGATG
ATF2_e11_150.9	GCTGGCTGTTGTAATGACTG
ATRX_D_e18_147.2	GGATTTCAGCACGGCGTTAG
ATRX_D_e19_178.0	AGCAACTGTGAAACGTCCTC
ATRX_D_e19_178.6	ATGTATAGAAATCTTGCTCA
ATRX_D_e20_138.2	TGAATTCTATACGATCAAGG
ATRX_H_e26_107.0	CAGCCAGTCCCTCATATCTC
ATRX_H_e27_109.0	ATTGACTATTACCGTTTAGA
ATRX_H_e28_178.3	TAATTATATTCGACGCTTCT
BAZ1A_BD_e25_118.0	TGAACAACTTGTTGTAGAAT
BAZ1A_BD_e25_118.1	TGGTACGACATGATGACAGC
BAZ1A_BD_e25_118.4	ATTCATGAACTCCTCCCTGT
BAZ1A_BD_e26_88.1	ACGAATTATATTTAAGGCAA
BAZ1A_BD_e26_88.4	TGATGATGTCATAGTAGTCT
BAZ1A_BD_e27_135.3	CATATTCAGGCTCAAAAGCT
BAZ1A_BD_e27_135.4	TTTGCTTCACTTGTGTTACG
BAZ1A_BD_e27_135.5	TTGCTTCACTTGTGTTACGA
BAZ1B_BD_e18_67.0	TCGTGAAGTACCGCTTCAGC
BAZ1B_BD_e18_67.4	AGCGGTACTTCACGATCTTG
BAZ1B_BD_e18_67.5	GGTACTTCACGATCTTGTGG
BAZ1B_BD_e19_237.1	TGATGTGATCACGCACCCCA
BAZ1B_BD_e19_237.13	CACAGAGCGGTAGCTCCCAC
BAZ1B_BD_e19_237.6	GCTGAGGTTTACAACTGCCG
BAZ2A_BD_e28_144.16	TAGGATTTTTGATGATGCGC
BAZ2A_BD_e28_144.17	CGGTACCCACTCACCAAACG
BAZ2A_BD_e28_144.19	CTCACCAAACGTGGGTTCAC
BAZ2A_BD_e29_135.7	GCCGCTTCTTCGAGAGCCGC
BAZ2A_BD_e29_135.8	CCGCTTCTTCGAGAGCCGCT
BAZ2A_BD_e29_135.9	CTTCTTCGAGAGCCGCTGGG
BAZ2B_BD_e36_144.1	TGGAAACTCATGAGGATGCA
BAZ2B_BD_e36_144.2	GTAAACTTGAAACTTGTTCC
BAZ2B_BD_e36_144.8	GGAACAAGTTTCAAGTTTAC
BAZ2B_BD_e37_139.0	AAACCTTTGCTCTAGATGTC
BAZ2B_BD_e37_139.2	GATTCTGATATAGGCAGAGC
BAZ2B_BD_e37_139.3	GCAGAGCTGGCCACAATATG
BPTF_BD_e23_85.10	GGAAAGGCCAGGCCATCTTA
BPTF_BD_e23_85.11	GAAAGGCCAGGCCATCTTAT
BPTF_BD_e24_187.1	TGAAAAGCTGACGGAATTTG
BPTF_BD_e24_187.3	ATTCGAGAACTTCTGCACAC
BPTF_BD_e24_187.7	TGGTAAAATGGGGAGTCACT

BRD1_BD_e5_95.10	GCTGCGCAAATATCCTGGCG
BRD1_BD_e5_95.7	CACGGGCTGCGCAAATATCC
BRD1_BD_e5_95.9	GGCTGCGCAAATATCCTGGC
BRD1_BD_e6_237.10	TGTTCTATAGAGCCGCGGTG
BRD1_BD_e6_237.11	CGCGGTGAGGCTGCGCGATC
BRD1_BD_e6_237.9	CACCGTGTTCTATAGAGCCG
BRD2_BD1_e2_109.12	CCTTGTGTAGGTATTGCAGC
BRD2_BD1_e2_109.3	TGTGGAAACATCAGTTCGCA
BRD2_BD1_e2_109.4	ATCAGTTCGCATGGCCATTC
BRD2_BD1_e3_138.2	ATAAAACAGCCTATGGACAT
BRD2_BD1_e3_138.3	TGGACATGGGTACTATTAAG
BRD2_BD1_e3_138.8	TTAATAGTACCCATGTCCAT
BRD2_BD1_e4_60.1	TGTCCTAATGGCACAAACGC
BRD2_BD1_e4_60.3	TTTCCAGCGTTTGTGCCATT
BRD2_BD2_e6_145.6	TGGGGTGCTTAATGATGTCA
BRD2_BD2_e6_145.7	TGATGTCATGGTAGTCATGC
BRD2_BD2_e7_129.0	AGATGGAGAACCGTGATTAC
BRD2_BD2_e7_129.6	ATTGCCACAACATCGTGATC
BRD2_BD2_e7_129.7	TTGCCACAACATCGTGATCT
BRD3_BD1_e2_112.5	GTTCAATTTGATTGCGTCCA
BRD3_BD1_e2_112.6	TTCAATTTGATTGCGTCCAC
BRD3_BD1_e2_112.7	ATTTGATTGCGTCCACGGGC
BRD3_BD1_e3_138.5	GAGTGCAAGCGAATGTATGC
BRD3_BD1_e3_138.7	TTAATAGTCCCCATATCCAT
BRD3_BD1_e3_138.8	TAATAGTCCCCATATCCATT
BRD3_BD1_e4_66.0	CACAGATGACATAGTGCTAA
BRD3_BD1_e4_66.4	CATTAGCACTATGTCATCTG
BRD3_BD1_e4_66.5	ATTAGCACTATGTCATCTGT
BRD3_BD2_e6_145.5	CGACATCATCAAGCACCCGA
BRD3_BD2_e7_129.0	CCGAGAGTACCCAGACGCAC
BRD3_BD2_e7_129.3	ATACAATCCCCCAGACCACG
BRD3_BD2_e7_129.8	ATGGCCACAACCTCGTGGTC
BRD4_BD1_e2_121.1	GAGTGGTGCTCAAGACACTA
BRD4_BD1_e2_121.4	TTCAGCTTGACGGCATCCAC
BRD4_BD1_e2_121.9	CTCTGAGCAGGTATTGCAGT
BRD4_BD1_e3_138.0	TAAGATCATTAAAACGCCTA
BRD4_BD1_e3_138.2	ATTAAAACGCCTATGGATAT
BRD4_BD1_e3_138.3	GGGAACAATAAAGAAGCGCT
BRD4_BD1_e4_68.0	TGGAGATGACATAGTCTTAA
BRD4_BD1_e4_68.1	AGTCTTAATGGCAGAAGCTC
BRD4_BD1_e4_68.2	ATTAAGACTATGTCATCTCC
BRD4_BD2_e6_148.11	TGATGTCACAGTAGTCGTGT
BRD4_BD2_e7_130.0	CCGTGAGTACCGTGATGCTC
BRD4_BD2_e7_130.13	CCTGAGCATCACGGTACTCA
BRD4_BD2_e7_130.14	CTGAGCATCACGGTACTCAC

BRD8_BD_e17_77.2	AGCTCTCCATACAAGCATGA
BRD8_BD_e18_69.0	GTTACAGATGACATAGCACC
BRD8_BD_e18_69.2	GGTGCTATGTCATCTGTAAC
BRD8_BD_e18_69.3	TGTCATCTGTAACAGGCTGC
BRD8_BD_e19_178.3	GGAGATGCAGCGAGATGTCT
BRD8_BD_e19_178.4	CTGCCATGTGATAGACATCA
BRD9_BD_e5_145.1	GTCACGGATGCAATTGCTCC
BRD9_BD_e5_145.7	TTCATGGTGCCAAAATCCAT
BRD9_BD_e5_145.9	GGAGCAATTGCATCCGTGAC
BRD9_BD_e6_111.2	GCGAAGAAGATCCTTCACGC
BRD9_BD_e6_111.4	CTTCGCCAACTTGTAGTACA
BRD9_BD_e6_111.5	AACTTGTAGTACACGGTATC
BRDT_BD1_e2_121.3	ATGGCCCTTTCAACGTCCTG
BRDT_BD1_e2_121.5	AGCATCCACAGGACGTTGAA
BRDT_BD1_e2_121.6	GCATCCACAGGACGTTGAAA
BRDT_BD1_e3_138.1	AAATACAATTAAGAAGCGCT
BRDT_BD1_e3_138.2	GGAGAATAAATATTATGCGA
BRDT_BD1_e4_68.0	TGGAGATGACATTGTTCTTA
BRDT_BD1_e4_68.1	ATAAGAACAATGTCATCTCC
BRDT_BD2_e6_150.1	TCCTGTTGACGTTAATGCTT
BRDT_BD2_e6_150.6	TGACAACGTCATAGTAGTTA
BRDT_BD2_e6_150.7	CCCAAAGCATTAACGTCAAC
BRDT_BD2_e7_129.2	AGATCACGAAGTTGTGACAA
BRDT_BD2_e7_129.3	ATTGTCACAACTTCGTGATC
BRDT_BD2_e7_129.4	GTCACAACTTCGTGATCTGG
BRPF1_BD_e6_118.11	GCTGCTCCAAGGTTTTGCGA
BRPF1_BD_e6_118.12	GCTCCAAGGTTTTGCGAAGG
BRPF1_BD_e6_118.5	TCTGTCTGAGGTAACCGAAT
BRPF1_BD_e7_237.10	GCAGTGCGGCTTCGTGAGCA
BRPF1_BD_e7_237.5	CTGCCTCAAGTATAACGCCA
BRPF1_BD_e7_237.9	AGCAGTGCGGCTTCGTGAGC
BRPF3_BD_e5_95.0	TGTTCTGTTGAGGACAACAC
BRPF3_BD_e5_95.3	AGAACCAGTCAACTTGAGTG
BRPF3_BD_e6_190.23	ACTCCTCCAAGGTGCGGTAC
BRPF3_BD_e6_190.25	CTCCAAGGTGCGGTACAGGT
BRPF3_BD_e6_190.4	GTCCCACCTGTACCGCACCT
BRWD1_BD1_e30_74.0	GAATGTGATAGAATTATCAG
BRWD1_BD1_e31_121.0	CATTCGAATTGTGTAAAGAT
BRWD1_BD1_e31_121.1	CGAATTGTGTAAAGATCGGT
BRWD1_BD1_e31_121.3	CTACTACAGTACAGTACTTC
BRWD1_BD1_e32_126.0	GAGGCTGTCTGCGTTAGTTT
BRWD1_BD1_e32_126.1	GATCTTGCAATTACACTCTC
BRWD1_BD1_e32_126.2	CTCAGGTTCGTTAAATGTTC
BRWD1_BD2_e35_98.1	ATTTAGACAACCTGTTGATT
BRWD1_BD2_e35_98.2	GGATATTCAACCAAATCAAC

BRWD1_BD2_e35_98.3	AAATCAACAGGTTGTCTAAA
BRWD1_BD2_e36_153.0	AGATATTATAGATACCCCAA
BRWD1_BD2_e36_153.1	ATAGATACCCCAATGGATTT
BRWD1_BD2_e36_153.5	GTAAGGGAAACTCTAGATGC
BRWD3_BD_e30_95.10	CCGTTCACATTCTTCGTCTC
BRWD3_BD_e30_95.8	AGAATGTGAACGGGTTATTC
BRWD3_BD_e30_95.9	GAATGTGAACGGGTTATTCA
BRWD3_BD_e31_121.1	CAACTGACCTCAATACCATC
BRWD3_BD_e31_121.3	ATTTTCAAGTCTCCGCCTGA
BRWD3_BD_e31_121.4	GTCTCCGCCTGATGGTATTG
BRWD3_BD_e32_126.2	GCTATATTGAACATAATGCC
BRWD3_BD_e32_126.3	GATGTCTTACTTCGATTTAT
BRWD3_BD_e32_126.5	GCTTTAACTATAGGACTGTC
BRWD3_BD2_e35_105.1	CATTTATGAACGTGAAGACT
BRWD3_BD2_e35_105.4	AGATCAGCTGGCTGTCGAAA
BRWD3_BD2_e35_105.5	AGTCTTCACGTTCATAAATG
BRWD3_BD2_e36_75.0	GATGAATCTTGTTGTCTTTC
BRWD3_BD2_e37_153.0	AGATGTTATAGATACTCCTG
BRWD3_BD2_e37_153.10	TTCACAGTGCTGAAGTCCAC
BRWD3_BD2_e37_153.4	AGGAAACTATGGTAGTCCTC
BRWD3_BD2_e38_95.0	CAGCAATCCAGAGTCAGAAG
CARM1_e5_111.13	CTTGGGCGGCAAAAAACGAC
CARM1_e5_111.6	AGCACGGAAAATCTACGCGG
CARM1_e5_111.7	ACGGAAAATCTACGCGGTGG
CARM1_e6_178.1	CAACCTGACGGACCGCATCG
CARM1_e6_178.13	GCTCTTCAACGAGCGCATGC
CARM1_e6_178.25	TGACCACGATGCGGTCCGTC
CECR2_BD_e12_92.0	AGACGTGGTAAAGGCTCACA
CECR2_BD_e12_92.1	TAAAGGCTCACAAGGATTCC
CECR2_BD_e12_92.7	GGGGCATAAGATTCATCCAC
CECR2_BD_e13_127.3	TGGAGGTTTATACTGTACCA
CECR2_BD_e13_127.7	GGAATTGTCGAAAGTATAAT
CECR2_BD_e13_127.9	CTTTCGACAATTCCTGAACA
CECR2_BD_e14_59.0	AGATGTCTGATAATTTAGAG
CECR2_BD_e14_59.2	TTTAGAGAGGTGTTTCCATC
CECR2_BD_e14_59.3	CTCTAAATTATCAGACATCT
CHD1_D_e11_210.0	TTGCATACTCGCTGATGAAA
CHD1_D_e12_90.0	GGACGCATCATCAGACCAAA
CHD1_D_e13_135.3	AGTGATAAGGAGACGATGAT
CHD1_H_e17_72.2	TCTAGATCATTTTAATGCTG
CHD1_H_e18_150.7	CAGGCTAGAGCCCATCGAAT
CHD1_H_e18_150.8	AGGCTAGAGCCCATCGAATT
CHD1L_D_e4_115.1	TGTAACATATGCAGGCGACA
CHD1L_D_e4_115.7	GTCGCCTGCATATGTTACAC
CHD1L_D_e6_82.1	TTGTTGTGGATGAAGCTCAC

CHD1L_H_e10_50.2	ATGCTAGTAGCTTATCCAGC
CHD1L_H_e11_74.2	CTCTGTAATCCATATAGTCT
CHD1L_H_e12_111.0	TACAGCTATGAGCGTGTGGA
CHD1L_H_e13_115.2	GCTGCCAGGGCTCATCGCAT
CHD2_D_e14_210.7	GGCACCAGAGATTAACGTAG
CHD2_D_e15_90.0	TGCGTTGAACTTCAATCTTT
CHD2_D_e16_135.13	GTCATCATTCTTCAACCGAT
CHD2_H_e19_153.1	GTTGACAAGACTTCGAGAAA
CHD2_H_e20_72.0	CGTCTGGATGGTTCCATCAA
CHD2_H_e21_150.10	CAAGCCCGAGCGCATAGAAT
CHD3_D_e14_79.5	CCTCTACTCACTCTACAAGG
CHD3_D_e15_201.10	GGTGACATACACGGGTGACA
CHD3_D_e16_138.12	TGATCAGCTCATACGATGTC
CHD3_D_e17_63.0	GATCATAAGTTGCTGCTGAC
CHD3_H_e20_86.3	TCAGCTTTCGCAGCATCTTC
CHD3_H_e21_118.4	GCGCATCGATGGTGGTATCA
CHD3_H_e22_125.0	TTCCTCCTGTCCACCCGAGC
CHD3_H_e23_50.0	TTTAGCCGGGCTCATCGGAT
CHD4_D_e16_201.15	GAACTCATTCTCTCGGATGA
CHD4_D_e17_138.8	CTGATTGTTCTTCAGCCGAT
CHD4_D_e18_55.3	GTGAGTAACCATTCAATACC
CHD4_H_e22_118.4	AAATACGAACGCATCGATGG
CHD4_H_e23_125.1	TCTTGCTTTCCACTCGAGCT
CHD4_H_e24_50.1	TTAGCAGAGCTCACCGGATT
CHD5_D_e15_201.32	AGGGGCGCGCTAACCAGGTA
CHD5_D_e16_138.10	CTGGTTGTTCTTGAGGCGGT
CHD5_D_e17_59.0	GATTACAAGCTGCTGCTGAC
CHD5_H_e20_95.3	CTGAAGAAACTGCGGGATGA
CHD5_H_e21_118.17	ACTTGTAGCCTTCGTACTCC
CHD5_H_e22_125.10	TCATCATCTACGACTCGGAC
CHD5_H_e23_50.4	TCTTGTTCTGGCCGATGCGG
CHD6_D_e12_244.11	GGGAGCGGGAGTTCCGGACA
CHD6_D_e13_177.11	CAAATGTTGTGATGACGACG
CHD6_H_e16_105.4	GGTAATCTTCTAGGATGTCG
CHD6_H_e17_196.20	TGGCTTACAGAACCGGTCGA
CHD6_H_e18_47.3	TCTGGCCTATGCGGTGACAT
CHD7_D_e16_118.9	GTCCAAGCAGCGCACCATCT
CHD7_D_e17_196.1	AGGATCGACGGCCGAGTAAG
CHD7_H_e12_244.14	ATGGGAGTCAAGCTAGTCGT
CHD7_H_e13_177.11	CATGGAATATTCCGCAGCTC
CHD8_D_e12_244.18	TCGCTCCCAGTTAGTAATTG
CHD8_D_e13_177.7	GCGTCAAACTTGTATGCGCC
CHD8_H_e16_135.0	GGTTCGTTCAGCCGGCAAAC
CHD8_H_e17_196.3	AGTTAGAGGCAACCTTCGAC
CHD8_H_e18_52.3	CAATTCGATGACATCGTGCC

CHD9_D_e12_244.8	GTGGACTGATATTAACGTTG
CHD9_D_e13_177.3	GAGAGCTTAATGCAATTGAA
CHD9_H_e16_135.3	GGTTCGTTGCCTTGACATTC
CHD9_H_e17_196.11	ACCCACCAGCTCGGGTACAC
CHD9_H_e18_50.2	CAATTCTGTGGCAACGAGCT
CLOCK_e21_210.10	TCTGGTTAGTAGGAACAACT
CLOCK_e21_210.12	TATTTATAGGTGCAAGTTGC
CLOCK_e21_210.3	GGTTCCGTTCAACTTTCTTC
CLOCK_e22_203.13	TGGACCATGCTTCCGGCTGC
CLOCK_e22_203.2	AGCCCCACTGTATAACACTA
CLOCK_e22_203.5	TGCAGTAACTACATTCACTC
CLOCK_e23_256.10	GTAAAAATTGTTGCGGTGGC
CLOCK_e23_256.17	TGGCCCATAAGCATAGTACT
CLOCK_e23_256.6	TAGTACTATGCTTATGGGCC
CREBBP_BD_e17_107.0	CCTAGAAGCACTGTATCGAC
CREBBP_BD_e17_107.5	GGATCTACAGGCTGCCGGAA
CREBBP_BD_e17_107.6	TGCCGGAAAGGTAATGACTC
CREBBP_BD_e18_210.14	GACGTCGTCCACGTACTGCC
CREBBP_BD_e18_210.15	ACGTCGTCCACGTACTGCCA
CREBBP_BD_e18_210.7	GGCAGTACGTGGACGACGTC
CREBBP_e25_147.2	TTTTGAGGAAATTGACGGCG
CREBBP_e25_147.7	AACAGAGCTTTGGTTCGATA
CREBBP_e25_147.8	ACAGAGCTTTGGTTCGATAT
CREBBP_e26_114.7	TCTCATGGTAAACGGCTGTG
CREBBP_e26_114.8	CATGGTAAACGGCTGTGCGG
CREBBP_e26_114.9	ACGGCTGTGCGGAGGCAACG
CREBBP_e27_166.2	ACCCAAGCCAAAACGACTGC
CREBBP_e27_166.7	TACCACTCCTGCAGTCGTTT
CREBBP_e27_166.8	CTCCTGCAGTCGTTTTGGCT
CREBBP_e28_168.12	TTCAAAATAGGGCAGTTCCT
CREBBP_e28_168.2	AAGGAACTGCCCTATTTTGA
CREBBP_e28_168.9	ATGCTCTCTTCTAACACATT
CREBBP_e29_162.10	TTGGACACGTTGGGCATGCT
CREBBP_e29_162.7	GGACAGGTCATTGGACACGT
CREBBP_e29_162.8	GACAGGTCATTGGACACGTT
DOT1L_e6_95.2	CGTCGAGAAAGCAGACATCC
DOT1L_e6_95.3	AGACATCCCGGCCAAGTATG
DOT1L_e6_95.7	CAGCAACCTGGAGCACGACC
DOT1L_e7_63.0	TGGACCGCGAGTTCAGGAAG
DOT1L_e7_63.3	TCATCCACTTCCTGAACTCG
DOT1L_e7_63.4	CTTCCTGAACTCGCGGTCCA
DOT1L_e8_56.0	GCGATTTCCTCTCAGAAGAG
DOT1L_e8_56.1	ATTTCCTCTCAGAAGAGTGG
DOT1L_e8_56.3	GCTCCCTCCACTCTTCTGAG
DOT1L_e9_80.5	GGAGCGGTTTGCAAACATGA

DOT1L_e9_80.7	TTCAGCTGGTGATCCACCTC
DOT1L_e9_80.8	GTGATCCACCTCAGGACCAA
EHMT1_e24_87.1	GCTGCAGCTCTACCGGACGC
EHMT1_e24_87.2	GCTCTACCGGACGCGGGACA
EHMT1_e24_87.3	CTCTACCGGACGCGGGACAT
EHMT1_e25_79.0	CTCAGAAGCCGACGTTCGAG
EHMT1_e25_79.1	AGAATCTTCCTCTCGAACGT
EHMT1_e26_176.0	AGGTTTACTGCATCGACGCG
EHMT1_e26_176.1	TGCATCGACGCGCGGTTCTA
EHMT1_e26_176.2	GCATCGACGCGCGGTTCTAC
EHMT2_e23_87.1	CTCTACCGAACAGCCAAGAT
EHMT2_e23_87.10	TCGCAGATGAAGGTCCCCTG
EHMT2_e23_87.11	CGCAGATGAAGGTCCCCTGT
EHMT2_e24_79.0	GGAGCTGATCTCTGATGCTG
EHMT2_e25_176.13	AGGCGATGCGTGGAAATCGC
EHMT2_e25_176.4	AGTTCCCGAGACATCCGGAC
EHMT2_e25_176.5	GTTCCCGAGACATCCGGACT
ELP3_e13_81.0	GTACGAGAGCTGCATGTGTA
ELP3_e13_81.1	TACGAGAGCTGCATGTGTAT
ELP3_e13_81.8	CACATGCAGCTCTCGTACTA
ELP3_e14_82.0	ATTTGGCATGCTGCTGATGG
ELP3_e14_82.4	CTAGAGAAGAACATGGGTCT
ELP3_e15_58.1	AAGATCGGCTACAGATTACA
ELP3_e15_58.2	GATCTTTCTATAATAATTCC
EP300_BD_e17_111.2	TTTGGAGGCACTTTACCGTC
EP300_BD_e17_111.6	AGGGTCCACAGGTTGACGAA
EP300_BD_e17_111.7	GGGTCCACAGGTTGACGAAA
EP300_BD_e18_196.14	AATATCATCGACATACTGCC
EP300_BD_e18_196.15	ATATCATCGACATACTGCCA
EP300_BD_e18_196.5	GGCAGTATGTCGATGATATT
EP300_e25_147.10	AAGAGGGCTTTGGTTCGGTA
EP300_e25_147.11	GGCTTTGGTTCGGTATGGAA
EP300_e25_147.3	GGCATGCATGTTCAAGAGTA
EP300_e26_114.0	TCTTCCGTCCTAAATGCTTG
EP300_e26_114.5	CAGTCCTCAAGCATTTAGGA
EP300_e26_114.6	GGAAGAAATGAACACTATCG
EP300_e27_166.3	ACCCAAGCCCAAGCGACTGC
EP300_e27_166.6	GTACCATTCCTGCAGTCGCT
EP300_e27_166.7	TACCATTCCTGCAGTCGCTT
EP300_e28_165.12	CAGAAATCACCCTCGAAATA
EP300_e28_165.2	AAGGAATTGCCTTATTTCGA
EP300_e28_165.4	TGATTTCTGGCCCAATGTTC
EP300_e29_162.10	GTTAGATACATTGGGCATCC
EP300_e29_162.8	TGAGAGGTCGTTAGATACAT
EP300_e29_162.9	GAGAGGTCGTTAGATACATT

EP400_D_e14_86.4	GCCCACCTAGCTTGTAACGA
EP400_D_e15_137.3	GAATTGAAACGTTGGTGTCC
EP400_D_e16_170.3	CCGCCTTCACACGAGTGCGC
EP400_H_e28_170.6	CATCGATTCTTACATAGGTG
EP400_H_e29_190.15	CGACGGTGTCCGCCTCTACA
ERCC6_D_e7_105.7	TACAGCAAGATCAGGACTCG
ERCC6_D_e8_136.9	AGAATTGCCACTCTGAACGG
ERCC6_D_e9_171.1	CTCCTACATTCGATTGATGC
ERCC6_H_e13_105.0	TTGGGTACTGGAAACGTTCT
ERCC6_H_e14_111.2	TCATTGTATCTCGTAATCAG
ERCC6_H_e15_120.3	TTTCTTCTGACCACGCGGGT
ERCC6_H_e16_52.2	CTATTCTCCATGCTCGCTCC
EZH1_e17_96.2	CCCCCTCTGATGTGGCCGGA
EZH1_e17_96.3	CCCCTCTGATGTGGCCGGAT
EZH1_e17_96.5	CGGATGGGGCACCTTCATAA
EZH1_e18_82.0	CAGGATGAGGCTGATCGACG
EZH1_e18_82.1	TGAGGCTGATCGACGCGGAA
EZH1_e19_81.0	ATTTTGTAGTGGATGCTACT
EZH1_e19_81.1	GTAGTGGATGCTACTCGGAA
EZH1_e20_85.1	GTGAATGGAGACCATCGGAT
EZH1_e20_85.2	TGAATGGAGACCATCGGATT
EZH1_e20_85.7	TGGCAAAGATCCCAATCCGA
EZH2_e16_96.1	CTGGCACCATCTGACGTGGC
EZH2_e16_96.2	CACCATCTGACGTGGCAGGC
EZH2_e16_96.8	CCCCAGCCTGCCACGTCAGA
EZH2_e19_85.3	AAAACAGCTCTTCGCCAGTC
GCN5_e11_127.10	TGCTGTCACCTCGAATGAGC
GCN5_e11_127.16	GACCCGCCCATCCTTGATCA
GCN5_e11_127.5	ATCAAGGATGGGCGGGTCAT
GCN5_e12_59.0	CCACCTGATGAACCACCTGA
GCN5_e12_59.1	GCTTGATGTGATACTCCTTC
GCN5_e12_59.4	CCTTCAGGTGGTTCATCAGG
GTF3C4_e3_131.0	GCCATTCACAGATCGCAAAC
GTF3C4_e3_131.3	GCCTGTTTGCGATCTGTGAA
GTF3C4_e3_131.4	TTACACAAACTACAGTGCTC
GTF3C4_e4_89.2	GATGCCGGGCAATGCTGTCA
GTF3C4_e4_89.5	ATATCAAACTCTGGCAGGAC
GTF3C4_e4_89.6	CAAACTCTGGCAGGACTGGT
HAT1_e3_76.0	GAGTATACCCATCAACTCTT
HAT1_e3_76.1	AGTATACCCATCAACTCTTT
HAT1_e3_76.3	AAGAGTTGATGGGTATACTC
HAT1_e4_121.2	ATCCTGTTATACTATATTGC
HAT1_e4_121.4	CAACACGGAACATTGTTGAC
HAT1_e4_121.5	TACCAGCAATATAGTATAAC
HAT1_e5_180.0	ATTAGACAAATCATTCCACC

HAT1_e5_180.4	TACTCAGTTCTCAGTCCAAC
HAT1_e5_180.9	GTATGAAGTAAGGTTCCGAA
HAT1_e6_73.0	GCTTTCGAGAATATCATGAA
HAT1_e6_73.1	GGCTTCAGACCTTTTTGATG
HBO1_e13_147.4	TTATAAGCTATCGCAGTTAC
HBO1_e14_107.2	TCAGATGCTCAAATACTGGA
HBO1_e14_107.4	TCCAGTATTTGAGCATCTGA
HBO1_e9_192.3	GCCGCTATGAGCTTGATACC
HBO1_e9_192.8	TGTGCCGGCGGAGTATCGTT
HDAC1_e2_113.6	TGAGTCATGCGGATTCGGTG
HDAC1_e3_118.4	CGACATGTTATCTGGACGGA
HDAC1_e4_75.0	GAGGACTGTCCAGTATTCGA
HDAC1_e5_139.9	TTACGTCAATGATATCGTCT
HDAC1_e6_142.2	TATTCACCATGGTGACGGCG
HDAC1_e7_93.0	GTTAACTACCCGCTCCGAGA
HDAC1_e8_109.7	TCCCTATCTGGGGATCGGTT
HDAC1_e9_140.6	TTCGTAACGTTGCCCGGTGC
HDAC10_e10_90.12	GTGGGCGAAGCACTCTGGCG
HDAC10_e11_107.3	ATGACAGTACAGACGCTGCT
HDAC10_e5_105.11	ACCCGTTGGCAGCCGCCCTC
HDAC10_e6_69.8	GCACATCCCAGTCCACGACG
HDAC10_e7_127.17	CCCATGCTCATAGCGGTGCC
HDAC10_e8_66.0	GATGGGAAACGCTGACTACG
HDAC10_e9_60.4	AGGATTTGACTCAGCCATCG
HDAC11_e3_101.12	GCCTCGTGTGCACCACCAGC
HDAC11_e4_117.8	TCCTCCTGTCTGGGTCCGAA
HDAC11_e5_43.3	GCGAGGCTGGGCCATCAACG
HDAC11_e6_77.6	GGCGAGCGTGATGTCCGCAT
HDAC11_e7_63.4	AAGATCAATGATGGTAGCCC
HDAC11_e8_97.1	CGACAAGCGTGTGTACATCA
HDAC11_e9_118.3	TCCGGATGGTCCGTGGCCGC
HDAC2_e5_139.3	GATATGGCTGTTAATTGGGC
HDAC2_e6_140.2	TACAACAGATCGTGTAATGA
HDAC2_e7_93.1	GGTATAGATGATGAGTCATA
HDAC2_e8_109.3	CTCATTATCTGGTGATAGAC
HDAC2_e9_140.4	TCCGTAATGTTGCTCGATGT
HDAC3_e10_65.1	TCTCTGGGCTGTGATCGATT
HDAC3_e11_90.1	AATATCCCTCTACTCGTGCT
HDAC3_e6_56.1	CGACATTGTGATTGGCATCC
HDAC3_e7_134.16	TGTCAATGTAGAGCACCCGA
HDAC3_e8_81.4	GGTCATCAATGCCATCCCGC
HDAC3_e9_74.6	CTACCTGGTTGATAACCGGC
HDAC4_e19_56.3	TGGGCGTGCTCTCCTCCGCA
HDAC4_e20_88.1	CTACTTCAACTCCGTGGCCG
HDAC4_e21_120.2	CCCTCCACCGCTACGACGAT

HDAC4_e22_98.4	CATGGCTTTCACCGGCGGCC
HDAC4_e23_119.4	GTCATCAGGCTTCGATGCCG
HDAC5_e20_56.2	TGGCTGTGGATTCCTCGGCG
HDAC5_e21_88.3	CGTGGGCAAGGTCCTCATCG
HDAC5_e22_120.11	GGTCATTGTAGAACGCCTGC
HDAC5_e23_97.7	GGACCCCCCCATTGGAGACG
HDAC5_e24_119.10	CATCAAACCCGGCGGAGACT
HDAC5_e25_134.18	CACAGATGGCGGTCAAGTCA
HDAC6_a1_e10_69.4	GCACATCCCAATCTACGATA
HDAC6_a1_e11_127.22	ACCTACCCTGCTCGTAGCGG
HDAC6_a1_e12_66.0	CGACTGGCAGCAGGACGTGC
HDAC6_a1_e13_60.3	CTGGATTTGATGCCCTGCAA
HDAC6_a1_e14_90.1	AGATGGCCGCCACTCCGGCA
HDAC6_a1_e8_98.7	AGGCTGGTGGATGCGGTCCT
HDAC6_a1_e9_105.4	CTCAACAGAAACACCGCATC
HDAC6_a2_e20_134.7	GGGCATGCCCACTGATAGTC
HDAC6_a2_e21_69.4	GCTGAGTTCCATTACCGTGG
HDAC6_a2_e22_193.13	CACAGGCTTCACCGTCAACG
HDAC6_a2_e23_150.7	GCTGCACGGGGGGATCCGCT
HDAC6_a2_e24_175.10	AGCTGCGCCAGTATCTGCGA
HDAC7_e14_134.17	TGGATGCGGCCGGCGTGCTC
HDAC7_e15_121.13	TGAGCGGGTTGGTGCCGTAG
HDAC7_e16_50.6	TCACAAACATCCGCTGTGCC
HDAC7_e17_108.0	ATTCCTCCAATGCAGCCCGC
HDAC7_e18_56.0	TTCGCTGTGGTGCGGCCCCC
HDAC7_e19_88.1	ACTCAGTGGCCATCGCCTGC
HDAC7_e20_120.12	GCCGTCGTCATGGCGATGCA
HDAC7_e21_98.10	AGGTCTGGACCCCCCCATGG
HDAC7_e22_119.18	TGGAGAGAACTCTCGGGCGA
HDAC8_e3_131.7	CATCCGGACTCCATAGAATA
HDAC8_e4_142.12	CTTTGCACATTCCGTCAATC
HDAC8_e5_113.2	TGGGAATATTACGATTGCGA
HDAC8_e6_78.4	CAGGGACACGGTCATGACTT
HDAC8_e7_109.1	GACGTGTCTGATGTTGGCCT
HDAC8_e8_173.14	GTGTTGCCAACTGCCATTGA
HDAC9_e17_56.1	CTGTGGATTCTTCAGCGTGA
HDAC9_e19_120.1	TCACTCCATCGCTATGATGA
HDAC9_e19_120.7	AGTTCCCTTCATCATAGCGA
HDAC9_e20_98.7	AAGGTACTCAACATCTCCCA
HDAC9_e21_119.13	GCAGATACTAAGACCATGTC
HDAC9_e22_134.4	CGTGTGGTGTTGGCTCTAGA
HELLS_D_e10_144.9	TCAAATGACGTGATTACCAC
HELLS_D_e11_135.2	ATATGAAGTGCCGTCTAATC
HELLS_D_e9_190.7	GTCTACACTTCCTAACTGGA
HELLS_H_e18_120.2	TCAGAGATTTCAACTTCAGC
HELLS_H_e19_117.1	TTCTTAGTGAGTACACGAGC
HELLS_H_e20_73.0	GAACCCCCAGTCGGATCTTC
HIF1AN_e2_251.15	TGGACCTCGGCTTAAAGTTC
HIF1AN_e2_251.26	TTTGTGTCGGTCAGCACCAC
HIF1AN_e2_251.6	ACTTTAAGCCGAGGTCCAAC
HIF1AN_e3_149.1	CAAACGCTCAATGACACTGT
HIF1AN_e4_146.7	TAGAGGCACTCGAACTGATC
HIF1AN_e4_146.8	AGGCACTCGAACTGATCCGG
HIF1AN_e4_146.9	GGCACTCGAACTGATCCGGA
HLTF_D_e15_144.6	GCCGGTTCTCTAATACGATC
HLTF_D_e16_139.7	CTGAGCATTTGGATTTCGTA
HLTF_H_e23_181.0	TGGATTTGTGTTTACTCGTT
HLTF_H_e23_181.1	TTTGTGTTTACTCGTTTGGA
HLTF_H_e23_181.2	TACTCGTTTGGATGGTTCCA
HLTF_H_e24_81.1	CTTGCTTCTGACCAAGTCTA
HR_e17_165.23	AGGCACCAGCACGGCCTCTC
HR_e17_165.25	GCACCAGCACGGCCTCTCCG
HR_e17_165.26	CACCAGCACGGCCTCTCCGG
HR_e18_129.18	TGACGCTGACTGTGCTCACC
HR_e18_129.2	GCATAAAGCAGGTGGCAGTC
HR_e18_129.3	CATAAAGCAGGTGGCAGTCA
HSPBAP1_e3_182.1	TGTGACCAGTCTAGTATTTC
HSPBAP1_e3_182.7	ATGGTCCAGAAATACTAGAC
HSPBAP1_e3_182.8	ACTAGACTGGTCACAGTTCC
HSPBAP1_e4_137.11	GGACTCCTATGGTTGTAACT
HSPBAP1_e4_137.15	ACCATAGGAGTCCAGATGAC
HSPBAP1_e4_137.7	TACATTGTGGATTGGCTCCT
HSPBAP1_e5_172.2	CATGCGGTTACACTGAGCCC
HSPBAP1_e5_172.4	CTTTGAGCTTTCCGGAACTG
HSPBAP1_e5_172.8	TCTTCATAAGGGATTCTAGT
INO80_D_e13_81.3	AGTACAGAGCATTGCCCTTC
INO80_D_e14_96.3	CAATTGTTAAGTGTAGACGC
INO80_D_e15_60.2	CTGATGACTTTTCTATCATG
INO80_D_e16_143.11	TGTATTGCCACTTGACCCGC
INO80_H_e27_128.7	GACCAGGATGATAGACCTAC
INO80_H_e28_95.5	GTCTCGCCTCTCCGAGATCT
INO80_H_e29_73.0	TTCCTGTTAAGCACACGAGC
INO80_H_e30_95.11	CTTTGTCTGCCCTAAGCGGT
JARID2_e12_115.1	TCCCTGGCTAAATATTGGCA
JARID2_e12_115.8	GACCATGCCAATATTTAGCC
JARID2_e13_106.4	CTGCAAGCCAATGGCACCCC
JARID2_e13_106.5	TGCAAGCCAATGGCACCCCA
JARID2_e14_158.16	CCGGGAAGCAGACGACAAAC
JARID2_e14_158.20	CTCTTTGCACAGCACCTCCG
JMJD1C_e21_131.3	TTTTTTGTACGTCCTGATCT

JMJD1C_e21_131.4	TACGTCCTGATCTAGGACCC
JMJD1C_e21_131.6	TCAGGACGTACAAAAAATCC
JMJD1C_e22_86.0	GCTGCTAAAGATCATGATAT
JMJD1C_e22_86.1	GTAAATATACTAGTTTATGT
JMJD1C_e22_86.2	CAACATCAGAAACTTCAATA
JMJD4_E1_139.0	GTTTTCCAGCGCCTTCACGC
JMJD4_E1_139.14	CCGTAGCAGGTGGTCGAAGT
JMJD4_E2_166.20	TGGGGTTCGAGTTGTATTCC
JMJD4_E3_126.0	CTGTGTACTTCTCGTCCGAC
JMJD4_E3_126.7	ACTACCGCTTTGTCTACGCG
JMJD4_E4_268.37	GCTGGGGAGGTCACGTCGTA
JMJD4_E4_268.42	GTTGCCGTGGCGGTCCCGCA
JMJD6_e3_280.11	ATCAAAGTGACCCGAGACGA
JMJD6_e3_280.3	TGGGATTCACATCGACCCTC
JMJD6_e3_280.45	TGAATCCCAGTTCCGGAGCG
JMJD6_e4_52.0	CGATAGTAGTGTCGAGATTG
JMJD7_e2_95.10	GCGGATAATGCACGGCCTGT
JMJD7_e3_254.41	ACGGCATCCGCGTAACCATC
JMJD7_e4_57.0	TGCCCGATGCTGTGAACTTC
JMJD7_e5_96.4	GGGATGAAGGGCCGGTCGCT
JMJD7_e6_77.10	TCAGTTAGCTGGTAGGTTGC
JMJD8_E3_49.6	CAGGATGACGGGCCTGACGA
JMJD8_E4_97.12	GCAACCTGTCGCGGGAGCAC
JMJD8_E5_69.9	CTGCTCCACATACTCCTGGA
JMJD8_E6_120.1	GGACAACAACTTCACCGAGT
JMJD8_E7_68.4	GGGTACTCAGAAGTGATCTA
JMJD8_E9_30.0	TGCTGTACTTCCCCGACCGC
KAT2A_BD_e17_64.2	TTGGGCCAGCAGGTTTTTGA
KAT2A_BD_e17_64.4	GAGGGTTGTGTAGAGCTGGT
KAT2A_BD_e17_64.6	GGGTTGTGTAGAGCTGGTCG
KAT2A_BD_e18_158.15	AGTTGGCGATGACCCGCTGC
KAT2A_BD_e18_158.19	GCTTCCGGGTCACGTAGTAG
KAT2A_BD_e18_158.5	CTGTCGCGAGTACAACCCCC
KAT2B_BD_e17_85.0	TCAAAGCGCTTGGCCCTTCA
KAT2B_BD_e17_85.3	TGAAGTTATAAGGTTCCCCA
KAT2B_BD_e17_85.8	TGAAGGGCCAAGCGCTTTGA
KAT2B_BD_e18_190.12	GGGGTTGTACTCTTTGCAAT
KAT2B_BD_e18_190.13	ACACGTAGTACCTATTCTTG
KAT2B_BD_e18_190.14	ATTCTTGAGGCGTTCACTCA
KDM1A_e11_140.8	GGGACACAGGCTTATTATTG
KDM1A_e13_135.3	GAACTCGGCAGTAATATCTC
KDM1A_e14_112.1	CAACTCTCTCCCTTAAGCAC
KDM1A_e15_133.6	AGTGCGACAGGTTCGCTACA
KDM1A_e16_188.13	GGTGGCACAAACTGAACGGC
KDM1A_e17_115.1	GTGTTTTGATCGGGTGTTCT
KDM1A_e18_128.5	TGTGATTGTTGGCCGATGCC
KDM1A_e19_147.2	TCGTTGGCGTGCTGATCCCT
KDM1A_e8_82.0	CGAGTTGCCACATTTCGCAA
KDM1A_e9_95.4	GTCCGTTGGCTTCATAAAGT
KDM1B_e11_138.6	TCTTTTAAAGGCGTCACAGT
KDM1B_e12_132.8	TCAATTATCACCGAGTACCC
KDM1B_e13_75.2	GTGCAGGTTACCACTACAGA
KDM1B_e14_117.3	GCTATCAACAGCTTAGGCGC
KDM1B_e15_126.10	GGCAAAAAGCCCTCGCTTGC
KDM1B_e16_123.4	GGGAGGCTGTCGCATCCGTG
KDM1B_e17_153.11	CAAGGAACCGTCTTTTTCGC
KDM2A_E10_56.0	CCTACAGACACATTAGTGTT
KDM2A_E10_56.4	CCAAACACTAATGTGTCTGT
KDM2A_E8_93.0	TACTGTCTAATGAGTGTTCG
KDM2A_E8_93.6	TTTGGTATCACATCCATCAA
KDM2A_E9_154.1	ACCTGGAGCTGTACGAGAAT
KDM2A_E9_154.10	TCAGCGCATTGAGCTCAAGC
KDM2B_e6_107.0	TGGACTGGGTGGACAACATG
KDM2B_e6_107.3	CATCTCTGCAATGGCGTTCG
KDM2B_e6_107.7	TGTTGTCCACCCAGTCCACC
KDM2B_e7_94.10	AAGTGCCTCCAAAGTCGATG
KDM2B_e7_94.3	ACTTTGGAGGCACTTCCGTT
KDM3A_e21_131.0	TTCCACTGCCCGAGTACACA
KDM3A_e21_131.1	CCCGAGTACACAAGGCGAGA
KDM3A_e21_131.13	CGCCTTGTGTACTCGGGCAG
KDM3A_e22_95.1	ACTCCTGAAGATCGGAAATA
KDM3A_e22_95.4	GCTAATGTCATGGTCTATGT
KDM3A_e22_95.5	GTTCCATATTTCCGATCTTC
KDM3B_e19_131.11	ATTGAGCCTGCCATCTCGTT
KDM3B_e19_131.2	ACGAGATGGCAGGCTCAATC
KDM3B_e19_131.7	CATAGGCGTTGTACATCTTG
KDM3B_e20_86.1	GTCTGATGCTGTTAATGTGA
KDM4A_e5_95.0	TGTGGAAGACATCCTTTGCT
KDM4A_e6_50.2	ACCTGAGCATGGAAAGCGGT
KDM4A_e6_50.3	TCCAACCGCTTTCCATGCTC
KDM4A_e7_104.4	ATATTTCTTCAGCATTAACG
KDM4A_e7_104.5	TCAGCATTAACGGGGAAATC
KDM4A_e8_95.6	AACCATGGTTAAAGCCGGCA
KDM4A_e8_95.8	AGCCGGCATGGTAACCATAA
KDM4B_e6_106.0	TGTGGAAGACCACCTTCGCC
KDM4B_e6_106.11	GTCCTCGGTGTGCCAGGCGA
KDM4B_e7_50.1	ACCAGAGCACGGCAAGCGCC
KDM4B_e7_50.4	TCCAGGCGCTTGCCGTGCTC
KDM4B_e8_104.11	TGAGGGTCATCTTATGCCGC
KDM4B_e8_104.12	GGTCATCTTATGCCGCAGGA

KDM4B_e8_104.5	AGTACGGGATCCCCTTCAGC
KDM4B_e9_96.5	ACCCGTGATTGAAGCCGGCG
KDM4B_e9_96.6	GAAGCCGGCGTGGTAGCCGT
KDM4B_e9_96.7	AAGCCGGCGTGGTAGCCGTA
KDM4C_e5_106.1	TGCATGGCACACCGAAGACA
KDM4C_e5_106.7	TTCGGTGTGCCATGCAAACG
KDM4C_e6_50.0	TATGCTATACCTCCGGAGCA
KDM4C_e6_50.1	TCAAGTCGTTTTCCATGCTC
KDM4C_e6_50.2	AGTCGTTTTCCATGCTCCGG
KDM4C_e7_104.5	GAAGAAATGCATCACACCCT
KDM4C_e7_104.6	AAGAAATGCATCACACCCTT
KDM4C_e7_104.8	CACACCCTTGGGAGCTGCTT
KDM4C_e8_96.1	TTCCCATATGGCTACCATGC
KDM4C_e8_96.5	AACCAGCATGGTAGCCATAT
KDM4D_e3_349.26	TCCCTTCAATCGCATAACTC
KDM4D_e3_349.27	CTTCAATCGCATAACTCAGG
KDM4D_e3_349.33	CCATGGTTTCAACTGCGCAG
KDM4D_e3_349.39	CTCCTGAGTTATGCGATTGA
KDM4D_e3_349.45	GGCCACCTTGTGCCGCAGGA
KDM4D_e3_349.63	CTCTGTATGCCAAGCAAACG
KDM5A_e11_79.1	TTCTTTTTGCTGGCACATTG
KDM5A_e12_163.16	GCTGGGATTCAAATAACTCG
KDM5A_e12_163.5	CATGAACCCCAACGTGCTAA
KDM5A_e12_163.8	CCATGCTCCATTAGCACGTT
KDM5A_e13_120.2	CTATCACTCTGGATTTAACC
KDM5A_e13_120.7	AATCCAGAGTGATAGGCACG
KDM5B_e11_86.2	ATTCAATTAACTACTTGCAC
KDM5B_e11_86.5	AGTTAATTGAATAGCTCCAG
KDM5B_e12_163.12	GCTGATGGAGGAGATCCGGC
KDM5B_e12_163.13	CTGATGGAGGAGATCCGGCT
KDM5B_e12_163.17	ATACCCTGGGACTCCATACC
KDM5B_e13_120.0	TACCGAACTAATCAGTGTGC
KDM5B_e13_120.1	ACCGAACTAATCAGTGTGCT
KDM5B_e13_120.2	CCGAACTAATCAGTGTGCTG
KDM5C_e11_95.9	GACCATGCCCACGTAGAGCC
KDM5C_e12_163.1	GTGAGCCGAAGACCTGGTAT
KDM5C_e12_163.15	GGCTGGCTATCAAATAGTTC
KDM5C_e12_163.18	AAGTGAGGGCACCCCATACC
KDM5C_e13_120.10	AAGCCGCTGTGGTAAGCACG
KDM5C_e13_120.3	CCAAGGCTACAACTTTGCCG
KDM5C_e13_120.6	CCTCGGCAAAGTTGTAGCCT
KDM5D_e13_79.1	AGCATTTTGTTGGCATATTG
KDM5D_e14_163.1	CTGGTATGGTGTACCCTCCC
KDM5D_e14_163.18	GGTACACCATACCAGGTCTT
KDM5D_e14_163.8	GATTCATGAGAGTGACAAGC
KDM5D_e15_118.1	TCCGCACAAACCAGTGTGCA
KDM5D_e15_118.5	CTTCAGCAAAATTGTAGCCT
KDM5D_e15_118.7	AAACCACTGTGGTAAGCACG
KDM6A_e25_115.7	AGAACACCCCAGTAACCTTC
KDM6A_e25_115.8	ACAAACCATTCACAGTCACC
KDM6A_e26_188.12	GTTCATTGGGTTCAGGCTAT
KDM6A_e26_188.17	CGCTGAATAAACCTATACAC
KDM6B_e18_115.4	TCGCGGTGCACGAGCACTAC
KDM6B_e18_115.6	GCGAACCACTCGCAGTCGCC
KDM6B_e18_115.7	CGAACCACTCGCAGTCGCCT
KDM6B_e19_188.0	GCACGGCGTGGACTACTTGA
KDM6B_e19_188.26	CGCTGCACGAAGCGGTACAC
KDM6B_e19_188.5	TACCGCTTCGTGCAGCGACC
KDM7A_e6_187.15	TGGCCAATAATTTTCCACCC
KDM7A_e6_187.17	GAAAGTTTTTTGGCTATATC
KDM7A_e6_187.7	GATTTCCACATTGACTTCGG
KDM7A_e6_187.8	ACTTCGGTGGAACTTCAGTC
KDM7A_e6_187.9	CAGTCTGGTACCATGTCCTC
KDM8_e4_133.0	GAGTATATCCAGGAGATCGC
KDM8_e4_133.7	TGAACTCGTTGACCGTCATG
KDM8_e4_133.8	GAACTCGTTGACCGTCATGA
KDM8_e5_45.0	CAAAGAGCTGGTGCTGAGCA
KDM8_e5_45.1	GCAAGGTACCCGACGTCCCT
KDM8_e6_150.10	CAATGCCTGGTTTGGTCCCC
KDM8_e6_150.12	AACCATCTCCCCACTACATC
KDM8_e6_150.27	CGCCCAGGCTGCAGTAGTCG
KDM8_e7_159.0	TGACGTGGAGAATCCCGACC
KDM8_e7_159.11	AGTATTTCACCGGGATGAAC
KDM8_e7_159.7	CTGGCATTACGTGCGGGCTC
KMT2A_BD_e16_174.15	CTGGTGGATCAGGTCCTTCG
KMT2A_BD_e16_174.16	TGGTGGATCAGGTCCTTCGG
KMT2A_BD_e16_174.22	TCTGTCTCGGGATTTAAGTC
KMT2A_BD_e17_111.1	GCAGCCATTAATTCAGATGG
KMT2A_BD_e17_111.4	GAAGGACTTGACCATGCTGT
KMT2A_BD_e17_111.6	CTGTCCTCCATCTGAATTAA
KMT2A_BD_e18_50.1	GGTTCAGTGTCAAAAAGTCC
KMT2A_BD_e18_50.2	GTGTCAAAAAGTCCAGGTTT
KMT2A_BD_e18_50.3	GACTTTTTGACACTGAACCA
KMT2B_e36_130.0	CTGTAAGCGCAACATCGACG
KMT2B_e36_130.2	GTAAGCGCAACATCGACGCG
KMT2B_e36_130.3	TAAGCGCAACATCGACGCGG
KMT2B_e37_218.0	CGGGTGCTATATGTTCCGCA
KMT2B_e37_218.11	GATGGGGAACTTGTAGTCGT
KMT2B_e37_218.25	GGCGGCATTGCCATGCATCG
KMT2C_e58_117.3	TTCGAAACGAAGTAGCCAAC

KMT2C_e58_117.5	CCCGATGTACTCAATGACCA
KMT2C_e58_117.6	CAATGTCTCGAGCAGCATAC
KMT2C_e59_74.1	GTGATTGACGCGACGCTCAC
KMT2C_e59_74.2	ATTGACGCGACGCTCACAGG
KMT2C_e59_74.3	TTGACGCGACGCTCACAGGA
KMT2C_e60_109.0	TTCGTGTGCACCTAATTGTG
KMT2C_e60_109.5	TCCTTTCTGGATTCTCCGAC
KMT2C_e60_109.7	CACAATTAGGTGCACACGAA
KMT2D_e50_153.15	GTAGATTTTCTCCCGCCGGT
KMT2D_e50_153.16	GGCCACCTCGTTCCGAATGA
KMT2D_e50_153.9	TGGCACCATCATTCGGAACG
KMT2D_e51_74.0	GTGATTGATGCTACGTTGAC
KMT2D_e51_74.1	ATTGATGCTACGTTGACCGG
KMT2D_e52_109.3	TCCAGCCGGCGAATCCCCAA
KMT2D_e52_109.6	TCCTTTGGGGATTCGCCGGC
KMT2D_e52_109.8	ACGACTTCGGCCACACAGTT
MINA_e2_254.19	GAACAGGGACCCATAGTATG
MINA_e3_120.4	AATGTGTACATAACTCCCGC
MINA_e4_129.12	CCACGCTGTACTCTCGTGCC
MINA_e5_104.4	ATCAAGCGGACACTCCTGCG
MINA_e6_103.8	AGTTACGGACCGGCATACCC
MINA_e7_172.5	ACACTTGCAGACCGGCTGGA
MLL1_e35_130.12	AAAAGACCCCGGCCATGGAT
MLL1_e35_130.6	TTCCCGCTTGTCAGTCTGGA
MLL1_e35_130.8	GATGGAGCGGATGACGTTGC
MLL1_e36_237.14	TCCTCGGTAGATCTTACGCA
MLL1_e36_237.6	CTCGGGTCATCAATATTGAT
MLL1_e36_237.8	GCGTAAGATCTACCGAGGAG
MLL5_e11_105.0	GCACTTATCATTGAATACAG
MLL5_e11_105.4	TATTCAATGATAAGTGCATC
MLL5_e11_105.5	TCAATGATAAGTGCATCAGG
MLL5_e12_118.4	GTGTTGATGCAAGGACTTTT
MLL5_e12_118.6	GGAATGAGGCTCGATTCATC
MLL5_e12_118.7	ATGAGGCTCGATTCATCAGG
MLL5_e13_110.1	TTCTATACACAGTATTCCAA
MLL5_e13_110.2	TCTATACACAGTATTCCAAA
MLL5_e13_110.3	ATTGCCTTTGATTTTGACTA
MOF_e10_135.9	GTGTGACACAGATCACGTGC
MOF_e5_164.14	CAATTTCGTAGTTCCCGATG
MOF_e6_90.7	TGCGGTAGATCTCTTTCCCG
MOF_e7_141.6	CCTGACTGAGGTGGACCGGC
MOF_e8_94.3	CCTACCAACGCCGCGGCTAC
MOF_e9_151.11	GAGATCCTGCGGGACTTCCG
MORF_e12_162.0	CTTGTTAGCCAAGCTCTTCC
MORF_e12_162.8	CATCATAATACAACGTTTTG
MORF_e12_162.9	AATACAACGTTTTGTGGTCC
MORF_e13_94.8	CTGGGGCATGATCATTATGC
MORF_e14_210.22	AGACGGCCCAGATCGGAGAG
MORF_e14_210.3	TGAAAAGCCTCTCTCCGATC
MOZ_e12_162.1	CAAAACCCTCTATTACGATG
MOZ_e12_162.7	ATGGCTCCACATCGTAATAG
MOZ_e12_162.8	TGGCTCCACATCGTAATAGA
MOZ_e13_94.4	ACCTGCCATAGCCCTTACGC
MOZ_e13_94.5	TAGCCCTTACGCTGGTATTG
MOZ_e14_210.0	TATTTGTTATCAAAGCGTGA
MOZ_e14_210.3	AGAGAAACCGTTATCTGATC
NAT10_e18_75.1	CCAGACTTTGGTGGTCTGTC
NAT10_e18_75.4	TGGTGGTCTGTCTGGTGGAA
NAT10_e18_75.5	GCTGTTCACCCAGATTATCA
NAT10_e18_75.6	CTGTTCACCCAGATTATCAA
NAT10_e18_75.7	CCAGACAGACCACCAAAGTC
NCOA1_e17_403.2	AATGTTGGCACAACGTCAGC
NCOA1_e17_403.26	GCTCACCATGCTGTTGCGGT
NCOA1_e17_403.31	TGCTAGTTGTGAAAACGGGC
NCOA1_e18_175.12	GTCTGGTGACTGATACCCGG
NCOA1_e18_175.13	GGTGACTGATACCCGGAGGC
NCOA1_e18_175.22	AGGGATTGGCATCGGCACCA
NCOA2_e19_235.10	GAGCAACCCTCGGATTCCCC
NCOA2_e19_235.14	GCATTTGCCTGGGGAATCCG
NCOA2_e19_235.16	CGAGGGTTGCTCATAGTTGC
NCOA2_e20_178.17	CGCCCATCCATTTATGTCGG
NCOA2_e20_178.18	GCCCATCCATTTATGTCGGA
NCOA2_e20_178.20	CCATCCATTTATGTCGGAGG
NO66_e1_980.106	AGTCACCCCAGGTATTGCGC
NO66_e1_980.131	GGCCTCGATGTCGTCGTAGT
NO66_e1_980.160	TCGTCCGTTGATGTAGCGAG
NO66_e1_980.162	CGTTGCGCAGCATCGAATCC
NO66_e1_980.7	GGATTCGATGCTGCGCAACG
NSD1_e19_117.0	TGAATTTGTGAATGAGTATG
NSD1_e20_142.13	TGGGACCAGCATCAATGATT
NSD1_e20_142.4	GTGAATGGAGATACCCGTGT
NSD1_e20_142.7	TAGTGCAAAAAGGCCTACAC
NSD1_e21_41.0	AACTACAACCTAGAATGTCT
NSD1_e21_41.1	ACATTCTAGGTTGTAGTTGA
NSD2_e19_73.7	GGTGGGGCCTGGTCGCCAAG
NSD2_e19_73.8	GTGGGGCCTGGTCGCCAAGA
NSD2_e20_117.0	GAATTTGTTAACGAGTACGT
NSD2_e20_117.3	CGTTGGGGAGCTGATCGACG
NSD2_e21_142.16	TGGGGCCAGCGTCTATTATA
NSD2_e21_142.6	GTGAATGGGGACACTCGTGT

NSD3_e20_117.1	AATTAGTTACACTGTTCTCG
NSD3_e20_117.2	ATTAGTTACACTGTTCTCGT
NSD3_e21_142.0	ATAATTGATGCCGGCCCAAA
NSD3_e21_142.11	TTGGGCCGGCATCAATTATA
NSD3_e21_142.4	GTGAATGGAGATGTTCGAGT
PADI4_e10_108.0	CCCACACAAAACGCTGCCCG
PADI4_e11_155.15	CAGGGGCAAGGAATACCCGC
PADI4_e12_145.10	AGCTCAGGAACTCGTCCACG
PADI4_e13_103.11	CTTGATCCCTTCGAACAGCA
PADI4_e14_70.0	GAGAGAACATAATTCATTTG
PADI4_e15_129.1	GAACCGCGAGCTGCTGAAGC
PADI4_e8_104.8	CACGCGTACACCTCCTGCGG
PADI4_e9_112.0	CCTGAAGTCAGTGACTACTC
PBRM1_BD1_e3_98.0	TAATACCATCCGAGACTATA
PBRM1_BD1_e3_98.1	CCGAGACTATAAGGATGAAC
PBRM1_BD1_e3_98.6	TTCATCCTTATAGTCTCGGA
PBRM1_BD1_e4_148.0	TCAACCAGACTATTATGAAG
PBRM1_BD1_e4_148.5	TGGATTTTCATCAAGTCAAT
PBRM1_BD1_e4_148.6	GAAACCACTTCATAATAGTC
PBRM1_BD1_e5_86.0	ATAAAGCCGCTTGCAAACTC
PBRM1_BD1_e5_86.3	CAAACTCATTTCTTGTTCGA
PBRM1_BD1_e5_86.5	TTGCAAGCGGCTTTATATTC
PBRM1_BD2_e6_117.1	GTTGTAGCTACAAATCCATC
PBRM1_BD2_e6_117.2	TCGCTAATGAGACGTCCTGA
PBRM1_BD2_e6_117.5	AGGATCTCCTTCAAGTAAGC
PBRM1_BD2_e7_69.1	ATCTCAAGACCATTGCCCAG
PBRM1_BD2_e7_69.2	GCAATGGTCTTGAGATCTAT
PBRM1_BD2_e8_99.1	GCCAAAACTTATAATGAGCC
PBRM1_BD2_e8_99.2	GCCAGGCTCATTATAAGTTT
PBRM1_BD2_e8_99.4	TGCGAGGAGATCTATATCTT
PBRM1_BD3_e12_158.2	AGGAGTTGTCGGAATAACCA
PBRM1_BD3_e12_158.3	GGAGTTGTCGGAATAACCAA
PBRM1_BD3_e13_140.1	TTTTAGAACTCGCTTGTAGA
PBRM1_BD3_e13_140.2	TCGCTTGTAGATGGCTGAAT
PBRM1_BD3_e13_140.4	ATTGGGCACATTATAGCGTT
PBRM1_BD4_e15_277.0	GTTCTTGAAGCTCGAGAGCC
PBRM1_BD4_e15_277.10	AGCTGATGTTCCGGAATGCC
PBRM1_BD4_e15_277.14	GCCCTCCTCATTATAGTGCC
PBRM1_BD5_e17_310.11	CTCCGGCTCATTGTATGTAC
PBRM1_BD5_e17_310.16	TTAATAGTCAGATAGTAGTC
PBRM1_BD5_e17_310.9	TGCCTGTACATACAATGAGC
PBRM1_BD6_e17_246.5	TTGAAAATAATCGCTACCGT
PBRM1_BD6_e17_246.7	AAGTATTGGAACGAGCAAGA
PBRM1_BD6_e17_246.8	CTTGAAATAAATCAAGCCGA
PCAF_e11_118.0	ATTAAAGATGGCCGTGTTAT

PCAF_e11_118.1	AAAGATGGCCGTGTTATTGG
PCAF_e11_118.7	TCTCTGTGAATCCTTGAGAT
PCAF_e11_118.9	AACAGATACCACCAATAACA
PCAF_e12_59.0	ATTCTTTCAAATGATTCATC
PHF2_e6_79.6	AGGCGCCCCCAGAGTCGATG
PHF2_e7_163.12	CAGACCGCCAGCGCTCATAC
PHF2_e7_163.13	AGACCGCCAGCGCTCATACA
PHF2_e7_163.9	ACTTGTAGCATTTGTCGACC
PHF2_e8_50.1	ACTCACCCCTGTGGACTGCC
PHF2_e8_50.3	GTGGACTGCCTGGCCTTCGC
PHF8_e7_187.10	AGGCAGTACTTCTGTACATT
PHF8_e7_187.12	GACAGCTTTCGAACAATCTT
PHF8_e7_187.13	CAATCTTCGGTGTCTCCACA
PHF8_e7_187.2	TGTCATGGGTCGAAAACTTG
PHF8_e7_187.8	AGCTATCTCGCACACTCATG
PhIP_BD1_e30_86.1	AGAATGGGGTACCAATCCCA
PhIP_BD1_e30_86.7	GTACCCCATTCTCCATCAAG
PhIP_BD1_e31_121.0	CTCAGCATTTGTGGCCCCCG
PhIP_BD1_e31_121.12	ATAGGCTTGCAGATCCACGG
PhIP_BD1_e31_121.5	TTAATTGTACTTAGATCCGT
PhIP_BD1_e32_126.0	GCGGGTTTCTTCCCTAATGT
PhIP_BD1_e32_126.4	GATTTCACAATAGGGCTTCC
PhIP_BD1_e32_126.5	TATATCGAACTTCCCACATT
PhIP_BD2_e35_80.0	TTCAGAGCCTTTCCGTCAGC
PhIP_BD2_e35_80.2	CAAGGAGATCTACCGGCTGA
PhIP_BD2_e35_80.3	AGATCTACCGGCTGACGGAA
PhIP_BD2_e36_153.0	AGACATCATTGACACTCCAA
PhIP_BD2_e36_153.1	TACCGTTAGAGAAACTTTAG
PhIP_BD2_e36_153.9	CTAACGGTAGCAAAATCCAT
PhIP_BD2_e37_80.0	GTTCTTCAAAGAAAGCAGAC
PRDM2_e3_47.0	CTTCTGCTGTTGACAAGACC
PRDM2_e3_47.1	GCTGTTGACAAGACCCGGAT
PRDM2_e4_104.3	AAAAAATTTGGGCCATTTGT
PRDM2_e5_153.4	ACTGGCTGCGATATGTGAAT
PRDM2_e5_153.5	CTGGCTGCGATATGTGAATT
PRDM2_e6_63.3	AGCTCCTGGTCTGGTACAAT
PRDM2_e6_63.7	AGCTCCTCGCCCGGCGCGAT
PRDM8_e8_135.1	GAACAGTACCGTATATCTTT
PRDM8_e8_135.11	AGCATTCTCAGGGATGTCGC
PRDM8_e8_135.12	ATTCTCAGGGATGTCGCAGG
PRDM8_e9_190.13	AACTAGTAACTCCTCGTCTT
PRDM8_e9_190.8	TGTTCTACCGCTCTCTCCGC
PRDM8_e9_190.9	CCGCAGGATTGCCAAAGACG
PRDM9_e10_135.11	GTATGGGGATGAATACGGCC
PRDM9_e8_158.10	TCTGCACTTTGGCCCTTATG

PRDM9_e8_158.18	CTGTAATTCGGCCCTCATAA
PRDM9_e8_158.7	TGAGGCATCTGATCTGCCGC
PRDM9_e9_68.4	AAGATAAATCCTGGGCCAAC
PRMT1_e4_73.0	CTGGACGTCGGCTCGGGCAC
PRMT1_e4_73.6	GCCGGGGCCCGCAAGGTCAT
PRMT1_e4_73.7	GACCTTGCGGGCCCCGGCCT
PRMT1_e6_143.13	TGCTCTATGCCCGGGACAAG
PRMT1_e6_143.14	CTTGTCCCGGGCATAGAGCA
PRMT2_e6_162.10	CCAGGATGACTTTATCCGTC
PRMT2_e6_162.11	CAGGATGACTTTATCCGTCA
PRMT2_e6_162.8	TCTTCTGTGCACACTATGCG
PRMT2_e7_165.10	GGATGTGGTGCTGCCCGAGA
PRMT2_e7_165.12	GCCCGAGAAGGTGGACGTGC
PRMT2_e7_165.21	TGTCAGCAAAGCCGTTCTGC
PRMT3_e10_100.0	GATACTATTACACTAATTAA
PRMT3_e10_100.1	TAGATGTTATCATATCTGAG
PRMT3_e10_100.2	ATAACATCTACTTTTTCTAC
PRMT3_e9_122.3	CTATGTTTGCTGCTAAAGCT
PRMT3_e9_122.5	TGCTAAAGCTGGGGCGAAGA
PRMT3_e9_122.6	GCTGGGGCGAAGAAGGTTCT
PRMT5_e11_123.12	CAGCATACAGCTTTATCCGC
PRMT5_e11_123.13	ATACAGCTTTATCCGCCGGT
PRMT5_e11_123.14	TATCCGCCGGTCGGCCTGCT
PRMT5_e12_106.10	CTGATGAGACTACGGTCACT
PRMT5_e12_106.3	CCGTAGTCTCATCAGACATG
PRMT5_e12_106.4	CGTAGTCTCATCAGACATGA
PRMT6_e1_281.34	TCAGCCACTTGGTTCGCGCG
PRMT6_e1_281.50	CTCTACCGCGTACACGCGCC
PRMT6_e1_281.6	CCGGCGCGTGTACGCGGTAG
PRMT7_e5_73.0	CACGGGACTCTTGTCAATGA
PRMT7_e5_73.1	GGGACTCTTGTCAATGATGG
PRMT7_e5_73.2	TCAATGATGGCGGTCACAGC
PRMT7_e6_79.2	TGGCTTTAGTGATAAGATTA
PRMT7_e6_79.3	TTCACAGCAGCATCAGCCAT
PRMT8_e3_68.3	TCCATGTTCGCTGCCAAGGC
PRMT8_e3_68.8	CCCTGCCTTGGCAGCGAACA
PRMT8_e4_64.0	CTACTCAGAGAAGATCATTA
PRMT8_e4_64.1	GATCATTAAGGCCAACCACT
PRMT8_e4_64.2	CTCTGAGTAGTCAGAAATAC
PRMT8_e5_143.10	TCAACACGGTGATCTTTGCC
PRMT8_e5_143.11	CAACACGGTGATCTTTGCCA
PRMT8_e5_143.13	AAAGATCACCGTGTTGAGCA
PRMT9_e4_168.0	TGAACTTGCCTGTGATGTCG
PRMT9_e4_168.4	GTTTGCTGCCACGACATCAC
PRMT9_e4_168.5	ATCACAGGCAAGTTCATACA
PRMT9_e4_168.6	GGCAAGTTCATACATGGTCT
PRMT9_e4_168.7	CATGGTCTTGGATAACTCAC
RAD54L_D_e7_237.21	ATCCTTAGATCCTCCATCGA
RAD54L_D_e8_125.10	GACTCCAACATGAAGGCGGA
RAD54L_D_e9_95.6	GGAGATGAGCACCCGCCGGC
RAD54L_H_e14_124.7	CTTTGTCACTGCTACGGCTT
RAD54L_H_e15_79.0	TACTTATACGTCCGCCTGGA
RAD54L_H_e16_180.9	TCTCATTGGGGCTAACCGGC
RAD54L2_D_e10_135.5	AGAATATCCGCTCTCGCCGC
RAD54L2_D_e7_118.9	GGACTGTTTTGGCTGGCGTG
RAD54L2_D_e8_134.9	CAGGCTTGTTGTCAGCCGGG
RAD54L2_D_e9_197.16	GACTGGGTGAGAACGCTTCT
RAD54L2_H_e13_119.0	CTTCTGACTAATTACCAGAC
RAD54L2_H_e14_119.4	CCCTGTCCACCTGGTACCGA
RAD54L2_H_e15_100.9	AATAAGCCGCTCCCTCTCAA
RAD54L2_H_e16_135.7	GTATGTCGGGTATACCGTTA
SCRAP_D_e14_137.13	CTCCCAGTTCAACATCACGC
SCRAP_D_e14_137.15	CAGTTCAACATCACGCTGGT
SCRAP_D_e15_170.7	GATAGCGCCAGTTCTTGCGA
SCRAP_D_e15_170.8	GCGCCAGTTCTTGCGACGGA
SCRAP_H_e28_170.10	TAGTAGATCCATCCAGGCGC
SCRAP_H_e29_197.18	CAATTCGGTGACAGCGGTCC
SETD1A_e17_120.4	CCGGAGCCGGATCCACGAGT
SETD1A_e18_138.0	GCGGGAGAAGCGCTACGTGC
SETD1A_e18_138.16	TGTCGTGGTCCACCCGGAAC
SETD1A_e18_138.17	GTAGCGCTTCTCCCGCATGT
SETD1A_e19_118.2	GCAGCCCATTGGCGTGGACG
SETD1A_e19_118.5	GATCTCCTCGTCCACGCCAA
SETD1A_e19_118.6	ATCTCCTCGTCCACGCCAAT
SETD1B_e15_95.7	ACCATCTCGTCAGCCGCGAT
SETD1B_e16_138.10	CGCGAAGTTGCCGCACTTGG
SETD1B_e16_138.11	GCACTTGGTGGCGTCGATGA
SETD1B_e16_138.9	GCGCGCGAAGTTGCCGCACT
SETD1B_e17_118.1	CTATGCCAAGGTGATCACGG
SETD1B_e17_118.2	GCAGCACATTAACGTCAATG
SETD1B_e17_118.3	TGACTATAAGTTCCCCATCG
SETD2_e5_124.1	TGTCCTAGAATATTGTGGAG
SETD2_e5_124.2	AGAGTTTAAAGCTCGAGTGA
SETD2_e5_124.5	GTACCTCTCCACAATATTCT
SETD2_e6_78.0	ATAATAGATGCCACTCAAAA
SETD2_e7_98.0	GGACTGTGAACGGACAACTG
SETD2_e7_98.2	GTGAACGGACAACTGAGGGT
SETD2_e7_98.3	TGAACGGACAACTGAGGGTT
SETD4_e6_175.12	CAGGTACGGAGCGAGTGCAC
SETD4_e6_175.16	GAAAGGCATTCCCGCTGCCT

SETD4_e6_175.19	GGGCCTCAGGTACACGGCTC
SETD4_e7_103.0	AAATTAGAACGACTTCACGT
SETD4_e7_103.2	GAAGAGGTATTCATCTGTTA
SETD5_e10_118.2	TGTGTGGATGCCCGTACTTT
SETD5_e10_118.4	GAGCATCATTACCGAAAGTA
SETD5_e10_118.5	AGCATCATTACCGAAAGTAC
SETD5_e11_110.0	GTGCGACACATGATTGCAGA
SETD5_e11_110.1	TGCGACACATGATTGCAGAT
SETD5_e9_105.1	ACTCTTATAATAGAGTATCG
SETD5_e9_105.2	CTCTTATAATAGAGTATCGT
SETD6_e4_118.10	GTTCTAGGGAGCGAACCCTG
SETD6_e4_118.16	ATCGGGGTGGGCTTCCATGA
SETD6_e4_118.19	TGAAGGGCAGCACGATGGAC
SETD6_e5_121.11	GGCACCATCACGGGGGAGTT
SETD6_e5_121.2	GGAGCCCAACTCCCCCGTGA
SETD6_e5_121.3	TTCTAGATTGGCGTTGTGAT
SETD6_e6_67.6	ATCTCATGGCCTTTAGGAAT
SETD7_e6_86.0	TGGAGAAGGACTTTTTTCAA
SETD7_e6_86.1	ACTTTTTTCAAAGGTAGCTG
SETD7_e6_86.2	CTTTTTTCAAAGGTAGCTGT
SETD7_e7_158.10	AGGCACAGTACTTGGATACG
SETD7_e7_158.11	GTACTTGGATACGTGGTTAT
SETD7_e7_158.13	GATACGTGGTTATAGGGCTC
SETD7_e8_95.11	CATTTGATGGGCCCAAAACG
SETD7_e8_95.12	ATTTGATGGGCCCAAAACGG
SETD7_e8_95.8	AGGGTGCGGATGCATTTGAT
SETD8_e7_190.11	ACGGGAGGCTCTGTACGCAC
SETD8_e7_190.14	TCTGAGCAAAACCTACTGCG
SETD8_e7_190.18	GTACAGAGCCTCCCGTTTCT
SETD8_e8_158.0	ACTAGAGAGACAAATCGCCT
SETD8_e8_158.5	GCCTCCCGAGACATCGCGGC
SETDB1_e20_55.0	ATCATTGATGCCAAGCTTGA
SETDB1_e20_55.1	TGCCAAGCTTGAAGGCAACC
SETDB1_e20_55.3	GCCCAGGTTGCCTTCAAGCT
SETDB1_e21_89.0	GTTTGTCCAGAATGTCTTCG
SETDB1_e21_89.8	CCAGGGGAAGCGAAGATCAT
SETDB1_e21_89.9	GGGTATCCACGAAGACATTC
SETDB1_e22_45.0	GGGCTGGGACAGAACTTACT
SETDB1_e22_45.1	GGCTGGGACAGAACTTACTT
SETDB1_e22_45.2	AAGTAAGTTCTGTCCCAGCC
SETDB2_e13_63.1	TTATTGGATGCCACAAAAGA
SETDB2_e14_89.3	CAACAGGAATTTTCCATTGG
SETDB2_e14_89.4	TTGGTGAAGAATGCCACCAA
SETDB2_e14_89.5	ACATTCTGTACCAAGAGATT
SETMAR_e2_380.19	TGACTCAATGGTACCTAAGT
SETMAR_e2_380.24	TAGGTACCATTGAGTCAATT
SETMAR_e2_380.25	ACCATTGAGTCAATTCGGAC
SETMAR_e2_380.4	AGGCTGGGGACTTCGTACCT
SETMAR_e2_380.5	ACCTTGGAATTTATACCGAA
SHPRH_H_e26_93.1	AAAATACAGCTTAGAGATCC
SHPRH_H_e26_93.4	CGTTGAGAAAACGAGTGCTT
SHPRH_H_e26_93.5	AAAACGAGTGCTTTGGCCCC
SHPRH_H_e27_179.10	AATTCGGTGCACCCTCCCTA
SHPRH_H_e27_179.16	TCATGGGCAGGGTTCAATAT
SHPRH_H_e27_179.9	ATAGGGAGGGTGCACCGAAT
SIRT1_e4_153.5	TCAATATCAAACATCGCTTG
SIRT1_e4_153.6	TCAAACATCGCTTGAGGATC
SIRT1_e5_148.2	TACCCAGAACATAGACACGC
SIRT1_e5_148.3	GAACATAGACACGCTGGAAC
SIRT1_e6_80.0	GTTGACTGTGAAGCTGTACG
SIRT1_e7_187.11	ATAGCAAGCGGTTCATCAGC
SIRT1_e7_187.14	GCTGGGCACCTAGGACATCG
SIRT2_e10_60.1	GCTCTGACAGTCTTCACACT
SIRT2_e11_56.0	ATACAGGAGAAGAAACGCGC
SIRT2_e12_77.3	GGTCATGGGTACCTCCTTGC
SIRT2_e13_52.4	CCTTGTTGATGAGCAGGCGA
SIRT2_e14_70.3	GGGATGATTATGGGCCTCGG
SIRT2_e6_107.11	TAGGTTGTCATAGAGGCCGG
SIRT2_e7_57.7	TCCTTGGCGAGGGCGAAGAA
SIRT2_e8_69.0	CTACTTCATGCGCCTGCTGA
SIRT2_e9_130.10	GGCACGAATACCCGCTAAGC
SIRT3_e3_210.0	GTACGATCTCCCGTACCCCG
SIRT3_e3_210.28	GGTACGGGAGATCGTACTGC
SIRT3_e4_101.10	GGGTCTTTGGCAGACTGTGC
SIRT3_e5_162.21	TTCACAACGCCGGTGCAGAC
SIRT3_e6_158.32	CCGCACGGCCTCGGTCAAGC
SIRT4_e2_341.15	CCCCAATCCGCCAGCGGTAC
SIRT4_e2_341.60	GTCTGGTATCCCCGATTCGG
SIRT4_e3_210.14	GTCCCAACCTGCGTTCAATG
SIRT4_e3_210.19	TCCCCGAAGAAAACGACATC
SIRT4_e4_158.10	GCCAAGTCATCCGACCGTGT
SIRT4_e4_158.9	CGCCAAGTCATCCGACCGTG
SIRT5_e5_210.25	GGGTGATGACCACGACTCGC
SIRT5_e5_210.7	GTTCTACCACTACCGGCGGG
SIRT5_e6_88.0	AAAACTCGATGTACCTCTTG
SIRT5_e7_54.0	AAGTTTCTCAACTGGGATGC
SIRT5_e8_124.3	CGACCTCACGTCGTGTGGTT
SIRT5_e9_116.7	GAATCTGTTCGTAGCTGGGG
SIRT6_e3_183.25	CGTGGGCCGCGCGCTCTCAA
SIRT6_e3_183.4	ACACCACCTTTGAGAGCGCG

SIRT6_e4_60.2	GCTCCACGGGAACATGTTTG
SIRT6_e4_60.3	CTTCCACAAACATGTTCCCG
SIRT6_e5_96.0	GTACGTCCGAGACACAGTCG
SIRT6_e5_96.1	TACGTCCGAGACACAGTCGT
SIRT6_e6_81.6	GGACCTGGCACTCGCCGATG
SIRT6_e6_81.7	TGGCCTCATCGGCGAGTGCC
SIRT6_e7_124.0	CGCCGACCTGTCCATCACGC
SIRT6_e7_124.20	CAGCGATGTACCCAGCGTGA
SIRT7_e4_70.10	CCATTAGGGCCCCGGTAGTC
SIRT7_e4_70.2	GCGTCTATCCCAGACTACCG
SIRT7_e5_73.1	GCAGACGGGTGATGCTCATG
SIRT7_e5_73.8	TGGCTCGGCCTCGCTCAGGT
SIRT7_e6_99.10	AGCTCGGAGATGGCCGTGCG
SIRT7_e6_99.9	GTTCCCGTGGAGCTCGGAGA
SIRT7_e7_210.3	TCCCAACAGGGAGTACGTGC
SIRT7_e7_210.4	CGTGCGGGTGTTCGATGTGA
SIRT7_e8_81.4	CGATGTAAAGCTTCGGCCGC
SIRT7_e8_81.6	AAAGCTTCGGCCGCCGGCTA
SMARCA2_BD_e28_218.0	CATCGAAGACGGCAATTTGG
SMARCA2_BD_e28_218.15	GGTGACAGTTTCTCAGCGGG
SMARCA2_BD_e28_218.19	CTCCAAATTGCCGTCTTCGA
SMARCA2_BD_e29_54.0	GGTGCCCAGTAATTCTCAGT
SMARCA2_BD_e29_54.2	TATTTCCAACTGAGAATTAC
SMARCA2_BD_e29_54.3	ATTTCCAACTGAGAATTACT
SMARCA2_D_e15_118.0	CGGAATCTTAGCCGATGAAA
SMARCA2_D_e16_67.2	TGACAAATGGGCTCCTTCTG
SMARCA2_D_e17_111.10	ACAAGGGAGCGACGCATGGC
SMARCA2_D_e18_135.13	AGTGGTGATTCTTCATTCGG
SMARCA2_H_e23_167.0	GCTGAACTGTATCGGGCCTC
SMARCA2_H_e24_164.7	TGGTCATCTTTGACAGCGAC
SMARCA2_H_e25_59.1	AAGACCGAGCTCACCGCATC
SMARCA4_BD_e31_52.1	GGATGCCGTGATCAAGTACA
SMARCA4_BD_e32_109.6	GCGGATGAGCTCGTAGTACT
SMARCA4_BD_e32_109.8	CTCGGGCAGCTCCTTTCGCG
SMARCA4_BD_e32_109.9	TCGGGCAGCTCCTTTCGCGA
SMARCA4_BD_e33_102.10	CGTTGAGGCTGCGGTACTTG
SMARCA4_BD_e33_102.5	GGTTGAAGGTCTGTGCGTTC
SMARCA4_BD_e33_102.8	CGTCCTTCTCTAGGTCGTTG
SMARCA4_BD_e34_59.1	AGTCGGTCTTCACCAGCGTG
SMARCA4_BD_e34_59.2	GAAGACCGACTGCAAGACGA
SMARCAD1_D_e12_100.3	CTGGCATACCTCTATCAGGA
SMARCAD1_D_e13_60.1	TTTATGGTGCCCTACTTTGA
SMARCAD1_D_e15_136.0	ATGACCGTAGTCTGTTTCGA
SMARCAD1_H_e21_119.0	TAGCCAATTTACCATGATGC
SMARCAD1_H_e22_183.0	TGATGAGTTTAATACCGATA

SMARCAL1_D_e10_66.2	CCGGCATAGCTGCTCGACAG
SMARCAL1_D_e8_135.4	CATCGCAGCCTTTTACCGGA
SMARCAL1_D_e9_159.0	CCCAGATTGCATCAACGTCG
SMARCAL1_H_e14_103.5	GCTCTTGCGTAATTGCGTCC
SMARCAL1_H_e15_183.0	GTGCAGCACATCCGCATCGA
SMARCAL1_H_e16_50.0	GAGGACCGCGTGCACCGCAT
SMYD1_e4_86.3	GCCTGGTGAACCATGACTGT
SMYD1_e4_86.7	TACAGTTGGGCCAACAGTCA
SMYD1_e6_73.0	GCCCTAGGCAAGATCTCAGA
SMYD1_e6_73.3	TCCTTCTGAGATCTTGCCTA
SMYD2_e6_55.0	AGATGAAGAACTTTCTCATT
SMYD2_e7_103.5	GCTGTACAGGAAATCAAGCC
SMYD2_e7_103.9	TTCTGCCAGGGTCCCTTTGT
SMYD3_e5_95.2	AAAGGCTTCAAAAAGGTCAA
SMYD3_e5_95.3	GCTTCAAAAAGGTCAAAGGC
SMYD3_e6_68.1	CATCTGTAATGCGGAGATGC
SMYD3_e6_68.3	GAGATGCAGGAAGTTGGTGT
SMYD3_e6_68.4	CTGCATCTCCGCATTACAGA
SMYD3_e7_103.2	GCGAGCAGTCCGAGACATCG
SMYD3_e7_103.4	GCAGTCCGAGACATCGAGGT
SMYD3_e7_103.7	CTCGGACTGCTCGCAGTAAG
SMYD4_e5_95.0	TGTGTCCTGACGTGACTATT
SMYD4_e5_95.4	ACAGCTTCAGTGTAACGCTC
SMYD4_e5_95.5	ACTCCCCAAATAGTCACGTC
SMYD4_e6_183.0	GGAGCATCGTTACCGACAGC
SMYD4_e6_183.6	GGCGTCACAGCGGATTAGAA
SMYD4_e6_183.9	TTCTAATCCGCTGTGACGCC
SMYD5_SET1_e10_57.0	GGAGAGTTTCTTAACTGTGA
SMYD5_SET1_e10_57.1	TTTCTTAACTGTGAAGGATC
SMYD5_SET1_e10_57.2	AGCAGCTCTGAAGCACAAAG
SMYD5_SET1_e11_95.0	CCTTTTGCATGTCACTGCTC
SMYD5_SET1_e11_95.2	GCTCTGGAGGATATTAAGCC
SMYD5_SET1_e11_95.9	CTGCATTGGGCACACAACTG
SMYD5_SET1_e2_73.5	GGGAGACCATCTTCGTAGAA
SMYD5_SET1_e2_73.6	CATCTTCGTAGAACGGCCCC
SMYD5_SET1_e2_73.8	CAGGGGCCGTTCTACGAAGA
SP100_BD_e27_72.1	AATCACAGTAGACCTTCAAG
SP100_BD_e27_72.2	CACAGTAGACCTTCAAGAGG
SP100_BD_e28_144.1	CAGATGTACACCCGAGTAGA
SP100_BD_e28_144.2	AGATGTACACCCGAGTAGAA
SP100_BD_e28_144.3	CCGAGTAGAAGGGTTTGTGC
SP100_BD_e29_79.0	GGAAGATAAATTCACCAGAC
SP100_BD_e29_79.3	CTGTACTTGAATTCCCAGTC
SP140_BD_e27_79.3	CCAAATGGGATTTAGACTGG
SP140L_BD_e19_73.2	CCAAATGGGACTTAGACTGG

SUV39H1_e3_105.1	TGCATCTTCCGCACGGATGA
SUV39H1_e3_105.10	AGCTTCGTCATGGAGTACGT
SUV39H1_e3_105.3	TTCCGCACGGATGATGGGCG
SUV39H1_e4_147.19	CGCCCTGACGGTCGTAGATC
SUV39H1_e4_147.4	GGGCCAGATCTACGACCGTC
SUV39H1_e4_147.5	GGCCAGATCTACGACCGTCA
SUV39H1_e5_130.10	GTGGCAAAGAAAGCGATGCG
SUV39H1_e5_130.11	TGCGGGGCAGCCGCTCGTCA
SUV39H1_e5_130.3	GCCACAAGAACCATCCGGGC
SUV39H2_e3_105.1	TTTCGAACTAGCAATGGACG
SUV39H2_e3_105.2	GAACTAGCAATGGACGTGGC
SUV39H2_e3_105.3	AACTAGCAATGGACGTGGCT
SUV39H2_e4_147.1	AGGACAGTTCTATGACAACA
SUV39H2_e4_147.3	AATCACGTATCTCTTTGATC
SUV39H2_e4_147.6	ACAGTGGATGCGGCTCGATA
SUV39H2_e5_130.0	TCCACAAGAACCATAAATGC
SUV39H2_e5_130.5	GTGGAAAACAATGCTATTCG
SUV39H2_e5_130.6	TTCGGGGAAGACGAGTATCG
SUV420H1_e6_45.0	CTCTTTTGTTGCAACTATTT
SUV420H1_e7_167.4	GAGAACATGCTACTTAGACA
SUV420H1_e7_167.5	TCAGTGTCATGTACTCCACA
SUV420H1_e7_167.7	AAACTGTGCTCAACTCTGGC
SUV420H1_e9_105.0	TCGAGATACAGCATGTGTGA
SUV420H1_e9_105.1	GCTCTAAGAGACATTGAACC
SUV420H1_e9_105.5	TATTATGGAGATGGGTTCTT
SUV420H2_e5_164.12	TCATGTACTCAACCCGCAAG
SUV420H2_e5_164.16	CATGGTTGATGAAGGCGGCT
SUV420H2_e5_164.18	CAGCTGAGCACTCCGCTTGC
SUV420H2_e7_105.13	GTCCCGGAGCACCTTCACGC
SUV420H2_e7_105.8	GGTGACATGCTTCTACGGCG
SUV420H2_e7_105.9	GTGACATGCTTCTACGGCGA
TAF1_BD1_e28_178.10	GGCGTTTACGCACGTTTTCG
TAF1_BD1_e28_178.11	CGTTTTCGCGGAGTGTTTGT
TAF1_BD2_e31_93.0	CAAAGTGATTGTCAATCCAA
TAF1_BD2_e32_85.3	GGATGATGTAAACCTTATTC
TAF1_BD2_e33_68.0	TCAGTATACTAAGACTGCCC
TAF1_BD2_e33_68.1	AACAGACGTTCACAATCTCC
TAF1_e13_174.15	GATCACCATCTTTGCCTGTG
TAF1_e15_201.1	CTTGATCATTCGGACAAGAC
TAF1_e15_201.13	AATATGCGTATTGGCCCTTT
TET1_e10_86.0	TCGAGAAGATAACCGCTCTT
TET1_e10_86.1	CGAGAAGATAACCGCTCTTT
TET1_e10_86.5	GCGGTTATCTTCTCGAGTTA
TET1_e7_86.2	GGATCAATTCTAAATCTTCT
TET1_e8_90.0	TGAAGATAACTTACAGAGTT
TET1_e8_90.1	GGAGCATACTGCTTATAAAT
TET1_e8_90.2	AACTCTGTAAGTTATCTTCA
TET1_e9_138.1	GTTGCCCGAGAATGTCGGCT
TET1_e9_138.17	TGCTGCCAAGCCGACATTCT
TET1_e9_138.18	GCTGCCAAGCCGACATTCTC
TET2_e10_86.0	AGAGAAGACAATCGAGAATT
TET2_e10_86.1	GAAGACAATCGAGAATTTGG
TET2_e10_86.3	ACGTGAAGCTGCTCATCCTC
TET2_e7_86.0	TCATGGAGCATGTACTACAA
TET2_e7_86.2	CCAAGGAAGTTTAAGCTGCT
TET2_e7_86.3	CAAGGAAGTTTAAGCTGCTT
TET2_e8_90.0	GCAAAACCTGTCCACTCTTA
TET2_e8_90.2	ATATGTTGGTGCCATAAGAG
TET2_e8_90.3	TTGGTGCCATAAGAGTGGAC
TET2_e9_138.0	CAGAGCACCAGAGTGCCGTC
TET2_e9_138.1	AGAGCACCAGAGTGCCGTCT
TET2_e9_138.14	TTCAGACCCAGACGGCACTC
TET3_e10_79.0	CAAGGAAGACAATCGCTGCG
TET3_e10_79.1	AAGGAAGACAATCGCTGCGT
TET3_e10_79.2	CTGCGTGGGCAAGATTCCCG
TET3_e7_86.1	TCAACGGCTGCAAGTATGCT
TET3_e7_86.3	CTCGCAAGTTCCGCCTCGCA
TET3_e7_86.4	TCGCAAGTTCCGCCTCGCAG
TET3_e8_90.2	CGCTCCCCTGTACAAGCGAC
TET3_e8_90.5	GAGGGGCCAGTCGCTTGTAC
TET3_e8_90.9	GTACAGGGGAGCGACTTCGG
TET3_e9_138.0	AATAGCGATTGACTGCCGTC
TET3_e9_138.1	ATAGCGATTGACTGCCGTCT
TET3_e9_138.2	TAGCGATTGACTGCCGTCTG
TIP60_e10_94.0	ATCAACGGAAGACTACAATG
TIP60_e10_94.13	GGGAGGCAGGGTTAGGATGC
TIP60_e10_94.5	TCAGCAGCTTGCCGTAGCCC
TIP60_e10_94.6	GCTTGCCGTAGCCCCGGCGC
TIP60_e10_94.8	CCGTAGCCCCGGCGCTGGTA
TRIM24_BD_e17_75.0	TTACTGCCATGAAATGAGCC
TRIM24_BD_e17_75.1	GAACAGGGTCTTGAAAAGCC
TRIM24_BD_e17_75.2	GAAAAGCCAGGCTCATTTCA
TRIM24_BD_e18_150.2	ATCTTCAGGTTTTGAGTACA
TRIM24_BD_e18_150.4	TTGATGGTTGACAAATCCAT
TRIM24_BD_e19_86.0	GATTCAGAAGTAGCCAATGC
TRIM24_BD_e19_86.1	TTCAAGTTTTATACCAGCAT
TRIM24_BD_e19_86.2	GCATTGGCTACTTCTGAATC
TRIM28_BD_e15_87.1	AGTCGGTAGCCAGCTGATGC
TRIM28_BD_e15_87.2	GTCGGTAGCCAGCTGATGCA
TRIM28_BD_e15_87.9	GCAGGGTTCGTGACAGAATA

TRIM28_BD_e16_138.19	CCTGGAGGCGGGCACGGATC
TRIM28_BD_e16_138.20	CTGGAGGCGGGCACGGATCA
TRIM28_BD_e16_138.7	GAACATGCGGCCCACATCCT
TRIM28_BD_e17_79.1	CTTCGAGACGCGCATGAACG
TRIM28_BD_e17_79.2	ACGCGCATGAACGAGGCCTT
TRIM28_BD_e17_79.3	GCGTCTCGAAGAAGCGCTGC
TRIM33_BD_e17_75.0	TGAATTAAGTATTGAATTCC
TRIM33_BD_e17_75.1	GGAATTCAATACTTAATTCA
TRIM33_BD_e17_75.2	TACTTAATTCATGGCAATAG
TRIM33_BD_e18_153.13	TTCACGGTGGATAAATCCAT
TRIM33_BD_e18_153.5	CTTGAAGATCAAACGGACAT
TRIM33_BD_e18_153.8	CCACAAAGTCATCCGGGATT
TRIM33_BD_e20_79.0	GATTCAGAAGTAGCTCAGGC
TRIM66_BD_e17_75.0	GGGGCTGACAGGTTCATGGA
TRIM66_BD_e17_75.1	GGGCTGACAGGTTCATGGAA
TRIM66_BD_e17_75.4	GCTGAGGTTATTGCAGCACA
TRIM66_BD_e18_153.17	GCTTCCTCCGGATGATTGAC
TRIM66_BD_e18_153.20	CCATGGGCCTCTTGATAATC
TRIM66_BD_e18_153.3	TGGACCTGTCAATCATCCGG
TRIM66_BD_e19_79.5	GTTCTTTGAGGGCTGGTTGA
TRIM66_BD_e19_79.6	CTGGTTGAAGGAGATCTACC
TRIM66_BD_e19_79.7	AGCCCTCAAAGAACACTTCC
TTF2_D_e10_120.6	GGATGAGCGCAATCATTGTC
TTF2_D_e11_151.10	CTGGCACGTGAATCCCGGTT
TTF2_D_e12_106.0	GATCACTACCTATAGCCTCG
TTF2_D_e13_158.13	ATCCAATATGATTCGAGCCC
TTF2_D_e20_151.8	GGGATTGACAGAGCCATCGA
TTF2_D_e21_83.3	GGTGTTGGTCTAAACCTGAC
TTF2_D_e22_75.1	GTGACCGAATTTACCGAGTA
UTY_e25_115.1	GTTAACATAAATATTGGTCC
UTY_e25_115.4	TGTTGTACCTGAAGATTATT
UTY_e25_115.5	GTTGTACCTGAAGATTATTG
UTY_e25_115.6	AGAACACCCCAATAATCTTC
UTY_e25_115.7	ACAAACCATTCACAATCTCC
ZMYND11_BD_e5_50.0	ATTCATTGTCTCCCGCATGA
ZMYND11_BD_e5_50.1	TTGTCTCCCGCATGAAGGAG
ZMYND11_BD_e6_93.3	ACAATAAACACCCGATGTAC
ZMYND11_BD_e6_93.5	ACACCCGATGTACAGGAGGC
ZMYND11_BD_e7_88.0	CACAATACCGTGATTTTCTA
ZMYND11_BD_e7_88.1	CATAGAAAATCACGGTATTG
ZMYND11_BD_e8_56.0	GTGAGCAAGCTGACATTGCG
ZMYND8_BD_e5_59.3	TCTGAATGGCAAACTTGAGC
ZMYND8_BD_e5_59.4	AATGGCAAACTTGAGCAGGT
ZMYND8_BD_e6_93.0	ATTCCAGAAGCCCGTTCCAT
ZMYND8_BD_e6_93.11	GTTCCAATGGAACGGGCTTC
ZMYND8_BD_e6_93.6	AAGATGTATTCCGCATAGTC
ZMYND8_BD_e7_88.1	TGGCTGCACAGAAGCCTTCC
ZMYND8_BD_e7_88.2	CCTTCCTGGCTGATGCAAAG
ZMYND8_BD_e7_88.6	AAATCCACTTTGCATCAGCC
Rosa26	GAAGATGGGCGGGAGTCTTC
neg01	GTAGCGAACGTGTCCGGCGT
neg02	GACCGGAACGATCTCGCGTA
neg03	GGCAGTCGTTCGGTTGATAT
neg04	GCTTGAGCACATACGCGAAT
neg05	GTGGTAGAATAACGTATTAC
neg06	GTCATACATGGATAAGGCTA
neg07	GATACACGAAGCATCACTAG
neg08	GAACGTTGGCACTACTTCAC
neg09	GATCCATGTAATGCGTTCGA
neg10	GTCGTGAAGTGCATTCGATC
neg11	GTTCGACTCGCGTGACCGTA
neg12	GAATCTACCGCAGCGGTTCG
neg13	GAAGTGACGTCGATTCGATA
neg14	GCGGTGTATGACAACCGCCG
neg15	GTACCGCGCCTGAAGTTCGC
neg16	GCAGCTCGTGTGTCGTACTC
neg17	GCGCCTTAAGAGTACTCATC
neg18	GAGTGTCGTCGTTGCTCCTA
neg19	GCAGCTCGACCTCAAGCCGT
neg20	GTATCCTGACCTACGCGCTG
neg21	GTGTATCTCAGCACGCTAAC
neg22	GTCGTCATACAACGGCAACG
neg23	GTCGTGCGCTTCCGGCGGTA
neg24	GCGGTCCTCAGTAAGCGCGT
neg25	GCTCTGCTGCGGAAGGATTC
neg26	GTCTTCCGCCGCTCAAGTTA
neg27	GCATGGAGGAGCGTCGCAGA
neg28	GTAGCGCGCGTAGGAGTGGC
neg29	GATCACCTGCATTCGTACAC
neg30	GCACACCTAGATATCGAATG
neg31	GTTGATCAACGCGCTTCGCG
neg32	GCGTCTCACTCACTCCATCG
neg33	GCCGACCAACGTCAGCGGTA
neg34	GGATACGGTGCGTCAATCTA
neg35	GAATCCAGTGGCGGCGACAA
neg36	GCACTGTCAGTGCAACGATA
neg37	GCGATCCTCAAGTATGCTCA
neg38	GCTAATATCGACACGGCCGC
neg39	GGAGATGCATCGAAGTCGAT
neg40	GGATGCACTCCATCTCGTCT

neg41	GTGCCGAGTAATAACGCGAG
neg42	GAGATTCCGATGTAACGTAC
neg43	GTCGTCACGAGCAGGATTGC
neg44	GCGTTAGTCACTTAGCTCGA
neg45	GTTCACACGGTGTCGGATAG
neg46	GGATAGGTGACCTTAGTACG
neg47	GTATGAGTCAAGCTAATGCG
neg48	GCAACTATTGGAATACGTGA
neg49	GTTACCTTCGCTCGTCTATA
neg50	GTACCGAGCACCACAGGCCG
PCNA_e2.1	GGACTCGTCCCACGTCTCTT
PCNA_e2.2	CTACCGCTGCGACCGCAACC
PCNA_e3.1	GCCGGCGCATTTTAGTATTT
PCNA_e3.2	CGAAGATAACGCGGATACCT
POLR2D_e2.1	TGAGAGTGCAGAGGACGAAC
POLR2D_e3.1	TGGGCAAAGGTTGGCCAAAC
POLR2A_e2.1	AAGCGAATGTCTGTGACGGA
POLR2A_e2.2	CAGGGGGTGATTGAGCGGAC
POLR2A_e10.1	GTACAATGCAGACTTTGACG
POLR2A_e10.2	TGGGGGGTGACAATCATGCG
POLR2A_e27.1	AGAGATTCCACCCATGGGAC
POLR2A_e27.2	CAACCCCTGCCTATGGCGCC
POLR2A_e28.1	GCTGGCATCTGACGCAGCAC
POLR2A_e28.2	GTAACCTGGGCTGAAGCCGC
RPL9_e3.1	GGACGCACAGTTATCGTGAA
RPL9_e3.2	AATGTAGAACTCAGCCTTCT
RPL9_e4.1	CTCCGGGTTGACAAATGGTG
RPL9_e4.2	GAAAGGAACTGGCTACCGTT
RPL23A_e2.1	TGCGGATCTTCTTCTTTTTG
RPL23A_e2.2	GTCGCAGTGTCTTCGGCCGC
RPL23A_e3.1	GAAGCTGTATGACATTGATG
CDK9_e2.1	CTGGCCGGTCTTGCGGTGCC
CDK9_e3.1	GATCTCCCGCAAGGCTGTAA
CDK9_e4.1	TTCCCCCTATAACCGCTGCA
CDK9_e4.2	GCTCGCAGAAGTCGAACACC
CDK9_e5.1	ATGTGCTTATCACTCGTGAT
CDK9_e5.2	CAACCGCTACACCAACCGTG
CDK1_e3.1	GGGTTCCTAGTACTGCAATT
CDK1_e4.1	AATCCATGTACTGACCAGGA
CDK1_e5.1	ACACAATCCCCTGTAGGATT
CDK1_e6.1	ACTCAACTCCAGTTGACATT
CDK1_e7.1	ACATGGGATGCTAGGCTTCC
RPA3_e1.1	CCGGCGTTGATGCGCGACCT
RPA3_e1.2	GCCGGCGTTGATGCGCGACC
RPA3_e1.3	GATGAATTGAGCTAGCATGC
RPA3_e2.1	AAATGGAACCATCGAGTTGA

Table 6: sgRNA sequences for pooled epigenome screen adopted from *Shi et al.* [79] and Hörmann *et al.* [317]

6.6 Alpha RRA scores

Gene symbol	KYSE-270	T.T	KYSE-30	KYSE-410	KYSE-510	COLO-680N	KYSE-150	KYSE-70	KYSE-140	KYSE-450
ACAT1	-0,003	-0,843	-0,015	-0,020	-0,018	-0,004	-0,011	-0,033	-0,007	-0,005
ASH1L	-0,272	-0,834	-1,192	-0,223	-1,933	-0,410	-1,149	-0,573	-0,467	-0,199
ATAD2	-0,938	-0,064	-1,178	-0,323	-1,395	-1,577	-0,242	-1,178	-0,099	-0,543
ATAD2B	-0,001	-0,067	-0,027	-0,496	-0,013	-0,068	-0,286	-0,029	-0,038	-0,048
ATAT1	-0,616	0,000	-0,826	-0,536	-0,543	-0,154	-2,004	-0,997	-0,083	-1,142
ATF2	-0,137	-0,125	-0,216	-0,419	-0,396	-0,350	-0,478	-0,411	-0,105	-0,402
ATRX	-0,207	-0,848	-2,790	-0,529	-0,562	-0,367	-0,957	-2,076	-0,906	-0,123
BAZ1A	-0,147	-0,240	-0,250	-0,368	-0,278	-0,478	-0,012	-0,090	-0,809	-0,038
BAZ1B	-2,535	-1,864	-1,932	-3,072	-1,025	-1,330	-2,999	-1,058	-2,509	-1,073
BAZ2A	-0,147	-0,233	-0,756	-0,097	-0,011	-0,003	-0,960	-0,644	-0,359	-0,850
BAZ2B	-0,002	-0,010	-0,012	-0,265	-0,079	-0,584	-0,023	-0,006	-0,020	-0,102
BPTF	-0,415	-0,815	-0,919	-0,314	-0,827	-2,484	-1,993	-1,438	-1,359	-1,784
BRD1	-0,419	-0,005	-0,439	-2,158	-0,764	-0,296	-0,136	-1,014	-0,084	-1,256
BRD2	-5,107	-3,585	-6,499	-5,560	-2,895	-3,388	0,000	-0,677	-1,207	-6,602
BRD3	-0,024	-0,597	-2,037	-0,328	-0,176	-0,512	-0,541	-0,582	0,000	-0,301
BRD4	-8,837	-6,992	-5,802	-7,376	-6,403	-6,625	-8,419	-6,388	-10,511	-9,624
BRD8	-3,744	-2,881	-1,672	-0,963	-2,563	-2,543	-2,167	-0,906	-3,280	-2,397
BRD9	-0,017	-0,094	-0,444	-1,401	-1,263	-0,032	-2,084	-0,009	-0,441	-0,188
BRDT	-0,528	-0,003	-0,117	-0,128	-0,073	-1,080	-0,373	0,000	-0,212	-0,028
BRPF1	-0,303	-0,580	-0,490	-1,057	-1,084	-0,244	-1,174	-0,365	-0,502	-0,227
BRPF3	-0,512	-0,049	-0,018	-0,132	-0,022	-0,055	-0,490	-0,004	-0,037	-0,014
BRWD1	-0,017	-0,021	-0,013	-0,045	-0,016	-0,004	-0,002	-0,035	-0,001	0,000
BRWD3	0,000	-0,009	-0,002	-0,079	0,000	-0,187	0,000	-0,035	0,000	-0,002
CARM1	-1,889	-1,541	-0,753	-0,310	-1,409	-0,459	-0,433	-2,089	-2,665	-1,601
CDK1	-2,590	-1,645	-2,119	-2,002	-2,293	-3,093	-2,484	-2,172	-2,779	-2,501
CDK9	-3,135	-3,885	-2,076	-1,921	-3,201	-4,661	-4,004	-4,172	-4,790	-3,818
CECR2	-0,002	-0,077	-0,050	-0,830	-0,084	-0,256	-0,674	-0,302	-0,161	-0,066
CHD1	-1,384	-0,192	-0,602	-0,392	-0,612	-0,118	-1,684	-0,025	-0,606	-0,967
CHD1L	-0,602	-3,012	-1,432	-0,401	-0,768	-0,471	-0,158	-0,404	-1,611	-0,255
CHD2	-0,389	-0,818	-0,005	-0,070	-1,601	-0,012	-1,557	-0,216	-0,127	-0,216
CHD3	-0,003	-0,001	-0,045	-0,662	-0,451	-0,504	-0,003	0,000	-0,064	-0,090
CHD4	-3,295	-2,314	-1,649	-2,910	-3,573	-3,851	-5,017	-3,842	-4,116	-4,279
CHD5	-0,240	-0,341	-0,055	-0,284	-0,452	-0,267	-0,167	-0,183	-0,226	-0,164
CHD6	-0,023	-0,036	-0,210	-0,306	-0,152	-0,339	-1,257	-0,021	-0,408	-0,029
CHD7	-0,145	-0,008	-0,089	-0,016	-0,005	-1,342	-0,250	-0,003	-0,057	-0,042
CHD8	-1,629	-1,314	-1,314	-2,358	-2,324	-0,937	-1,843	-0,802	-2,399	-2,569
CHD9	-0,080	-0,069	-0,213	-0,028	-0,009	-0,715	-0,169	-0,178	-0,126	-0,010
CLOCK	-0,007	-0,188	-0,084	-0,003	-0,142	-0,309	-0,118	-0,001	-0,007	-0,004
CREBBP	0,000	-1,247	-2,782	-0,038	-0,040	-0,240	-0,112	0,000	-0,013	0,000
DOT1L	-4,878	-4,189	-0,886	-1,879	-0,498	-0,537	-0,990	-2,881	-0,019	-0,595
EHMT1	-0,004	-3,178	-1,554	-0,981	-3,021	-0,570	-0,740	-2,808	-5,168	-2,546
EHMT2	-1,189	-3,732	-1,800	-1,089	-2,981	-0,891	-1,157	-2,063	-5,097	-3,402
ELP3	-4,432	-5,243	-4,329	-4,943	-5,205	-5,182	-5,145	-5,158	-6,331	-5,243
EP300	-1,886	-6,650	-0,520	-2,531	-0,200	-4,735	-2,349	-7,164	0,000	-7,846
EP400	-2,885	-2,722	-2,820	-1,195	-2,185	-2,043	-3,755	-2,121	-3,737	-2,831
ERCC6	-0,231	-0,205	-1,206	-0,192	-0,350	-0,791	-0,404	-0,009	-0,005	-0,228
EZH1	-1,071	-0,258	-0,509	-0,249	-0,284	-0,905	-0,061	-0,185	-0,481	-0,273
EZH2	0,000	-0,169	-0,041	-0,211	-1,477	-2,303	-0,027	-1,640	-0,010	-0,018
GTF3C4	-4,926	-4,134	-3,220	-4,517	-4,901	-2,041	-5,428	-3,175	-6,130	-5,323
HAT1	-0,011	-0,215	-0,008	-0,218	-1,099	-0,349	-0,323	-0,022	-0,001	-0,401
HDAC1	-2,008	-2,651	-2,874	-0,123	-0,211	-0,676	-1,814	-2,018	-1,163	-1,054
HDAC10|MAPK12	-0,003	-0,004	-0,152	-0,123	-0,136	-0,203	-0,043	-0,428	-0,018	-0,070
HDAC11	-0,330	-0,232	-0,963	-0,663	-0,760	-0,603	-0,066	-0,034	-0,138	-0,614
HDAC2	-0,016	-0,156	-0,464	-0,764	-0,337	-0,116	-0,229	-0,752	-0,227	-0,130
HDAC3	-1,752	-4,714	-4,292	-5,680	-2,966	-3,342	-4,279	-5,555	-4,868	-3,922
HDAC4	-0,014	-0,367	-0,286	-0,342	0,000	-0,004	0,000	-0,091	-0,017	-0,001
HDAC5	-0,047	-0,501	-0,140	-0,212	-0,148	-1,580	-0,243	-0,483	-0,436	-0,614
HDAC6	-0,567	-1,210	-0,128	-0,132	-0,095	-0,137	-1,995	-0,940	-0,055	-0,204
HDAC7	-1,152	-0,574	-0,571	-0,476	-0,045	-0,494	-1,390	-3,121	-0,535	-0,633
HDAC8	-0,655	-1,405	-0,617	-0,221	-0,209	-0,010	-0,305	-0,831	-3,524	-0,032
HDAC9	-0,163	-0,002	-0,239	-0,014	-0,009	-0,010	-0,312	-0,003	-0,136	-0,001
HELLS	-0,060	-0,042	-0,403	-0,001	-0,166	-0,010	-0,215	-0,766	-0,039	-0,102
HIF1AN	-1,020	-0,737	-0,757	-0,507	-0,692	-0,872	-0,891	-1,340	-0,288	-0,768
HLTF	-0,130	-0,064	-0,009	-0,162	-0,236	-0,392	-0,014	-0,070	-0,015	-0,120
HR	-0,108	-0,051	-0,089	-0,315	-0,029	-0,084	-0,005	-1,057	-0,286	-0,153
HSPBAP1	-0,003	-0,021	0,000	-0,152	-0,015	-0,174	-0,173	-0,042	-0,003	0,000
INO80	-3,391	-2,889	-3,617	-3,234	-2,828	-2,981	-4,891	-3,983	-4,550	-3,964
JARID2	-0,684	-0,065	-0,064	-0,415	-0,268	-0,226	-0,090	-0,001	-0,005	-0,196
JMJD1C	-0,001	-0,002	-0,011	-0,036	-0,019	-0,005	-0,002	-0,062	-0,027	-0,001

JMJD6	-1,826	-1,202	-1,441	-0,878	-1,269	-1,297	-1,283	-1,077	-1,759	-0,874
JMJD7|JMJD7-PLA2G4B	-0,786	-0,009	-0,512	-0,525	-0,710	-0,397	-0,954	-0,695	-0,440	-0,661
KAT2A	-3,733	-3,411	-1,796	-3,032	-5,078	-4,390	-2,395	-4,307	-3,502	-1,577
KAT2B	-0,003	-0,075	-0,046	-0,079	0,000	-0,048	0,000	-0,001	-0,004	-0,091
KAT5|RNASEH2C	-3,962	-2,946	-2,159	-4,002	-3,316	-3,658	-4,270	-4,145	-4,356	-4,642
KAT6A	-0,018	-0,115	-0,410	-0,069	-0,645	-0,153	-0,573	0,000	-0,650	-0,755
KAT6B	-0,180	-0,227	-0,115	-0,272	-0,035	-0,401	-0,444	-0,018	-0,004	-0,130
KAT7	-0,403	-1,093	-1,218	-1,671	-2,499	-1,572	-1,802	-3,369	-1,098	-1,590
KAT8	-2,805	-2,555	-2,586	-3,535	-3,291	-1,673	-3,367	-3,134	-3,702	-3,754
KDM1A	-2,836	-3,127	-2,885	-5,163	-4,051	-0,825	-1,314	-1,062	-0,848	-2,748
KDM1B	-0,226	-0,002	-0,184	-0,001	0,000	-0,199	-0,170	0,000	-0,012	-0,001
KDM2A	-0,087	-3,580	-1,433	-2,349	-1,965	-0,518	-0,055	-0,105	-1,854	-1,767
KDM2B	-0,265	-1,463	-0,333	-1,934	-1,228	-0,081	-0,874	-0,162	-0,698	-1,283
KDM3A	-0,687	-0,392	-0,037	-0,954	-0,258	-0,167	-0,011	-0,101	-0,037	-0,105
KDM3B	-1,976	-2,480	-1,890	-1,881	-1,809	-0,950	-2,671	-2,138	-2,191	-1,465
KDM4A	-2,746	-0,714	-0,699	-0,175	-0,019	-0,045	-0,085	-1,577	-1,281	-0,860
KDM4B	-0,362	-0,582	-0,028	-0,091	-0,280	-0,492	-0,166	-0,625	-0,035	-0,449
KDM4C	0,000	-0,271	-0,243	-0,088	-0,149	-0,108	-5,411	-0,097	-0,001	-0,203
KDM4D	-0,009	-0,380	-0,642	-0,004	-0,186	-1,343	-0,083	-0,069	-0,895	-0,280
KDM5A	-0,065	-0,198	-0,199	-0,028	-0,060	-0,001	-0,287	-0,006	-0,815	-1,157
KDM5B	-0,178	-0,936	-0,404	-1,023	-0,640	-0,086	-0,437	-0,003	-0,205	-0,637
KDM5C	-0,051	-0,002	-0,045	-0,254	-0,260	-0,007	-0,505	-0,045	-0,028	-0,002
KDM5D	-0,005	-0,108	-0,011	-0,028	-0,010	-0,177	0,000	-0,008	-0,108	-0,004
KDM6A	-0,397	-0,238	-0,884	-0,004	-0,716	-0,610	-0,003	-1,259	-0,011	-0,862
KDM6B	-0,476	-0,233	-0,555	-0,187	-0,295	-1,977	-0,458	-0,055	-0,574	-0,535
KDM7A	-0,050	-0,001	-0,303	-0,428	-0,045	-0,778	-0,125	-0,166	-0,037	-0,021
KDM8	-4,028	-3,780	-2,621	-3,844	-3,968	-2,762	-4,243	-4,741	-4,739	-5,169
KMT2A	-0,475	-0,025	-0,299	-0,638	-0,201	-0,050	-0,227	-2,535	-0,043	-0,168
KMT2B|IGFLR1	-0,725	-2,612	-0,738	-1,904	-2,847	-1,428	-0,324	-0,742	-1,928	-1,000
KMT2C	-0,029	-0,037	-0,482	-0,026	-0,241	-1,118	0,000	-0,648	-0,023	-1,877
KMT2D	-1,805	-1,228	-1,311	-0,384	-0,569	-1,691	-0,968	-2,831	-1,449	-1,076
KMT2E	0,000	-0,065	-0,428	-0,694	-0,067	-0,047	0,000	-0,402	-0,259	-1,174
KMT5A	-2,507	-2,397	-2,289	-3,259	-2,496	-1,809	-2,562	-2,671	-3,226	-2,630
KMT5B	-0,814	-2,770	-1,297	-0,035	-0,723	-0,344	-0,035	-0,190	-1,563	-0,070
KMT5C	-0,162	-0,005	-0,020	-0,453	-3,207	-0,067	-0,081	-0,070	-0,031	-0,100
NAT10	-2,260	-2,100	-1,785	-2,110	-1,893	-1,795	-2,874	-2,516	-3,516	-2,445
NCOA1	-0,381	-0,109	-0,209	-0,510	-0,663	-0,217	-0,359	-0,008	-0,077	-0,374
NCOA2	-0,022	-0,002	-0,080	-0,646	-0,524	-0,529	-0,605	-0,068	-0,003	-0,080
NCOA3	-0,578	-0,532	-0,741	-2,576	-0,401	-0,655	-0,416	-0,463	-0,545	-0,097
NEG_CONTR	0,000	0,000	0,000	-0,008	-0,017	0,000	0,000	0,000	0,000	0,000
NSD1	-0,979	-1,998	-5,045	-2,751	-3,817	-2,697	-4,743	-2,170	-3,123	-3,573
NSD2	-0,635	-0,569	-0,005	-0,924	-2,289	-0,049	-0,657	-1,932	-1,048	-0,583
NSD3	-0,007	0,000	-0,461	-0,011	-0,048	-0,453	-0,013	-0,009	-0,366	-0,402
PADI4	-0,171	-0,080	-0,291	-0,897	-0,047	-0,370	-0,566	-0,726	-0,246	-0,138
PBRM1	-0,725	-2,828	-1,495	-0,232	-0,569	-0,786	-0,005	-2,766	-0,475	-0,060
PCNA	-1,919	-1,433	-1,355	-1,007	-1,848	-1,755	-2,247	-1,781	-2,208	-2,280
PHF2	-0,247	-0,124	-1,222	-0,462	-0,246	-1,940	-1,387	-0,002	-1,397	-0,222
PHF8	-0,019	-0,043	-0,020	-0,173	-0,282	-0,010	-0,004	-0,304	-0,001	-0,383
PHIP	-2,422	-0,020	-0,303	-0,068	-0,111	-0,105	-0,203	-0,233	-0,037	-2,125
POLR2A	-1,417	-3,624	-2,717	-2,604	-3,880	-1,728	-4,307	-4,354	-4,577	-3,530
PRDM2	-0,023	-0,492	-0,356	-0,006	-1,561	-0,148	-0,150	-0,206	-0,102	-0,019
PRDM8	-0,313	-0,275	-0,677	-0,248	-0,032	-0,223	-0,302	-0,064	-0,072	-0,330
PRDM9	-0,160	-0,031	-0,020	-0,308	-0,297	-0,998	-0,080	-0,899	-0,267	-0,003
PRMT1	-2,082	-2,148	-2,055	-1,098	-2,198	-2,282	-2,471	-2,508	-2,856	-3,316
PRMT2	-0,722	-0,217	-0,954	-0,277	-0,555	-0,647	-0,019	-0,591	-0,443	-1,227
PRMT3	-0,174	-0,072	-0,129	-0,008	-0,207	-0,032	-0,124	-0,455	-0,001	-0,023
PRMT5	-3,618	-3,376	-1,276	-4,929	-4,824	-3,768	-5,063	-4,537	-5,017	-5,704
PRMT7	-0,997	-1,391	-2,587	-0,539	-0,011	-2,039	-1,104	-2,316	-1,355	-0,764
PRMT8	-0,034	-0,020	-0,161	-0,214	-0,300	-0,166	-0,092	-0,002	-0,003	-0,003
PRMT9	-0,068	-0,535	-0,010	-0,296	-1,968	-0,757	-0,217	-0,799	-0,367	-0,107
RAD54L2	-1,527	-0,684	-0,676	-0,046	0,000	-1,826	-0,329	-0,024	-2,434	-0,457
RIOX1|HEATR4	-0,525	-0,025	-0,350	-0,015	-0,395	-0,289	-0,500	-0,365	-0,155	-0,374
RIOX2	-0,004	-0,013	-0,005	-0,373	-0,116	-0,024	-0,112	-0,183	-0,137	-0,921
RPA3	-2,629	-1,919	-2,529	-2,910	-2,713	-1,587	-3,070	-1,816	-3,055	-3,328
RPL23A	-1,825	-2,200	-1,480	-1,331	-1,639	-1,352	-1,639	-1,500	-1,792	-1,888
RPL9	-3,234	-2,284	-2,004	-3,070	-2,852	-2,872	-3,242	-2,775	-3,371	-3,692
SETD1A	-4,275	-2,824	-2,456	-2,035	-0,009	-1,494	-4,257	-2,198	-3,788	-2,605
SETD1B	-6,903	-1,039	-2,659	-3,035	-5,593	-0,650	-4,309	-1,520	-0,388	-6,075
SETD2	-2,463	-3,600	-3,012	-3,711	-1,602	-2,853	-2,975	-2,934	-3,557	-1,849
SETD4	-0,187	-0,080	-0,302	-0,232	-0,345	-0,052	-0,484	-0,006	-0,257	-0,010
SETD5	-0,002	-0,120	-0,080	-0,461	-0,407	-0,589	-0,010	-0,206	-0,016	-0,388
SETD6	-0,234	-0,144	-0,035	-0,371	-1,478	-0,818	-0,843	-0,051	-0,170	-0,233
SETD7	-0,830	-0,136	-0,721	-0,050	-0,615	-0,418	-0,216	-0,161	-0,276	-0,273

SETDB1|CERS2	-2,353	-3,497	-1,874	-0,503	-2,343	-2,408	-1,468	-3,448	-2,854	-0,063
SETDB2	-0,005	-0,184	-0,216	-0,027	-0,013	-0,003	-0,005	-2,008	-0,008	-0,188
SETMAR	-0,004	-0,012	-0,001	-0,229	-0,031	-0,009	-0,032	-0,585	-0,008	-0,056
SHPRH	-0,227	-0,311	-0,495	-0,689	-0,727	-0,245	-0,344	-0,413	-0,512	-0,187
SIRT1	-0,586	-0,008	-1,540	-0,007	-0,004	-0,132	-1,587	-0,439	0,000	-0,357
SIRT2	-0,020	-0,311	-0,160	-1,620	-0,105	-1,858	-0,221	-0,221	-0,450	-0,156
SIRT3	-0,049	-0,134	-0,031	-0,240	-0,180	-0,003	-0,016	-0,175	-0,006	-0,016
SIRT4	-0,864	-0,208	-0,614	-0,580	-0,232	-0,465	-0,229	-0,601	-0,135	-0,473
SIRT5	-0,018	-0,193	-0,798	-0,812	-0,339	-0,328	-0,057	-0,004	-0,080	-0,254
SIRT6	-1,279	-1,586	-1,160	-0,686	-0,610	-0,207	-0,273	-0,774	-2,051	-1,533
SIRT7	-0,079	-0,655	-1,239	-0,519	-0,025	-0,308	-0,012	-0,058	-0,862	-0,264
SMARCA2	-0,176	-0,593	-0,111	-0,497	-0,859	-0,232	-0,114	-0,674	-0,289	-0,468
SMARCA4	-4,503	-4,499	-3,724	-3,108	-3,397	-1,322	-1,640	-1,215	-1,310	-0,489
SMARCAD1	-0,040	-0,011	-0,137	-0,107	-0,368	-0,813	-0,010	-0,002	-0,001	-0,332
SMARCAL1	-0,378	-0,009	-0,266	-0,403	-0,023	-0,734	-0,137	-0,059	-0,152	-0,585
SMYD1	-0,023	-0,544	-0,086	-0,058	-0,041	-0,611	-0,312	-0,078	-0,054	-0,013
SMYD3	-0,001	-0,738	-0,111	-0,202	-0,166	-0,281	-0,525	-0,248	-0,005	-0,083
SMYD4	-0,012	-0,008	-0,183	-0,305	-0,195	-0,326	-0,197	-0,039	-0,024	-0,033
SMYD5	-0,283	-0,011	-0,013	-0,066	-0,009	-0,018	-0,088	-0,142	-0,019	-0,192
SNAP47|JMJD4	-1,202	-1,000	-1,351	-1,113	-1,463	-0,160	-0,131	-0,539	-0,343	-0,765
SP100	-0,036	-0,069	-0,034	-0,019	-0,001	-0,096	-0,218	-0,033	-0,061	-0,042
SRCAP	-1,859	-2,275	-0,514	-1,810	-3,398	-3,842	-2,993	-2,686	-3,419	-3,315
STUB1|JMJD8	-0,192	-0,100	-0,162	-0,211	-1,582	-0,064	-0,091	-0,075	-0,315	-0,139
SUV39H1	-1,046	-0,093	-0,123	-0,234	-0,196	-0,167	-0,050	-0,496	-0,018	-0,108
SUV39H2	-0,104	-0,007	-0,092	-0,075	-0,092	-0,089	-0,227	-0,200	-0,032	-0,698
TAF1	-2,143	-2,875	-2,043	-2,467	-6,026	-3,060	-3,419	-4,009	-6,138	-3,760
TET1	-0,059	-0,002	-0,003	-0,098	-0,001	-0,071	-0,083	-0,007	-0,026	-0,011
TET2	-0,003	0,000	-0,045	-0,018	-0,288	-0,140	-0,017	0,000	-0,022	-0,006
TET3	-0,745	-0,815	-0,491	-0,583	-0,328	-0,076	-1,192	-0,735	-1,531	-2,386
TRIM24	-0,001	-0,056	-0,004	-0,003	0,000	-0,001	-0,047	-0,038	0,000	-0,138
TRIM28	-3,398	-1,127	-1,471	-2,964	-1,698	-0,960	-0,919	-3,172	-1,059	-1,464
TRIM33	-0,065	-3,485	-0,127	-0,126	-0,004	-0,331	-0,309	-0,030	-0,012	0,000
TRIM66	0,000	-0,019	-0,202	-0,005	-0,004	-0,120	-0,091	-0,085	-0,009	-0,043
TTF2	-4,361	-3,388	-3,457	-3,293	-4,285	-3,475	-2,493	-4,107	-5,050	-4,936
UTY	-0,001	-0,007	-0,013	-0,119	-0,036	0,000	-0,004	-0,003	-0,043	-0,016
ZMYND11	-0,079	-0,004	-0,004	-0,005	-0,067	-0,001	-0,202	-0,003	-0,143	-0,036
ZMYND8	-0,700	-0,091	-0,018	-1,486	-0,017	-0,062	-1,070	-0,081	-1,165	-0,026

Table 7: α-RRA scores obtained from CRISPR-screens

6.7 Chromatogram

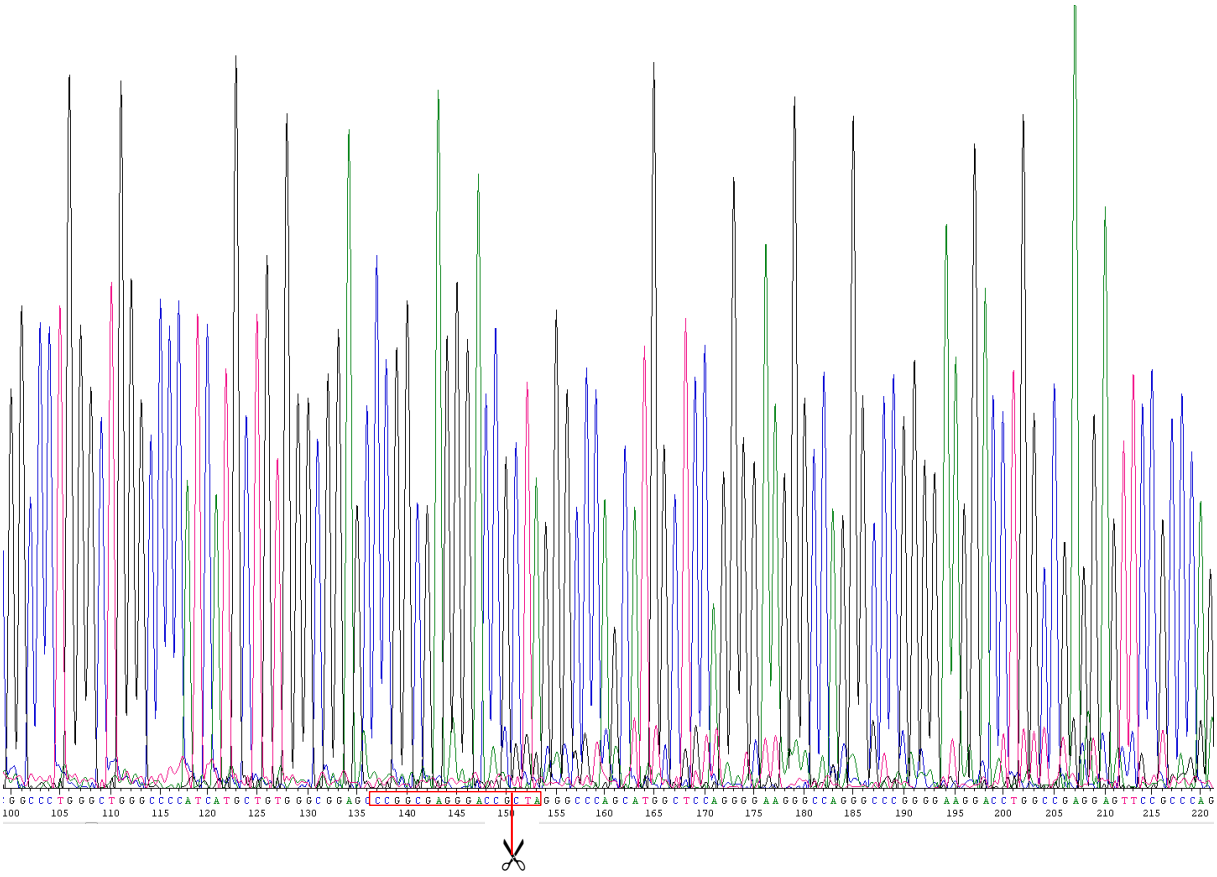


Figure 55: Chromatogram for endogenous SMARCA4-KO obtained by Sanger sequencing in KYSE-30-SWAP cell line.

Sequence obtained by reverse sequencing. Double strand break is introduced three nucleotides away from the PAM sequence (NGG). The sgRNA targeted sequence is indicated by the red rectangle. Sufficient editing is indicated by the separation of the two sequenced strands obtained, starting at the introduced cut site.

6.8 Abstract

Esophageal squamous cell carcinoma (ESCC) is the most prevalent type of esophageal cancer which is the sixth leading cause of deaths due to cancer worldwide, with a five-year survival rate of 15-25%. Therapy options are limited due to the aggressive nature of the disease and late diagnosis, accompanied with the presence of distant metastasis. Next generation sequencing has allowed the identification of gene mutations in diverse cancer subtypes thereby defining oncogenes as well as tumor suppressor genes. As a result, patients with an oncogenic “driver” mutation can therefore be treated by targeted therapy, exploiting the protein alteration. In contrast, reconstituting a tumor suppressor is (currently) therapeutically not feasible. However, this can be circumvented by harnessing the concept of synthetic lethality. By inhibiting the functional redundant partner, tumor growth can be impaired whereby sparing other cell types of the human body. In wild-type cells, the synthetic lethal factor is still present, which compensates for the loss of the individual partner. CRISPR-Cas9 screens targeting specific domains within the epigenome, were conducted with the aim to identify novel targets in ESCC. SMARCA4, the catalytic subunit of the chromatin remodeling complex SWI/SNF (BAF), was identified as a strong hit which exclusively scored in cell lines with SMARCA2^{low} (mutual exclusive ATPase) background indicating synthetic lethal dependencies. Rescue experiments confirmed the importance of the ATPase-domain of SMARCA4 for cell survival. The paralog dependency could be highlighted by either re-expression of SMARCA2 in SMARCA2^{low} cell lines, reverting the dependency, or by knocking-out SMARCA2 in SMARCA2-proficient cell lines, transforming them into SMARCA4-dependent ones. In addition, cell lines from different indications including colon, pancreas, and ovarian carcinoma were identified as SMARCA2^{low} and SMARCA4-dependent, which might allow for extension of the therapeutic concept to additional indications with a defined patient selection biomarker (SMARCA2^{low}). To date, no selective SMARCA4 inhibitors are available. Therefore, a domain-swap strategy was used, exchanging the bromodomain of SMARCA4, with the one of BRD9. This allowed pharmacological inhibition of chimeric SMARCA4 by using a BRD9 bromodomain directed PROTAC (dBRD9). Targeted degradation as well as impaired viability and thereby confirms SMARCA4 as a potential target in SMARCA2^{low} ESCC.

6.9 Zusammenfassung

Das Plattenepithelzellkarzinom des Oesophagus (ESCC) ist die dominant auftretende Form des Oesophaguskarzinoms, welches zu den sechst häufigsten durch Krebs hervorgerufenen Todesursachen weltweit zählt. Die Überlebensrate über fünf Jahre liegt zwischen 15-25%. Die Therapiemöglichkeiten sind aufgrund der hohen Aggressivität und der späten Diagnose limitiert. Moderne Sequenzierungsmethoden („next generation sequencing“) machten es möglich neue „Onkogene“ und „Tumorsuppressorgene“ zu identifizieren. Patienten mit „Driver-Mutationen“ können deshalb mit zielgerichteter Therapie behandelt werden. Derzeit ist es therapeutisch (noch) nicht möglich, ein Tumorsuppressorgen wiederherzustellen. Um diese Limitierung zu umgehen, nutzt man das Prinzip der „Synthetischen Letalität“. Bei der Inhibierung des funktional redundanten Partners kann das Tumor Wachstum inhibiert, gesunde Zellen aber verschont werden. In Wildtyp Zelllinien ist der „synthetisch letale“ Partner präsent, welcher den Verlust des individuellen Partners kompensieren kann. Das Ziel war, mit Hilfe von „CRISPR-Cas9-Screens“ basierend auf Domänen gerichtete Inhibierung des Epigenoms neue „Targetgene“ im ESCC zu identifizieren. SMARCA4, die katalytische Untereinheit des „Chromatin Remodeling Komplexes“ SWI/SNF (BAF) wurde als ein dominanter „Hit“ ausschließlich in SMARCA2-niedrig (redundante ATPase) exprimierenden Zelllinien identifiziert. Die Abhängigkeit war auf synthetisch letale Interaktion zurückzuführen. Durch „Rescue“-Experimente wurde die Bedeutung der ATPase-Domäne von SMARCA4 gezeigt. Die paraloge Abhängigkeit konnte sowohl mittels Re-expression von SMARCA2 in einer SMARCA2-niedrig exprimierenden Zelllinie und folgender Reversion der Abhängigkeit, als auch durch SMARCA2 „knock-out“ in einer SMARCA2-exprimierenden Zelllinie und Generierung einer SMARCA4 Abhängigkeit, analysiert werden. Um das therapeutische Konzept mit einem definierten Biomarker auszudehnen, wurde die Abhängigkeit von SMARCA4 auch in zusätzlichen Krebsindikationen (Kolon-, Pankreas- und Ovarialkarzinomen) mit geringer SMARCA2 Expression gezeigt. Ein „Domänen-SWAP“ durch Austausch der SMARCA4-Bromodomäne mit der BRD9-Bromodomäne, wurde verwendet um SMARCA4 pharmakologisch mit einem BRD9-PROTAC (dBRD9) zu degradieren. Die Anwendung dieses BRD9 protein-degradierenden Moleküls inhibierte die Proliferation und bestätigt SMARCA4 als einen potentiellen Angriffspunkt in ESCC.