Effect of CRISPR-based epigenetic interference on cis-regulatory gene networks

Loh, Xinpeng

2018

Loh, X. (2018). Effect of CRISPR-based epigenetic interference on cis-regulatory gene networks. Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/80404>

<https://doi.org/10.32657/10220/46576>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**EFFECT OF CRISPR-BASED EPIGENETIC
INTERFERENCE ON
CIS-REGULATORY GENE NETWORKS**

By

LOH XINPENG

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University in partial
fulfilment of the requirement for the degree of Master of Science

July, 2018

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude towards my supervisor, Asst Professor Amartya Sanyal, for his guidance and the opportunities for development throughout the course of working in his lab. This work was supported by Nanyang Technological University's Nanyang Assistant Professorship grant and Singapore Ministry of Education Academic Research Fund Tier 1 grant to Asst Professor Amartya Sanyal.

I would also like to extend my heartfelt appreciation to my mentor Dr. Aravinda Chavalmane Krishnappa Naik, who has provided unwavering support throughout my project.

Finally, I would like to thank my lab members Miss Siti Rawaidah Binte Mohammad Muzaki, Mr Costerwell Khyriem, Miss Yao Chen and Dr. Vijayachandra Kinnimulki for their technical expertise, assistance and kind encouragements in the lab.

CONTENTS

TABLE OF FIGURES	4
ABBREVIATIONS	7
ABSTRACT	8
INTRODUCTION.....	9
HYPOTHESIS AND AIMS	21
MATERIALS AND METHODS	28
EXPERIMENTAL RESULTS	39
DISCUSSION	65
CONCLUSION	69
REFERENCES.....	70

TABLE OF FIGURES

Figure 1 Schematic illustration of CRISPR interference and activation systems	13
Figure 2 HBB gRNA designed to target HS2 enhancer in the locus control region. ...	23
Figure 3 TAL1 gRNA designed to target the erythroid enhancer in the stem cell leukemia locus.....	23
Figure 4 Copy number analysis of ABCC1 and INO80C loci in Doxorubicin sensitive and tolerant SCLC cells	26
Figure 5 ABCC1-Ups gRNA designed to target the enhancer region upstream of ABCC1 drug transporter encoding gene	26
Figure 6 ABCC1-Int gRNA designed to target intragenic regions of ABCC1 drug transporter encoding gene	27
Figure 7 INO80C-Enh gRNA designed to target the enhancer region upstream of INO80C chromatin remodeler protein encoding gene	27
Figure 8 Schematic illustration of Krüppel associated box (KRAB) domain fused with dead Cas9 protein.....	30
Figure 9 Schematic illustration of Lysine-Specific Demethylase 1 (LSD1) fused with dead Cas9 protein.....	30
Figure 10 Schematic illustration of Histone Deacetylase 1 (HDAC1) fused with dead Cas9 protein	31
Figure 11 Schematic illustration of Heterochromatin Protein 1 alpha (HP1 α) fused with dead Cas9 protein.....	31
Figure 12 Schematic illustration of CRISPR-based epigenome editing plasmid assembly.....	40
Figure 13 Construction of gRNA and dCas9-epigenetic modifier recombinant plasmids.....	41
Figure 14 Schematic illustration of the screening process for desired clones.....	42
Figure 15 Detecting recombinant CRISPRi plasmids.	43
Figure 16 Determining orientation of insert in recombinant CRISPRi plasmids	44
Figure 17 Restriction enzyme validation of recombinant CRISPRi plasmids.	45
Figure 18 Schematic summary of assembled recombinant CRISPRi constructs targeted at the LCR and SCL loci.....	46

Figure 19 Schematic summary of assembled recombinant CRISPRi constructs against the ABCC1 super-enhancer region	47
Figure 20 Schematic summary of assembled recombinant CRISPRi constructs against the INO80C super-enhancer region	48
Figure 21 Expression of dCas9- epigenetic modifier fusion proteins in K562 cells....	51
Figure 22 Quantification of globin gene expression in K562 cells transfected with HBB gRNA	54
Figure 23 Quantification of SCL locus gene expression in K562 cells transfected with TAL1 gRNA.....	55
Figure 24 Quantification of ABCC1 locus gene expression in low Doxorubicin tolerant cells transfected with ABCC1-Ups gRNA.....	58
Figure 25 Quantification of ABCC1 locus gene expression in low Doxorubicin tolerant cells transfected with ABCC1-Int gRNA	59
Figure 26 Quantification of ABCC1 locus gene expression in high Doxorubicin tolerant cells transfected with ABCC1-Ups gRNA.....	62
Figure 27 Quantification of ABCC1 locus gene expression in high Doxorubicin tolerant cells transfected with ABCC1-Int gRNA	63
Figure 28 Quantification of INO80C locus gene expression in high Doxorubicin tolerant cells transfected with INO80C-Enh gRNA.....	64

LIST OF TABLES

Table 1 Sequences of gRNA and the respective upstream PAM site	28
Table 2 Buffer composition for restriction enzyme digestion.	32
Table 3 Primer set sequences for various screening reactions.	34
Table 4 Primer sets sequences for quantitative analysis of target genes.	37

ABBREVIATIONS

CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats

CRISPRa, Clustered Regularly Interspaced Short Palindromic Repeats Activation

CRISPRi, Clustered Regularly Interspaced Short Palindromic Repeats Interference

dCas9, dead-CRISPR associated protein 9

Dox, Doxorubicin

gRNA, guide RNA

H69-150DT, Low Doxorubicin tolerant SCLC cells

H69AR, High Doxorubicin tolerant SCLC cells

HDAC1, Histone Deacetylase 1

HP1 α , Heterochromatin protein 1

K562, Chronic Myelogenous Leukemia cells

KRAB, Krüppel associated box

LSD1, Lysine-specific histone demethylase 1

NCI-H69, Doxorubicin sensitive SCLC cells

PAM, Protospacer Adjacent Motif

SCLC, Small Cell Lung Carcinoma

HBB, Beta globin gene locus

HBG, Gamma globin subunit gene

HBE, Epsilon globin subunit gene

LCR, Locus Control Region

SCL, Stem Cell Leukemia

TAL1, T-cell Acute Lymphocytic Leukemia protein 1

ABCC1, ATP Binding Cassette Subfamily C Member 1

INO80C, INO80 Chromatin remodeling complex subunit C

ABSTRACT

The revolutionary CRISPR-Cas9 genome engineering technology has been repurposed into a docking tool by fusing various effector proteins to the catalytically inactive Cas9. The modified system can regulate transcription by editing the epigenetic code, thereby modifying the local chromatin structure. While several reports have harnessed the potential of the CRISPR interference (CRISPRi) system to silence genes, only limited types of effector domains have been tested to date. Here, we expanded the repertoire of effector domains in the CRISPRi toolbox by incorporating novel epigenetic modulators - Histone Deacetylase 1 and Heterochromatin Protein 1 alpha - to study gene regulatory networks (GRNs). Efficiencies of these novel epigenetic effector domains were successfully demonstrated by targeting enhancers of globin genes in leukemia cells, as well as drug resistance genes in a chemoresistant model. Integration of new epigenetic modulators with CRISPRi would present new possibilities for understanding GRNs in various cell types and context.

Graphical abstract

INTRODUCTION

Introduction to epigenetics

Epigenetics serves as the underlying molecular basis that resolved the long-standing inadequacy of genotype to explain phenotypic variations (Enríquez, 2016). This dynamic branch of genetics is associated with studying basic molecular processes responsible for stable and heritable changes in gene expression or cellular phenotype not caused by alteration of DNA sequences (Goldberg et al., 2007).

Epigenetic studies in the last two decades have unraveled novel understanding of the main arms that constitute the field: DNA modifications, histone post-translational modifications and non-coding RNAs (Enríquez, 2016). These mechanisms act cooperatively to choreograph the highly complex organization of chromatin structure.

At the first level of organization, the human chromatin fiber is packaged in a series of nucleosome core particles, with each particle composed of ~146 base pair of DNA wrapped around eight core histone proteins – two copies of the four different histone types (H3, H4, H2A and H2B) (Young et al., 1986). While the histone globular core is surrounded by DNA, the N-terminal tails of the histone proteins extend out of the nucleosomes. This makes them accessible to molecular machinery that can induce post-translational modifications including acetylation, methylation, phosphorylation, sumoylation, etc. Several components of the protein machinery that recognizes, adds or removes these post-translational modifications have come under scrutiny to gain mechanistic understanding of chromatin structure regulation (Gillette & Hill, 2015). The expression and composition of post-translational modification complexes is intricately controlled, as it engages in cross-talk with DNA and dynamically affects the cell's transcriptional program (Torres & Fujimori, 2015). Histone acetyltransferases are one of the examples of enzymatic 'writers', which add acetyl groups to histones, while histone deacetylases are 'erasers' that antagonize this process by removing the acetylation marks (Gillette & Hill, 2015). Addition or removal of acetyl groups directly influences the strength of electrostatic attraction between histone proteins and DNA, consequently affecting chromatin compaction (Workman & Kingston, 1998). Alternatively, the effect can be mediated via histone 'readers', such as chromodomain, which have high affinity for recognizing specific post-translational modifications (Yun et al., 2011). These modules are usually a part of effector protein complexes containing chromatin modifying or adaptor proteins to aid in localizing these enzymes (Gillette & Hill, 2015).

These modifications collectively exert effects on endogenous genomic DNA replication, repair and transcriptional control (Anand & Marmorstein, 2007). Considering their effect on such key processes, specific post-translational modifications have been associated with distinct cellular phenotypes. For instance, the methylation of histone H3 lysine 9 (H3K9me) is the characteristic feature of silenced promoters, while active promoters are associated with histone H3 lysine 4 methylation (H3K4me) and histone H3 lysine 27 acetylation (H3K9ac) marks. However, the transcriptional effects of many epigenetic modifications and their role in disease phenotypes remain poorly understood (Enríquez, 2016). Technological limitations have impeded researchers' ability to maneuver local and genome-wide epigenetic regulation. This highlights the need for genome manipulation tools that enable alteration of locus- specific epigenetic mark to functionally dissect their significance in normal development and disease.

State-of-the-art genome editing tools

A diverse array of tools has been employed for genetic modification purposes over the years. These tools are broadly classified depending on their intended target identification method; *via* protein domains attributed or nucleotide-directed specificity.

Commonly used protein-based genome editing tools include recombinases, integrases, Zinc Finger Nuclease Protein (ZFP) modelled from eukaryotic transcription factors, and the Transcription Activator-Like Effector Nucleases (TALENs) adapted from the plant pathogens under the genus *Xanthomonas* (Hsu et al., 2014). However, recombinases and integrases have limited application potential because the catalytic mechanism necessitates pre-existing recognition sites in the genome (Esvelt & Wang, 2014). ZFN and TALEN protein domains can be customized by manufacturing them in required combination and then affixed to one half of the FokI nuclease domain for site directed nucleolytic activity (Bitinaite et al., 1998). However, this protein engineering process is costly and entails a labor-intensive manufacturing process (Pabo et al., 2001). In the assembly of ZFP, each module is specific for a single DNA triplet code. Hence, a series of modules would be required to target the ZFP to a sequence composed of an array of DNA triplets. In a similarly extensive screening process, TALEN proteins are made up of multiple repeats of 35 amino acid long domain. The individual domain can recognize only a

single base pair of the targeted DNA sequence, making TALE arrays bulky and labor intensive to assemble (Hsu et al., 2014).

Popular nucleotide-directed approaches include small interference RNA (RNAi) and the recent Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) – CRISPR associated protein 9 (Cas9) tool derived from the bacterial adaptive immune system. RNAi is highly susceptible to severe off-target effects (Xu et al., 2014), and is more adept at hindering cytosolic mRNA targets because interference is facilitated by cytoplasmic Argonaute proteins (Gilbert et al., 2014a).

CRISPR-Cas9 is an RNA-guided editing tool that has greatly streamlined and revolutionized the process of customizable site-specific gene editing in both prokaryotes (Jiang et al., 2013) and eukaryotes (Mali et al., 2013b; Cong et al., 2013). CRISPR has stemmed from adaptive immunity studies on bacteria and Archaea which were found to utilize RNA guided DNA endonucleases to cleave extraneous DNA (Marraffini & Sontheimer, 2010; Wiedenheft & Sternberg, 2012). Three main types of CRISPR systems were discovered, categorized based on sequence homology and endonuclease structure. As opposed to type I and III, the type II CRISPR system is most commonly utilized for genome engineering applications due to the structural simplicity of its Cas9 ribonuclease protein (Xu et al., 2014). It is cost effective and highly versatile because it can be tailored to needs by simply replacing the guide RNA sequence (gRNA) which is designed complementary to the desired genomic site. Compared to previous targeting techniques, it has also demonstrated higher efficiency, site specificity (Lo & Qi, 2017) and potential for multiplexed gene modification by incorporating an array of guide RNAs against multiple gene targets (Hsu et al., 2014).

Mechanism of CRISPR

The functional type II CRISPR-Cas9 complex comprises of the Cas9 endonuclease, a CRISPR RNA (crRNA) and a trans-activating CRISPR RNA (tracrRNA). In response to invading genetic material, short fragments of the foreign DNA are transcribed and processed into crRNAs. The crRNAs and tracrRNAs organize into a hetero-duplex that dictates the sequence to be cleaved by Cas9. The Cas9 endonuclease, in turn, requires an existing conserved protospacer adjacent motif (PAM) sequence upstream of the intended target's crRNA binding region to execute a double strand nick (Gaj et al., 2013).

In the synthetic CRISPR-Cas9 system adapted for genome editing purposes, the crRNA and tracrRNA have been reengineered into a single guide RNA (gRNA) via a 'GAAA' linker. This produces a simplified, one-component system crucial in improving guiding efficiency (Mali et al., 2013b; Cong et al., 2013). The 20-nucleotide gRNA can be manipulated against any genomic target that is adjacent to a PAM sequence, which can vary with the type of Cas9 being engaged (Mojica et al., 2009; Xu et al., 2014). The Cas9 from *Streptococcus pyogenes* is most commonly employed, and biochemical analysis of this particular Cas9 suggests the importance of PAM sequence in CRISPR-Cas9 complex progression from target binding to strand cleavage. The gRNA-DNA duplex annealing begins at the PAM site, then advances to the upstream target sequence for actual guide sequence complementarity. The initial PAM binding and gRNA that corresponds to target site activates the Cas9 domains HNH and RuvC responsible for nucleolytic function (Sternberg et al., 2014; Hsu et al., 2014). The length and complexity of the PAM sequence also influence the overall available target sites. The 5' 'NGG' requirement imposed by *S. pyogenes* Cas9 allows it to act on the human genome, on average, at every 8 base pairs. The targeting space is further expanded as the *S. pyogenes* Cas9 also accommodates target sites with a 5' 'NAG' PAM sequence, though at a lower efficiency (Hsu et al., 2013).

Hence, Cas9 is directed to its intended site via complementary base pairing between the gRNA and target sequence in order to generate a sequence-specific DNA double strand break (Garneau et al., 2010). The double strand break is repaired by the error-prone non-homologous end joining repair pathway, often resulting in insertions and deletions at the target site. Alternatively, if an appropriate repair template is provided, it can also be incorporated to the site of edit through homology directed repair (Lo & Qi, 2017). Besides its ability to cause loss-of-function mutations, the CRISPR based technology has also been repurposed into a genome docking tool, further extending its range of utility (**Fig. 1**) (Xu et al., 2014; Qi et al., 2013). By synthetically fusing dCas9 with other functional enzymatic domains or effector proteins, transcription can be regulated at a locus of interest (Qi et al., 2013). The versatile substitution of the modular component affixed to dCas9 led to the development of diverse CRISPR-based activation (CRISPRa) and interference (CRISPRi) systems to study complex gene regulatory networks (Beerli et al., 2000).

Figure 1 Schematic illustration of CRISPR interference and activation systems. In the synthetically modified *S. pyogenes* Cas9 system, the target sequence must be immediately adjacent to a 5'-NGG PAM site. The gRNA forms complementary pairing with the opposite strand of the target sequence, to anchor the complex to the desired site. Cas9 cleavage activity is ablated by mutating the RuVC and HNH domains, resulting in a dead Cas9 (dCas9) that is unable to induce a double strand break yet retains its ability to bind to target site. The dCas9 can also be tethered to a transcriptional activator or repressor for further transcriptional control. Some promoters can impose restrictions on target site selection and gRNA design. In the case of an RNA-polymerase III dependent U6 promoter, a single 'G' nucleotide is required at the 5' end of the target sequence to initiate transcription.

Epigenetic additions to the genome engineering toolbox

The concept of epigenetic modification surfaced initially with the advent of ZFN and TALENS. Manipulation of gene expression through histone or DNA modification using TALE fusion with chromatin modifiers (Mendenhall et al., 2013; Beerli et al., 2000) or similar ZFN constructs (Snowden et al., 2002; Krichevsky et al., 2007) paved the way for applying similar strategy in the CRISPR-Cas9 context.

CRISPR-based transcriptional activation

The dCas9 protein has been coupled with several transactivation domains, such as the NF- κ B p65 subunit, Herpes simplex Viral Protein (VP16), or its multiple tandem repeats variants VP64 and VP160 (Gao et al., 2014). While these tools have demonstrated significant gene activation, systemic comparisons show a large variation in activation potency, in the range of 10 to 100-fold difference, across various genes and cell types (Chavez et al., 2016). This prompted the use of synergistic epigenetic modifiers to supplement existing CRISPR based strategies and improve gene activation efficiency.

In 2015, a research group tethered dCas9 to the catalytic core of the histone acetyltransferase (HAT) enzyme p300. The p300 acetylates histone H3 lysine 27 (H3K27) at enhancers and promoters, thereby inducing transcriptional activation. Such fusion of epigenetic modulator exhibited stronger activation potency when directed to proximal and distal regulatory elements by a single gRNA, rivalled against other similar fusion constructs engineered using activation domains of transcription factors (Hilton et al., 2015).

Another study in 2016 reported the concurrent use of two epigenetic activators, histone methyltransferase PRDM9 and DOT1L catalytic core with dCas9 to enrich target sites for histone 3 lysine 4 trimethylation (H3K4me3) and histone 3 lysine 79 methylation (H3K79me) respectively. This study demonstrated the additive effect of both dCas9-PRDM9 and dCas9-DOT1L to cooperatively hypermethylate the target locus and sustain strong reactivation of silenced genes (Cano-Rodriguez et al., 2016).

CRISPR-based transcriptional repression

Compared to CRISPRa, fewer CRISPRi systems have been established and characterized. In fact, the sole use of dCas9 protein and a corresponding gRNA targeted to a gene promoter is sufficient for transcriptional repression as it prevents transcriptional initiation or elongation by limiting RNA polymerase access (Xu et al., 2014). This mode of repression exhibited up to 99.9% gene silencing efficiency in prokaryotes (Qi et al., 2013), but yielded modest levels (60-80%) of suppression when evaluated in the context of mammalian cells (Gilbert et al., 2014a). Given the elevated complexity of endogenous gene regulation in eukaryotes, it is plausible that the dCas9-sgRNA complex alone is inadequate to impede the action of transcription machinery in mammalian cells. To increase the repression efficiency, dCas9 protein is fused to various types of repressive effector domains, including epigenetic modulators.

Gilbert and group first demonstrated the use of epigenetic-based CRISPRi technology in 2014 by fusing Krüppel-associated box (KRAB) domain of KOX1 with dCas9. They successfully modified the expression of genes encoding for cell surface transmembrane proteins in HeLa cells (Gilbert et al., 2014b). Gersbach's group also used dCas9-KRAB to epigenetically modulate erythroid gene expression (Thakore et al., 2015). The KRAB domain recruits a plethora of protein complexes, such as the KRAB-associated protein 1 (KAP1) co-repressor to the specified locus (Groner et al., 2010). The study examined the chromatin remodeling capacity of dCas9-KRAB when directed to the distal HS2 enhancer of locus control region (LCR) responsible for transcriptional activation of globin subunit genes in the erythroid cells. The complex resulted in decreased chromatin accessibility and facilitated deposition of repressive methylation marks (H3K9me3) within the designated stretch of HS2 enhancer. This ensued in the targeted silencing of globin genes, with little off-target effects detected (Thakore et al., 2015).

Kearns and group pitted the enhancer repressive properties of dCas9-KRAB against a novel histone demethylase fusion construct – dCas9 reconstituted with Lysine-Specific Demethylase 1 (LSD1) (Kearns et al., 2015). The LSD1 protein was first characterized in 2004 and it catalyzes the removal of methyl group from methylated H3K4 (active histone mark) (Shi et al., 2004). The study compared the ability of KRAB and LSD1 fusion complexes to repress pluripotency maintenance genes in embryonic stem cells, by means of downregulating Tbx3 transcription factor

expression. The findings revealed major differences in repressive mechanism; the dCas9-LSD1 complex did not cause pronounced chromatin structure perturbations, but instead, repressed the target enhancer in a manner independent of the effector domain. In contrast, instead of repressing enhancer activity, the dCas9-KRAB was reported to induce promoter silencing or heterochromatin spreading. This stands in contradiction to the conclusions drawn from the dCas9-KRAB HS2 enhancer study (Thakore et al., 2015), suggesting that epigenome editing via dCas9-KRAB-mediated histone modification may be context and cell-type dependent.

The CRISPR-based epigenome modification allows unparalleled precision in the spatial control of dynamic chromatin reorganization (Beerli et al., 2000), and is certainly poised to become a potent tool in bridging the gaps between chromatin state, gene regulation and cellular phenotypes. By further expanding and recombining dCas9 nuclease with novel epigenetic modifiers, complex distal regulatory elements can be functionally characterized and utilized to restore the epigenome. Hence, our study proposes to tether novel epigenetic modulator alternatives to existing combinations of dCas9-KRAB and dCas9-LSD1 and to interrogate their repressive properties on enhancer regions. We selected epigenetic enzyme Histone Deacetylase 1 (HDAC1) and effector protein Heterochromatin Protein 1 alpha (HP1 α) as novel candidates to integrate into the dCas9-gRNA complex, due to their well-established roles in chromatin repression (Johnstone, 2002; Eissenberg & Elgin, 2014). Additionally, existing dCas9-epigenetic effectors (KRAB and LSD1) also recruit HP1 α and HDAC1 in their mechanism of action (Sripathy et al., 2006; Anand & Marmorstein, 2007), suggesting a pivotal role of the latter in sustained, robust gene repression. The next section will expand on the structural annotations and functions of each epigenetic effector in our study.

Krüppel associated box (KRAB) domain

KRAB is a small, 75 amino acid residue protein originally discovered as a conserved amino terminus module amongst a prominent group of human zinc finger transcription factors (Bellefroid et al., 1991). KRAB domain is a part of a family of powerful repressors that decommission activators, transcriptional apparatus or induce local chromatin remodeling (Hanna-Rose & Hansen, 1996).

The KRAB domain can be functionally divided into two adjacent subdomains, KRAB-A and KRAB-B, encoded by separate exons and have complementary functions. The KRAB-A subdomain has been shown to be necessary and sufficient for transcriptional repression (Witzgall et al., 1994), while subdomain B plays an auxiliary role in the assembly of multi-molecular silencing complex through its interactions with transcriptional intermediary factor 1 β (Abrink et al., 2001; Barde et al., 2009). The KRAB domain itself does not contain a DNA-binding motif (Witzgall et al., 1994), hence, it is a portable module that can be utilized in conjunction with other target recognition motifs, such as the dCas9/gRNA complex.

The mechanism of KRAB-mediated transcriptional repression has been well elucidated, in large part due to the identification of its universal co-repressor KAP-1 (Barde et al., 2009). KRAB-A subdomain binding to the Ring B-BOXED Coiled Coil of KAP1 is required to initiate gene silencing. Upon KRAB-KAP1 complex binding at a specific locus, KAP1 operates as a scaffold to recruit a cascade of downstream enzymes associated with a repressive epigenetic state *via* protein-protein interactions (Schultz et al., 2002). This includes heterochromatin protein 1 (HP1), histone deacetylase complex NCOR1, histone methyltransferase SETDB1, and Nucleosome Remodeling Deacetylase (NuRD) co-repressor to transcriptionally silence the target site through chromatin remodeling (Sripathy et al., 2006). While KRAB has not been functionally implicated in many molecular processes, its co-repressor KAP1 is involved in important processes such as DNA double strand break repair and regulation of self-renewal (Enríquez, 2016).

Lysine-Specific Demethylase 1 (LSD1)

LSD1 belongs to the amine oxidase family of enzymes, characterized by their highly conserved N-terminal alpha helical SWIRM and C-terminal Flavin Adenine Dinucleotide (FAD)-dependent amine oxidase domain (Chen et al., 2006). The SWIRM domain, named after the chromosomal proteins Swi3p, Rsc8p, Moira, is present in many chromosomal proteins and hypothesized to anchor protein complexes to nucleosomal substrates (Anand & Marmorstein, 2007). The FAD-dependent amine oxidase domain is further sub-divided into substrate binding and FAD binding subdomains, both crucial to forming the site of catalytic action for the enzyme. LSD1 is responsible for the demethylation of H3K4me1 and H3K4me2 active histone marks (Shi et al., 2004; Anand & Marmorstein, 2007).

The central Tower domain is highly varied amongst amine oxidases, and in the case of LSD1, it forms a helical binding platform for LSD1 co-repressor element silencing factor (coREST) complex. Cooperatively, the coREST-LSD1 complex performs the histone demethylation. The N-terminal of the coREST protein was also found to harbor a histone deacetylase interacting domain. Hence, LSD1 functions as part of a multi-protein transcription repressor complex involving both histone demethylase and deacetylase components (Anand & Marmorstein, 2007).

However, LSD1 behaves as a co-repressor or co-activator when associating with different complexes and substrates. As part of coREST and NuRD repressive complexes, it represses transcription by demethylating active marks H3K4me1/2 (Shi et al., 2003). But when interacting with nuclear hormone receptors, LSD1's substrate specificity switches from H3K4me1/2 active histone marks to H3K9me1/2 inactive histone marks. Under these circumstances, it operates as a co-activator by demethylating the heterochromatin epigenetic marks (Garcia-Bassets et al., 2007).

LSD1 is also involved in the demethylation of non-histone proteins, including p53 transcription factor in DNA damage response (Huang et al., 2007), E2F protein in DNA-damage induced apoptosis, MYPT phosphatase in mammalian cell cycle (Cho et al., 2011) and Dnmt1 enzyme needed for maintaining DNA methylation during DNA replication (Wang et al., 2009). This indicates the involvement of LSD1 in a plethora of cellular activities.

Histone deacetylase 1 (HDAC1)

Genome-wide chromatin acetylation status is balanced by the activity levels of histone acetyltransferase (HAT) and histone deacetylase (HDAC). HDAC1 is a class I, zinc-dependent deacetylase naturally occurring as part of multi-protein repressor complexes including coREST, NuRD and Sin3 (Ruijter et al., 2003). HDAC1 comprises almost entirely of the highly conserved zinc-dependent deacetylase domain, with minimal N or C terminal domain. It is a 482 amino acid protein in human, including multiple serine residues that are subjected to phosphorylation by kinases (Delcuve et al., 2012).

The HDAC1 proteins initially form homodimers, before associating with their respective protein partners in different cell types (Taplick et al., 2001). Each repressor complex has a variable composition of effector proteins and transcription factors capable of further recruiting a specific subset of enzymatic modules. In addition to HDAC1, the coREST complex also engages histone demethylase LSD1, while the Sin3 complex mobilizes several nucleosome remodeling enzymes, as well as DNA and histone methyltransferases. HDAC1 can also act in conjunction with the Polycomb Repressive Complex 2 (PRC2) to catalyze H3K27 trimethylation. Thus, HDAC1 could have diverse roles, substrate specificity and genomic localization depending on its interacting partners (Delcuve et al., 2012).

HDAC1 is also known as 'canonical' HDAC due to their conventional strong activity on histone proteins. They deacetylate lysine residues on the N-terminal tails of all 4 types of core histone proteins – H2A, H2B, H3 and H4. While acetylation removes the positive charges, deacetylation restores the positive charge on the histone amine group. This results in enhanced ionic interactions between the histone and negatively charged DNA. Consequently, the chromatin structure undergoes compaction, which in turn silences gene transcription by limiting the access of transcriptional machinery (Johnstone, 2002). However, recent genome-wide studies have implicated them in the deacetylation of non-histone proteins, albeit at a lower substrate specificity (Haberland et al., 2009). This includes transcription factors and co-regulators associated with cell cycle progression control, cell differentiation, apoptosis and metabolism (Mihaylova & Shaw, 2013). Hence, HDACs have indirect functions in regulating cellular homeostasis (Minucci & Pelicci, 2006).

Heterochromatin protein 1 alpha (HP1 α)

HP1 α was first identified in *Drosophila melanogaster* as a key factor in mediating euchromatic gene silencing when these genes were randomly displaced near a heterochromatin region, in a process later known as position effect variegation (Eissenberg & Reuter, 2009). Several homologues of the protein were later discovered, all of which possess the highly conserved N-terminal chromodomain and C-terminal chromoshadow domain (Maison & Almouzni, 2004). The HP1 proteins also contain a highly variable hinge region reported to have protein and nucleic acid binding roles, potentially conferring each isoform the target binding selectivity (Maison & Almouzni, 2004).

HP1 α physically links a specified histone mark to a non-histone effector protein through the amino and carboxyl terminal domains. The chromodomain of HP1 α recognizes and binds to methylated lysine 9 residues on histone 3 (H3K9me₂/me₃) (Eissenberg & Elgin, 2014) while the chromoshadow interacts with enzymatic proteins, such as histone methyltransferases SU(VAR)3-9 which catalyzes methylation of histone H3K9. This self-sustaining succession of recognizing a heterochromatic H3K9me mark and recruiting methyltransferases to further act on neighboring lysine residues stabilizes and propagates the spread of heterochromatin near the site of initiation (Eissenberg & Elgin, 2014). HP1 α proteins also self-associate to form homodimers through the chromoshadow domain while the chromodomain remain bound to the H3K9me marks, leading to further spreading of heterochromatin (Eissenberg & Elgin, 2014) When a methyl group occupies a histone lysine residue, it precludes acetylation at that site. Additionally, binding of HP1 α interferes with demethylation of the histone proteins. HP1 α also employs Cohesin through its interacting partners to promote higher order chromatin condensation (Hahn et al., 2013). Taken together, these mechanisms operate cooperatively to promote a stable heterochromatin state (Maison & Almouzni, 2004).

HYPOTHESIS AND AIMS

The simple, highly efficient and customizable nature of the CRISPR system poises it as a convenient tool for studying endogenous *cis*-regulatory gene networks.

Systematic mapping of the human epigenome has identified several cell type-specific enhancers involved in development, differentiation and disease (ENCODE Project Consortium, 2012). These CRISPR-based epigenome editing tools can be targeted to an identified enhancer region, induce local chromatin structure modification and potentially influence the transcription level of target gene.

As discussed in the preceding introductory segment, the repressive function of existing dCas9-epigenetic effectors, KRAB and LSD1, acts through the recruitment of HP1 α and HDAC1. This implies a pivotal role of HP1 α and HDAC1 in chromatin compaction. Therefore, **we hypothesize that novel CRISPRi strategies using HDAC1 and HP1 α will display enhanced transcriptional repression potency than the existing KRAB and LSD1-based CRISPRi systems.**

This study aims to expand the repertoire of CRISPRi-based epigenetic modulator tools that can downregulate enhancer-driven gene expression in different models.

- **Specific aim I:** Generate recombinant CRISPRi constructs containing HDAC1, HP1 α , LSD1 and KRAB effectors for targeted disruption of gene enhancer
- **Specific aim II:** Proof-of-concept study to analyze efficacies of CRISPRi constructs using well-established erythroid-specific *cis*-regulatory gene activation models
- **Specific aim III:** Examine efficacies of CRISPRi constructs in repressing transcription of novel variant enhancer-driven genes differentially expressed in a chemoresistant model

Specific aim I: Generate recombinant CRISPRi constructs containing HDAC1, HP1 α , LSD1 and KRAB effectors for targeted disruption of gene enhancer

Previous attempts to apply CRISPRi technique in the lab by co-transfecting two plasmids, one encoding for the enhancer-specific gRNA and another for dCas9-epigenetic modifier, showed inconsistent results in suppressing enhancer-driven gene expression. One possible reason attributed to this observation is the failure in delivery of both plasmids, which are indispensable for the formation of CRISPRi complex, into the recipient cells during co-transfection. Hence, we sought to overcome the limitations posed by co-transfection by integrating both gRNA expression cassette and dCas9 expression elements into a single construct to facilitate greater ease of transfection.

Single vector systems targeting the globin gene locus (aim II) and drug resistant gene locus (aim III) were established in-house and assessed for their repressive potency.

Specific aim II: Proof-of-concept study to analyze efficacies of CRISPRi constructs using well-established erythroid-specific *cis*-regulatory gene activation models

We first designed gRNAs to direct the dCas9-epigenetic modulator constructs to two well-characterized enhancer clusters in K562 chronic myelogenous leukemia (CML) cells: the HS2 enhancer of LCR in the beta-globin (*HBB*) locus (**Fig. 2**) and the erythroid enhancer in stem cell leukemia (*SCL*) locus (**Fig. 3**).

The LCR contains five distinct erythroid-specific enhancer regions, hypersensitive sites HS1-5, which regulate different globin gene cluster expression throughout different stages of erythroid cell development (Tuan et al., 1992). Specifically, the hypersensitive site 2 (HS2) has exhibited many crucial enhancer properties, such as harboring transcription factor binding sites (McDowell & Dean, 1999) and chromatin looping to globin gene promoters (Vakoc et al., 2005). On the other hand, the *SCL* locus comprises 2 distinct enhancer regions – the hematopoietic stem cell enhancer and the erythroid enhancer. Both regions are enriched with multiple H3K4me2 transcriptionally active marks and H3K27ac active enhancer marks, but is functionally unique to cell lineages at distinct stages of hematopoiesis (Dhami et al., 2010). In K562 cells, the erythroid enhancer at the +50 kb region of the *SCL* locus (Dhami et al., 2010) directly regulates expression of transcription factor *TAL1* and

neighboring gene *STIL* (Zhou et al., 2013). TAL1 is an important regulator of hematopoiesis and erythroid differentiation (Hall et al., 2003), likely *via* enhancing globin gene expression. It interacts with GATA-1 erythroid factor, as well as chromatin loop-mediating proteins Ldb1 and LMO2 to promote looping contacts between globin genes and their respective HS enhancers (Yun et al., 2014). The TAL1 chromatin looping hub also brings the *cis*-regulatory elements in the *SCL* locus and a *TAL1* flanking gene known as *STIL* into close proximity, thereby enhancing *STIL* expression (Zhou et al., 2013).

Hence, we propose to employ four CRISPRi constructs generated in-house for targeted disruption of LCR and *SCL* enhancer. By disrupting well studied *cis*-regulatory circuitry in beta-globin and *SCL* loci, we will quantify the downregulation of globin genes, *TAL1* and *STIL* transcripts as an assessment of the repressive potential of dCas9-HDAC1/HP1 α fusion constructs compared to the existing dCas9-LSD1/KRAB tools.

Figure 2 HBB gRNA designed to target HS2 enhancer in the locus control region (red arrow). LCR is characterized by H3K27ac ChIP-Seq tracks shown from UCSC genome browser.

Figure 3 TAL1 gRNA designed to target the erythroid enhancer in the stem cell leukemia locus (red arrow). SCL is characterized by H3K27ac ChIP-Seq tracks shown from UCSC genome browser.

Specific aim III: Examine efficacies of CRISPRi constructs in repressing transcription of novel variant enhancer-driven genes differentially expressed in a chemoresistant model

The validated CRISPRi constructs can be used in suppressing genes whose misexpression is associated with aberrant *cis*-regulatory mechanisms. This phenomenon is observed in the instance of cancer drug resistant cells, which acquire substantial alterations to its transcriptional profile. Epigenetic manipulation is pertinent and crucial to re-sensitize these cells to chemotherapy drugs.

Cancer cells can elude drug treatment through genetically intrinsic mechanisms or environmentally acquired alterations, whereby initially drug sensitive cells undergo transcriptional reprogramming in response to a therapeutic agent (Kerbel et al., 1994). While the former has an underlying genetic basis, studies have associated epigenetic-mediated transcriptional change in acquired drug resistance (Gottesman, 2002a). Drug resistance is a multi-faceted phenomenon that can be caused by accelerated drug efflux, drug metabolism, alterations or amplifications of drug targets, evading apoptosis or inhibiting normal cell homeostasis processes (Gottesman, 2002a).

Among these contributing factors, increase in drug efflux was cited as a leading cause of multi-drug resistance, in which cancer cells become simultaneously desensitized towards multiple cytotoxic drugs. Enhanced drug efflux is often caused by overexpression of ATP-binding cassette transporters, such as Multidrug Resistance-associated Protein 1 (MRP1/ABCC1) (Gottesman et al., 2002b; Cole et al., 1992). Our lab previously identified enhancer-driven regulation of genes conferring drug resistance through dynamic alterations of chromatin structure. The study was performed in the context of Small Cell Lung Cancer (SCLC), which has been affiliated with tobacco consumption as its main predisposing factor (myVMC, 2015). Lung cancer is the world's leading cause of cancer death and a runner-up amongst the list of common cancers diagnosed in both genders (Siegel et al., 2017). SCLC accounts for approximately 15-17% of all lung cancer cases (Kahnert et al., 2016). It is a highly aggressive, metastatic cancer known for its high growth rates (Hamilton & Rath, 2015) and quick progression into drug resistant phenotype (Chen et al., 2012). As clinically significant symptoms are slow to surface, the cancer would have metastasized in approximately 67-75% of patients at the point of diagnosis (Conrad, 2018). Unlike other forms of lung cancer, such as non-small cell lung

cancer (NSCLC), SCLC treatment improvements have been stagnant in the past few decades (Wang et al., 2017). Consequently, the 5-year survival rate of SCLC remains low at 6.4% and the median survival unchanged at 7 months over the last 2 decades (Wang et al., 2017). The poor prognosis, limited treatment options and high mortality rates suggest that SCLC will persist as health threat. Thus, it would be clinically beneficial to identify molecular factors aptly poised for therapeutic intervention and enhance understanding of the mechanism driving disease progression.

The lab carried out gene expression and enhancer profiling of two SCLC cell lines representing distinct states of Doxorubicin (Dox) drug sensitivity and identified enhancer-mediated overexpression of two crucial transcripts – *ABCC1* (drug transporter gene) and *INO80C* (chromatin remodeler gene). *ABCC1* was focally amplified in the low Doxorubicin tolerant cell line H69-150DT (Dox^{low}; 150nM Doxorubicin resistant) (**Fig. 4B**), while both genes were focally amplified in the high Doxorubicin tolerant cell line H69AR (Dox^{high}; EC50: 2.23μM), in contrast to its Doxorubicin sensitive counterpart NCI-H69 (EC50: 30nM) (**Fig. 4A & 4C**). ChIP-seq results also showed increased enrichment of H3K27ac active enhancer marks in both Dox^{low} and Dox^{high} cell line upstream of the transcription start site and in intragenic regions (**Fig. 5-7**) (Unpublished data).

Presence of multiple enhancer enrichment clusters is collectively characterized as ‘super-enhancer’ region (Pott & Lieb, 2014). As not all the genes present in the super-enhancer region are upregulated, this suggests the crucial presence of variant enhancer loci at work to induce differential gene expression during the cellular state transition to a drug resistant phenotype. Hence, we will utilize the validated CRISPRi constructs containing epigenetic modifying enzymes to occlude the variant enhancer region from functionally acting on genes that facilitate drug resistance. We designed gRNAs against the enhancer regions using information from H3K27ac ChIP-Seq peaks and presence of high-occupancy target genomic regions with unusually high number of transcription factor binding sites. Our study intends to repress these drug-resistance associated *cis*-regulatory pathways in both Dox^{low} and Dox^{high} carcinoma cells.

Figure 4 Copy number analysis of *ABCC1* and *INO80C* loci in Doxorubicin sensitive and tolerant SCLC cells. Dox^{high} and Dox^{low} represent varying levels of tolerance towards cytotoxic drug Doxorubicin. Cells from Dox^{low} H69-150DT were created through selection using incremental concentration of Doxorubicin against drug sensitive NCI-H69 cells and measured for copy number gain against a drug sensitive age match control (H69-150AMC). (A) Dox^{high} and (B) Dox^{low} tolerant cells both demonstrated amplification of *ABCC1* drug transporter gene. (C) Dox^{high} also demonstrated a proportionate upregulation of *INO80C* chromatic remodeler gene. Copy number gains in various cells were plotted with reference to human lymphoblastoid cell line GM12878 (diploid control).

Figure 5 *ABCC1*-Ups gRNA designed to target the enhancer region upstream of *ABCC1* drug transporter encoding gene (red arrow) based on in-house H3K27ac enrichment signals (Input subtracted signal) in Dox^{low} tolerant H69-150DT cells, Dox^{high} tolerant H69AR cells and Dox sensitive NCI-H69 cells. H3K27ac ChIP-Seq signal tracks were uploaded to UCSC genome browser.

Figure 6 ABCC1-Int gRNA designed to target intragenic regions of ABCC1 drug transporter encoding gene (red arrow) based on in-house H3K27ac enrichment signals (Input subtracted signal) in Dox^{low} tolerant H69-150DT cells, Dox^{high} tolerant H69AR cells and Dox sensitive NCI-H69 cells. H3K27ac ChIP-Seq signal tracks were uploaded to UCSC genome browser.

Figure 7 INO80C-Enh gRNA designed to target the enhancer region upstream of INO80C chromatin remodeler protein encoding gene (red arrow) based on in-house H3K27ac enrichment signals (Input subtracted signal) in Dox^{high} tolerant H69AR cells and its Dox sensitive counterparts NCI-H69 cells. H3K27ac ChIP-Seq signal tracks were uploaded to UCSC genome browser.

MATERIALS AND METHODS

gRNA design

The HBB gRNA design was adapted from the previous study on HS2 enhancers (Thakore et al., 2015). We utilized gRNA named CR4 designed to target HS2 enhancer target sequence in the study, which demonstrated consistent downregulation of the globin subunit genes. The genomic location of *TAL1* enhancer was taken from a previous study using the TALEN effector protein array targeted at erythroid enhancer in the *SCL* locus (Mendenhall et al., 2013). We designed the nucleotide sequences of gRNA using the CRISPR mediated editing, repression and activation (ERA) web server tool. Similarly, the gRNA targeted at *ABCC1* and *INO80C* enhancer were designed based on H3K27ac ChIP-seq ‘peaks’ in drug resistant SCLC cell lines in the lab (unpublished data) using the CRISPR-ERA tool (Fig. 5-7).

CRISPR-ERA uses a fast algorithm to search for genome wide gRNA binding sites and evaluates their efficiency and specificity using a set of rules summarized from published data for CRISPR repression. For each target site, CRISPR-ERA first searches all targetable 20 nucleotide sites in humans adjacent to a PAM sequence. Each target sequence is then assessed based on sequence features such as GC content, presence of poly-thymidine gRNA transcription terminator, distance from target gene transcription start sites and genome wide off target binding sites (Liu et al., 2015).

Table 1 Sequences of gRNA and the respective upstream PAM site

Target Enhancer	gRNA sequence (5’—3’)	PAM Sequence
sgHBB	GAAGGTTACACAGAACCAGA	AGG
sgTAL1	TCCGTGGCTGCCAGTCTGAG	GAG
sgABCC1-Ups	GGCAGGTACCCGTCGGCCAA	AGG
sgABCC1-Int	GTCGGCCAATCCTCCCCAGC	AGG
sgINO80C-Enh	GAGGGTTCCGAAAGCATTGC	AGG

Plasmid DNA construction

The hU6 promoter, gRNA scaffold and *Streptococcus pyogenes* dCas9 encoding plasmid pLV-hU6-sgRNA-hUbc-dCas9-KRAB-T2A-Puro is available from Addgene #71236 (MA, USA). The sgRNA cassette was removed via PacI restriction enzyme digestion and re-ligating the vector. For constructing the epigenetic modulators KRAB (**Fig. 8**), LSD1 (**Fig. 9**), HDAC1 (**Fig. 10**) and HP1 α (**Fig. 11**) domain was removed and replaced with other epigenetic modulators using the NheI sites flanking the KRAB domain. The HDAC1 sequence was amplified from GM12878 isolated cDNA. LSD1 was amplified from EMM67 plasmid available from Addgene #49043 (MA, USA). HP1 α was synthesized by Bio-Basic Inc (Ontario, Canada) and cloned in pUC-57 vector.

The hU6 promoter, HBB/TAL1 gRNA and scaffold was synthesized and cloned into plasmid pUC-57 by Bio-Basic Inc (Ontario, Canada). sgRNA for ABCC1/INO80C synthesized and cloned into pU6-sgGFP-NT1 (Addgene #46914 (MA, USA) by Bio-Basic Inc (Ontario, Canada). To shuttle the U6 promoter along with sgRNA scaffold from pU6-sg ABCC1/INO80C vector into PacI site of vector containing dCas9 and epigenetic modifier, PCR primers PacI-U6-sgRNA-F 5'-GGTTAATTAAACAGTGCAGGGGAAAGAATAGTAGACA-3' and PacI-U6-sgRNA-R 5'-GGTTAATTAAATACTAATGCATGGCGGTAATACGGT-3' were used to introduce PacI sites flanking the gRNA expression cassette. The PCR products were cut with PacI enzyme, gel purified and ligated to vector containing dCas9 and epigenetic modifier. DNA Gel Extraction Kit (Bio-Basic Inc, Ontario, Canada) was used to purify the DNA fragment for cloning and was performed according to manufacturer's recommendations. The sequences of all constructs were validated by DNA Sanger sequencing (Axil Scientific, Singapore) prior to transfection.

Figure 8 Schematic illustration of Krüppel associated box (KRAB) domain fused with dead Cas9 protein. The recombinant plasmid encodes for dCas9-KRAB and gRNA expression cassette under the control of Human Ubiquitin C (HubC) promoter and RNA polymerase III U6 promoter respectively. The KRAB domain is further divided into subdomain A and B, both of which work cooperatively to result in transcriptional repression. Neither subdomains possess a catalytic function; instead it serves as a scaffolding platform to recruit repressor complexes to act on the targeted locus, such as KRAB Associated Protein-1 and Transcriptional Intermediary Factor 1β.

Figure 9 Schematic illustration of Lysine-Specific Demethylase 1 (LSD1) fused with dead Cas9 protein. The recombinant plasmid encodes for dCas9-LSD1 and gRNA expression cassette under the control of Human Ubiquitin C (HubC) promoter and RNA polymerase III U6 promoter respectively. The LSD1 enzyme comprise of 3 functionally distinct subdomains. The Flavin Adenine Dinucleotide (FAD) dependent amine oxidase domain confers the enzyme with the demethylase activity and catalyzes reaction on lysine residue 4 on histone 3. The Tower domain acts as a scaffolding domain that can mobilize co-repressors such as co-repressor element silencing factor (coREST) and Nucleosome Remodeling Deacetylase (NuRD).

Figure 10 Schematic illustration of Histone Deacetylase 1 (HDAC1) fused with dead Cas9 protein. The recombinant plasmid encodes for dCas9-HDAC1 and gRNA expression cassette under the control of Human Ubiquitin C (HubC) promoter and RNA polymerase III U6 promoter respectively. HDAC1 is made up of the classical deacetylase catalytic domain conserved across all 3 classes of histone deacetylases, as well as serine residues that are susceptible to phosphorylation by protein kinases.

Figure 11 Schematic illustration of Heterochromatin Protein 1 alpha (HP1α) fused with dead Cas9 protein. The recombinant plasmid encodes for dCas9-HP1α and gRNA expression cassette under the control of Human Ubiquitin C (HubC) promoter and RNA polymerase III U6 promoter respectively. The chromodomain of HP1α acts as an epigenetic reader while the chromoshadow domain trigger the recruitment of epigenetic writers. HP1α is not directly responsible for histone modifications. Instead, its interacting partners such as demethylases SU(VAR)3-9 help to induce stable repression through heterochromatin spreading.

Restriction enzyme digestion

Constructs were subjected to restriction enzyme digestion, in the recommended buffer as listed by New England BioLabs (MA, USA) and incubated at 37°C overnight. For sequential digestion, DNA precipitation was performed in the transition step between the two restriction enzyme reactions. After the first overnight restriction digestion, 10% of 3M Sodium Acetate pH 5.2 and 2.5 volumes of cold absolute ethanol were added. The mixture was incubated at -70°C for 1h, then pelleted down by centrifuging at 18,000g for 20min at 4°C. The pellet was briefly washed with 70% ice-cold ethanol, then centrifuged at 18,000g for 20min at 4°C

again to obtain the precipitated DNA pellet. The DNA pellet was resuspended in nuclease-free water for the succeeding restriction digestion reaction. The buffer used in each reaction mixture is in line with the conditions needed for maximum enzyme activity.

Table 2 Buffer composition for restriction enzyme digestion.

Enzyme	Buffer (1X)	Buffer composition
PacI	CutSmart® Buffer	50mM Potassium Acetate, 20mM Tris-acetate, 10mM Magnesium Acetate, 100µg/ml BSA
NheI	CutSmart® Buffer	50mM Potassium Acetate, 20mM Tris-acetate, 10mM Magnesium Acetate, 100µg/ml BSA
Ascl	CutSmart® Buffer	50mM Potassium Acetate, 20mM Tris-acetate, 10mM Magnesium Acetate, 100µg/ml BSA
NotI	NEBuffer™ 3.1	100mM NaCl, 50mM Tris-HCl, 10mM MgCl ₂ , 100µg/ml BSA
MluI	NEBuffer™ 3.1	100mM NaCl, 50mM Tris-HCl, 10mM MgCl ₂ , 100µg/ml BSA
KpnI	NEBuffer™ 1.1	10mM Bis-Tris-Propane-HCl, 10mM MgCl ₂ , 100µg/ml BSA
SbfI	CutSmart® Buffer	50mM Potassium Acetate, 20mM Tris-acetate, 10mM Magnesium Acetate, 100µg/ml BSA

Ligation and transformation

Vectors (50ng) were ligated to insert DNA at a 1:9 molar ratio, using 1U of T4 DNA ligase (Invitrogen, CA, USA), 2.5mM Adenosine Triphosphate (ATP) and 1X T4 Polynucleotide Kinase Reaction Buffer (70mM Tris-HCl, 10mM MgCl₂, 5mM DTT) (New England BioLabs, MA, US). Temperature cycle ligation was performed in a thermal cycler set to the following: 65°C for 5min; (Gradual 0.1°C ramp decrease to 10°C for 30s; 0.1°C incremental increase to 26°C for 30s; 10°C for 50s) x 50; 16°C for 120min; then cooled to 10°C. Ligated products were then heat-shock transformed

into One Shot™ TOP10 Chemically Competent *E. coli* (Thermo Fisher Scientific, Singapore) for screening.

Plasmid DNA purification

Plasmid DNA was isolated using Plasmid DNA purification Kit (Bio-Basic Inc, Ontario, Canada) for cloning, and purified using Maxiprep kit (Qiagen, Germany) for transfection according to manufacturer's recommendations. Briefly, bacterial cells were collected by centrifugation at 8,000g for 2min then subjected to alkaline lysis with sodium dodecylsulfate (SDS). Lysis reaction was neutralized and genomic DNA was precipitated. The supernatant containing the plasmid DNA was then purified using DNA-binding resin columns and two rounds of ethanol-based washes to remove salt impurities. Lastly, plasmid DNA was eluted in 35µL of TE buffer (10mM Tris-HCl, 1mM EDTA pH 8.0).

PCR amplification

After transformation of ligated clones, bacterial colonies were screened for clones containing the gRNA expression cassette. Colony PCR was performed using primer sets flanking U6-sgRNA. All primers used in this study were synthesized by Integrated DNA Technologies (IA, USA) using following PCR cycling condition: 95°C for 10min; (95°C for 20s; 62°C for 20s; 72°C for 30s) x 35; 72°C for 5min; then cooled to 10°C. PCR products were then analyzed for positive clones using gel electrophoresis. From the selected positive clones plasmid DNA was extracted and subjected to a second PCR reaction to verify the orientation of gRNA expression cassette upstream of the HubC promoter. Following PCR cycling condition was used: 95°C for 5min; (95°C for 20s; 62°C for 20s; 72°C for 30s) x 35; 72°C for 5min; then cooled to 10°C. Same positive clones were also confirmed for the presence of epigenetic modifier using primer sets unique to each epigenetic modifier. For all primer sets same PCR cycling conditions were used except for dCas9 F2 and HDAC1 NheI R: 95°C for 5min; (95°C for 20s; 62°C for 20s; 72°C for 90s) x 35; 72°C for 5min; then cooled to 10°C. All PCR products were analyzed by agarose gel electrophoresis.

Table 3 Primer set sequences for various screening reactions.

Reaction	Forward primer name/ sequence (5'—3')	Reverse primer name/ sequence (5'—3')
Screening for gRNA expression cassette	U6sgRNA-F GGGACAGCAGAGATCCA GTTT	U6sgRNA-R- GTCTGACGTGGCAGCGCT CGC
Validate gRNA insert orientation (sgHBB & sgTAL)	U6sgRNA-F GGGACAGCAGAGATCCA GTTT	U6 pSPgRNA –F TAAGGTACCCGTCTCACAG G
Validate gRNA insert orientation (sgABCC1 & sgINO80C)	U6sgRNA-R- GTCTGACGTGGCAGCGC TCGC	PacI-U6sgRNA R- GGTTAATTAATAACTAATG CATGGCGGTAATACGGT
Screening for sgRNA-dCas9- HDAC1	dCas9-F2- AGAATCGACCTCTCTCAG CTC	HDAC1 NheI R- CTAGCTAGCGGCCAACTTG ACCTCCTCCTT
Screening for sgRNA-dCas9- HP1 α	dCas9-F2- AGAATCGACCTCTCTCAG CTC	HP1 α - R- CTTGTCCCTTAACCACGCG C
Screening for sgRNA-dCas9- KRAB	dCas9-F2- AGAATCGACCTCTCTCAG CTC	KRAB-R- CGAGGCGCACCGTGGGCT TGT

Cell lines and transfections

Human Chronic Myelogenous Leukemia cell line (K562), Small Cell Lung Carcinoma cell line (NCI-H69) and its Doxorubicin drug resistant counterpart (H69AR) were procured from the American Type Culture Collection (ATCC, VA, USA). The NCI-H69 cells were also subjected to gradual, incremental dosage of Doxorubicin drug challenge in-house to select for small cell lung carcinoma cells in low drug tolerant state of 150nM Doxorubicin concentration (H69-150DT). K562, NCI-H69, H69-150DT cells were cultured in RPMI 1640 medium (Nacalai tesque, INC, Japan) containing 2mM L-glutamine, 10% fetal bovine serum and 1% Penicillin Streptomycin (GE Healthcare, IL, USA) kept in a humidified incubator at 37°C containing 5% CO₂. H69AR were cultured in similar conditions with 20% of fetal bovine serum (GE Healthcare, IL, USA).

K562 cells were transfected with dCas9 epigenetic modifier constructs via the Neon® Transfection System (Thermo Fisher Scientific, Singapore) and Lipofectamine 3000 (Invitrogen, CA, USA), according to manufacturer's instructions. Briefly, 1.0×10^6 cells were transfected with 10µg of DNA for Neon® electroporation with recommended settings of 3 pulses at 1450V each, 10ms width apart. Cells were then immediately resuspended in RPMI media at a density of 0.5×10^6 cells/ml. For Lipofectamine transfection, 5.0×10^5 cells were transfected with 4µg of DNA in 250µL of Gibco™ Opti-MEM™. Cells were resuspended in RPMI media at a density of 5.0×10^5 cells/ml for 24h, then in 2.5×10^5 cells/ml for 72h. Approximately 96h after transfection, 1.0×10^6 cells each were harvested for protein and RNA analyses, while the remaining cells were selected using 0.75µg/mL puromycin (#P8833; Sigma-Aldrich, MO, USA) in growth medium.

H69-150DT and H69AR cells were transfected using Lipofectamine 3000 (Invitrogen, CA, USA) as per the manufacturer's instructions. Briefly, 1.0×10^6 H69-150DT suspension cells were transfected with 4µg of DNA in 250µL of Gibco™ Opti-MEM™. Cells were resuspended in RPMI media at a density of 1.0×10^6 cells/ml for 24h, then in 5.0×10^5 cells/ml for another 24h. For H69AR adherent cells, 0.8×10^6 cells were seeded in RPMI media (20% FBS) for 48h prior to transfection. Approximately 48h after transfection, cells were selected using 0.5µg/mL puromycin containing media to enrich the transfected cells. After 96h in puromycin selection media, cells were harvested for RNA analysis.

Cell lysis and western immunoblotting

Cells at a density of 1.0×10^6 cells were centrifuged at 400g for 5min. Cells were lysed in RIPA buffer (10mM Tris-Cl pH 8.0, 1mM EDTA, 0.5mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 1.4M NaCl) containing protease inhibitor cocktail (Sigma-Aldrich, MO, USA) on ice for 1h. Cellular debris were removed by centrifugation at 13,000g for 15min at 4°C. Protein content of the lysates were determined using Qubit protein assay kit as per manufacturer's instructions (Thermo Fisher Scientific, Singapore). 50µg of protein from cellular extracts were boiled with Laemmli sample buffer for 10mins and resolved using sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in SDS Running Buffer (25mM Tris, 192mM Glycine, 3.5mM SDS). The separated proteins were transferred to nitrocellulose membrane by wet transfer in SDS Transfer Buffer (25mM Tris, 192mM Glycine, 20% Methanol) for 120min at 4°C using a constant current of 150mA. Membranes were blocked in 5% non-fat dry milk in PBST (0.1% Tween20 in PBS) for 1h at room temperature. Membranes were then incubated with monoclonal anti-FLAG M2 primary antibody (#SLBQ6349V; Sigma-Aldrich, MO, USA) in 0.2% BSA overnight at 4°C with gentle agitation. After washing in PBST, the membranes were incubated with the horseradish peroxidase conjugated anti-mouse secondary antibody from Dako (Agilent Technologies, Glostrup, Denmark) for 1h at room temperature. Western Lightning® Plus enhanced chemiluminescence reagent (PerkinElmer, MA, USA) was used to visualize the immunoreactive bands using FUJIFILM LAS400IR software (SPD Scientific, Singapore).

Quantitative Reverse Transcriptase-PCR Analysis of genes

Cells (1.0×10^6) were collected by centrifugation at 400g for 5min. Total RNA extraction was performed using EZ-10 DNAaway RNA Miniprep Kit (BioBasic Inc Ontario, Canada) according to the manufacturer's instructions. Reverse transcription polymerase chain reaction (RT-PCR) was performed using MultiScribe™ Reverse Transcriptase (ThermoFisher Scientific, Singapore) in a thermal cycler using following conditions: 25°C for 10min; 37°C for 120min; 85°C for 5min. Quantitative RT-PCR was then performed for *HBE*, *HBG1/2*, *TAL*, *STIL*, *ABCC1*, *FOPNL*, *INO80C*, *GALNT1*, *dCas9*, *sgHBB*, *sgTAL*, *sgABCC1-Ups*, *sgABCC1-Int*, *sgINO80C-Enh*, *Puromycin*, *B2M* and *SDHA* using KAPA SYBR® Fast Universal qPCR Master Mix. The amount of cDNA template utilized in each reaction is 5ng equivalent of total RNA, assuming 100% reverse transcription efficiency. Thermal cycle was set to the

following: 95°C for 3min; (95°C for 15s; 65°C for 30s) x 40; followed by melting curve analysis.

Primers were designed and synthesized using the RealTime qPCR Assay Entry web tool (Integrated DNA Technologies, IA, USA). All primers are targeted at exon spanning sequences to prevent amplification of genomic DNA. The designed primer sequences as indicated below:

Table 4 Primer sets sequences for quantitative analysis of target genes.

Primer name	Forward primer sequence (5'—3')	Reverse primer sequence (5'—3')
HBE1	AGATGAATGTGGAAGAG GCTG	TTACCCAGGAGCTTGAA GTTC
HBG1/2	GCAAGGTGAATGTGGAA GATG	CAGGAGCTTGAAGTTCTC AGG
TAL	ATGCCTTCCCTATGTTCA CC	AGGATCTCATTCTTGCTG AGC
STIL	TGGGAATTTGGCTGTCTG G	ACCGAAAGTCAGGTATCT TGC
ABCC1	CCTATTACCCAGCATCG TG	TGGTGACCTGCAATGAG TAAG
FOPNL	TTGCGTGGAACATAAGGAT GG	GGTGCATTTTGGTAACAT TGGG
INO80C	AATGGTGCCCTCTGAGTT TAG	TCTGTGTAGTTGGCAAGC AG
GALNT1	AACATGACAGGAGAACA GTGG	CCTCCCCAAATATCCATT CCAG
dCas9	TTCAAGACCGAGATTACA CTGG	GCTTGTCGCTGTTTCCTTT TC
gRNA HBB	GAAGGTTACACAGAACC AGA	GCACCGACTCGGTGCCA CTTT (Universal gRNA-R)
gRNA TAL	TCCGTGGCTGCCAGTCT GA	GCACCGACTCGGTGCCA CTTT (Universal gRNA-R)
gRNA ABCC1-Ups	GGCAGGTACCCGTCGGC CAA	GCACCGACTCGGTGCCA CTTT (Universal gRNA-R)

gRNA ABCC1-Int	GTCGGCCAATCCTCCCC A GC	GCACCGACTCGGTGCCA CTTT (Universal gRNA-R)
gRNA INO80C-Enh	GAGGGTTCCGAAAGCAT TGC	GCACCGACTCGGTGCCA CTTT (Universal gRNA-R)
Puromycin	GCTGCAAGAACTCTTCCT CAC	CAGGAGGCCTTCCATCT GTT
B2M	GCTATGTGTCTGGGTTTC ATCC	ACCTCCATGATGCTGCTT AC
SDHA	AGACTACAAGGTGCGGA TTG	ACAGTCAGCCTCGTTCAA AG

Statistical analysis

All graphs are plotted and analyzed using GraphPad Prism 7.0 software. The values in graphs are expressed as mean $\pm$ standard deviation (SD) and P-values were calculated by multiple unpaired student's t-test.

EXPERIMENTAL RESULTS

Specific aim I: Generate recombinant CRISPRi constructs containing HDAC1, HP1 α , LSD1 and KRAB effectors for targeted disruption of gene enhancer

The gRNA expression cassette insert was assembled in different ways, depending on the gRNA type. The sgHBB and sgTAL1 gRNA expression cassette was released from the pUC-57 plasmid and ligated into the pLV-hUbC- dCas9-KRAB-T2A-Puro plasmid in reverse orientation upstream of the HubC promoter via PacI sites (**Fig. 12A & 13A**), in order to ensure no reciprocal interference in RNA polymerase binding between the two promoters. On the other hand, due to the lack of flanking PacI restriction sites in sgABCC1 and sgINO80C plasmids, primers encoding for PacI sites were used to introduce restriction sites in a PCR amplification reaction (**Fig. 12B & 13B**). The PCR product was then subjected to PacI restriction digest to produce sticky ends complementary to the pLV-hUbC- dCas9-KRAB-T2A-Puro plasmid vector (**Fig. 12C & 13C**). This process of isolating complementary gRNA expression cassette for ligation into CRISPR/dCas9 encoding plasmid is repeated for the remaining epigenetic modulator plasmids.

Following T4 DNA ligase mediated cloning reaction, the ligation mixture was transformed into *E. coli* TOP10 cells and plated in LB-agar plate with Ampicillin (**Fig. 14A**). Positive clones were identified through 2 rounds of PCR screening (**Fig. 14B, 14C, 15, 16**), checked with diagnostic restriction digestion (**Fig. 14D & 17**) before validation by Sanger sequencing. This process of plasmid screening was repeated for the remaining gRNA cassettes and dCas9-epigenetic modulator fusion construct. As a result, 25 recombinant plasmids were synthesized (**Fig. 18-20**), which also include the null epigenetic modifier control dCas9-puro.

Figure 12 Schematic illustration of CRISPR-based epigenome editing plasmid assembly. The gRNA expression cassette flanked by *PacI* restriction sites was prepared through either direct *PacI*-directed excision from plasmid (A) or PCR-primer introduced *PacI* sites (B). The gRNA insert was then inserted into a similarly *PacI* cut dCas9-epigenetic modulator encoding plasmid (C) and ligated by T4 DNA ligase. This results in a single 15kb plasmid that encodes both the dCas9/ epigenetic modifier and gRNA expression cassette needed to guide the CRISPRi complex to the target site (D). As a selection marker for transfection, a puromycin resistance cassette was fused to the dCas9 and all epigenetic modulators using a T2A ribosome skipping peptide.

Figure 13 Construction of gRNA and dCas9-epigenetic modifier recombinant plasmids.

(A) Pac1 cut pUC57 vector releases a linearized insert of 439bp comprising of the U6 promoter, *TAL1* gRNA and gRNA scaffold. (B) Alternatively, a pair of Pac1 containing primer set was used to amplify a 730bp fragment of the *ABCC1-Ups* and *INO80C-Enh* gRNA and scaffold from the respective template. (C) The dCas9-epigenetic modifier encoding plasmids were also cut with Pac1 restriction enzymes to create complementary sticky ends suitable for ligation with the gRNA inserts. All restriction enzyme digested fragments and PCR products were analyzed using ethidium bromide stained agarose gel electrophoresis to validate the desired fragment size. All linearized fragments would separate into their actual size on the agarose gel, while the original circularized plasmid would not resolve to size.

Figure 14 Schematic illustration of the screening process for desired clones. The ligation reaction contains a mixture of re-ligated, linear un-ligated and recombinant DNA products. (A) To sieve out the successful gRNA, dCas9-epigenetic modifier recombinant plasmids, the mixture is transformed into *Escherichia coli* bacteria cells and selected using Ampicillin. (B) Colony PCR was performed using primer sets specific to the gRNA expression cassette in order to screen for insertion of gRNA expression cassette. The identified positive clones from the preliminary colony PCR were subjected to DNA extraction and purification. (C) A second PCR was performed to screen for gRNA expression cassettes inserted in the reverse orientation upstream of the HubC promoter. (D) The identified positive clones were put for *PacI* and *NheI* restriction digestion to verify presence of gRNA and epigenetic modifier domain. (E) Sequences of all positive clones were validated by Sanger sequencing.

Figure 15 Detecting recombinant CRISPRi plasmids. PCR products of preliminary colony PCR using primer set flanking the gRNA expression cassette. A positive clone containing *TAL1* gRNA (A) expression cassette appears as 561bp fragment, while a negative clone gives an 114bp fragment. A positive clone containing *INO80C-Enh* gRNA (B) expression cassette appears as 761bp fragment, while a negative clone gives a 122bp fragment. All PCR products were analyzed using ethidium bromide stained agarose gel electrophoresis to validate the desired fragment size.

Figure 16 Determining orientation of insert in recombinant CRISPRi plasmids.

Following DNA extraction of identified positive clones expressing the respective gRNA, an orientation determining PCR was performed using primers set that would only permit amplification in the event of accurate reverse insertion. A positive orientation clone for *TAL1* gRNA (A) would appear as a 441bp fragment, and 664bp fragment for *INO80C-Enh* gRNA (B). All PCR products were analyzed using ethidium bromide stained agarose gel electrophoresis to validate the desired fragment size.

Figure 17 Restriction enzyme validation of recombinant CRISPRi plasmids. Restriction digestion was performed using *PacI* and *NheI* enzymes which have sites flanking the gRNA and epigenetic modulator respectively. The respective gRNA and epigenetic modulators can then be verified as the individual fragments resolve to size. The TAL1 gRNA fragment (A) is 440bp, KRAB domain 315bp, HP1 α 579bp, HDAC1 1449bp and LSD1 2559bp (B). All restriction enzyme digested fragments and PCR products were analyzed using ethidium bromide stained agarose gel electrophoresis to validate the desired fragment size.

Figure 18 Schematic summary of assembled recombinant CRISPRi constructs targeted at the LCR and SCL loci. These plasmids integrate the epigenetic modifiers and gRNAs into single construct. Each cluster contains 4 epigenetic modifiers (HDAC1, LSD1, HP1a, KRAB) fused with gRNAs specific for their intended enhancers. Shown above is the sgHBB (A) and sgTAL1 (B) designed for K562 cells.

Figure 19 Schematic summary of assembled recombinant CRISPRi constructs against the *ABCC1* super-enhancer region. These plasmids integrate the epigenetic modifiers and gRNAs into single construct. Each cluster contains 4 epigenetic modifiers (HDAC1, LSD1, HP1a, KRAB) fused with gRNAs specific for their intended enhancers. Shown above is the sgABCC1-Ups (A) and sgABCC1-Int (B) designed for cell lines H69-150DT and H69AR.

Legend:

Figure 20 Schematic summary of assembled recombinant CRISPRi constructs against the *INO80C* super-enhancer region. These plasmids integrate the epigenetic modifiers and gRNAs into single construct. Each cluster contains 4 epigenetic modifiers (HDAC1, LSD1, HP1a, KRAB) fused with gRNAs specific for their intended enhancers. Shown above is the sglNO80C-Enh designed for H69AR cells.

Specific aim II: Proof-of-concept study to analyze efficacies of CRISPRi constructs using well-established erythroid-specific *cis*-regulatory gene activation models

Validation of dCas9-epigenetic modifier fusion protein expression

As previous studies identified Cas9/sgRNA delivery strategy as a limiting factor to overall engineering efficiency (Liang et al., 2015; Yu et al., 2016), it is important to determine the mode of transfection that may lead to improved outcomes. To this end, we considered the use of two popular DNA delivery systems – electroporation and a cationic lipid-based DNA delivery method.

Under appropriate parameters, electric pulses can induce transmembrane potential differences, as well as field effects on water and membrane dipole. This reversibly destabilizes membrane structures and thereby ensures survival of transiently permeabilized cells (Villemeijane & Mir, 2009). However, electroporation has had varying successes in mammalian cells as uncontrolled high intensity pulses may result in permanent, long lived membrane pores that leads to loss of cell viability (Covello et al., 2014). On the contrary, Lipofectamine is based on lipid nanoparticle directed DNA uptake. When mixed in appropriate buffer conditions, the cationic liposome forms a complex with the negatively charged nucleic acid through electrostatic interactions. The positively charged liposome enables DNA to overcome the initial electrostatic repulsion with the negatively charged cell plasma membrane (Dalby et al., 2004). Compared to other lipid nanoparticle-based technology, Lipofectamine's liposomes are formulated with a neutral co-lipid that assists in liposome-membrane fusion to allow nucleic acid release into the cytoplasm (Dalby et al., 2004). Additionally, Lipofectamine also lead to optimal transfection outcomes by mediating the penetration through the nuclear membrane and steering away from inefficient intracellular DNA transport mechanisms (Cardarelli et al., 2016).

To draw comparisons between plasmid delivery systems, the sequenced positive clones were transfected into K562 cells using Neon® electroporation system as well as Lipofectamine 3000 (Thermo Fisher Scientific, Singapore).

Following transfection of the CRISPRi constructs, we sought to determine if the recombinant plasmid alone is sufficient for dCas9-epigenetic fusion protein expression. The transfected cells were immunoblotted with antibody against the FLAG epitope upstream of the dCas9-epigenetic modulator. After 4 days post-transfection, the Neon® transfected cells were collected for protein analysis (**Fig. 21A & 21B**). The Lipofectamine transfected cells were further subjected to puromycin stable selection before harvesting the cells for dCas9/dCas9-fusion protein expression confirmation (**Fig. 21C-21F**). While our immunoblotting results indeed demonstrated the expression of dCas9 fusion proteins in transfected cells, the level of expression was not uniform. Expression intensities of the bulkier dCas9-HDAC1 and dCas9-LSD1 proteins were consistently observed to be lesser than that of the other samples, irrespective of the transfection method. The unequal expression of the fusion constructs suggests a dependency upon the epigenetic modulator type and size.

The Neon® transfected cells demonstrated limited protein expression of dCas9 and epigenetic modulator (**Fig. 21B**). Of note is the low protein expression in sgTAL-dCas9-HDAC1 and sgTAL-dCas9-LSD1 transfected cells. Moreover, apart from the protein band of expected size, additional bands of considerable intensities were noticed in some of the Neon® transfected cells. The sgHBB-dCas9-HDAC1 and sgHBB-dCas9-LSD1 Neon® transfected cells exhibited 2 protein fragments, with molecular weights corresponding to that of dCas9 protein and dCas9-HDAC1/LSD1 fusion protein. This may be indicative of partial degradation of the fusion protein construct (**Fig. 21A**). Interestingly, immunoblotting analyses of stably selected clones from the Lipofectamine transfected counterparts showed only a single protein band of expected molecular weight in appropriate cell lysates (**Fig. 21C-21F**). While previous studies have shown electroporation's superior transfection efficiency (Liang et al., 2015), our results suggest that a lipid nano-particle directed delivery of plasmid may minimize the effects of protein degradation. This allows for a more accurate comparison of the epigenetic effector proteins.

Nevertheless, the immunoblotting results assure the expression of the respective dCas9-epigenetic effector protein, thereby verifying the presence of CRISPRi complexes in transfected cells.

Figure 21 Expression of dCas9- epigenetic modifier fusion proteins in K562 cells. (A) After 96h of NEON electroporation transfection, dCas9-sgHBB cell lysates were subjected to immunoblotting analysis. (B) Similarly, dCas9-sgTAL electroporated cell lysates were taken for protein analysis after 96h. (C, D) Post transfection by Lipofectamine, sgHBB transfected cells were put in selection media containing 0.75ug/mL puromycin. Stable cell lysates from the erythroid K562 cells were then subjected to immunoblotting analysis. (E, F) Similarly, dCas9-sgTAL electroporated cell lysates were taken for protein analysis after 96h. In all the set ups, the expected protein sizes were estimated by taking the molecular weight per amino acid as 110Da: dCas9-Puro: 155kDa, dCas9-HDAC1: 208kDa, d-cas9-LSD1: 249kDa, dCas9-HP1a: 176kDa, dCas9-KRAB: 167kDa. UT represents untransfected cell lysates.

Quantifying knockdown activity of CRISPRi constructs

After the successful introduction of CRISPRi complexes via cationic lipid-based plasmid delivery means, we examined the functional consequence of this intervention. In a proof-of-concept study, the activity of each epigenetic effector was screened by qRT-PCR analysis of globin subunit genes *HBG* and *HBE* expression to demonstrate the disruption of HS2 enhancer by different CRISPRi systems. Due to high sequence similarity, the designed primers cannot distinguish between *HBG1* and *HBG2* transcripts. As *HBB* is not highly expressed in CML K562 (Dean et al., 1983), its expression was not interrogated for comparison.

Transcript quantification of *HBB* gRNA transfected cells revealed lower mRNA expression levels in cells expressing both epigenetic effector and gRNA, relative to the control expressing dCas9 alone. This suggests that the combined activity of dCas9 and epigenetic modifier domain repressed expression of globin subunit genes. Additionally, sgHBB-dCas9-LSD1 was shown to have the strongest effect on repression of *HBE* and *HBG1/2* genes (**Fig. 22A & 22B**). Relative to the respective non-gRNA transfected control, *HBE* and *HBG1/2* gene expression was reduced by 50-80% by LSD1. The dCas9-KRAB complex shows similar repressive potency of 50-70% on both *HBE* and *HBG1/2*. In contrast, the sgHBB-dCas9-HDAC1 and sgHBB-dCas9-HP1 α transfected cells show modest repression of 30-60% and 20-40% respectively on both globin subunit genes.

Primers against *dCas9* and *gRNA* were utilized to measure levels of plasmid uptake by the K562 cells (**Fig. 22C**). Our results reveal a large variation in dCas9 and gRNA expression, implying a discrepancy in recombinant plasmid uptake for different epigenetic modifier protein. Alternatively, there could also be differential stability of dCas9 protein and sgRNA when embedded in different context. In both cases, it would cause disparities in the ratio of available dCas9/sgRNA complex, which in turn affects repression efficiency.

Similarly, we also targeted the dCas9-epigenetic effector complex using sgTAL1 guide RNA to repress the erythroid enhancer in K562 cells. Quantification of erythroid enhancer-driven genes *TAL1* and *STIL* transcripts revealed no significant repression activity. Differences in mRNA expression were not observed across all the epigenetic modifiers (**Fig. 23A & 23B**), despite reasonable levels of *dCas9* and *gRNA* expression (**Fig. 23C & 23D**). This could be attributed to poor gRNA design, as the TAL1 gRNA carries the low Cas9 affinity 5' 'GAG' PAM sequence (Hsu, et al., 2013), possibly leading to sub-optimal target binding efficiency.

Nevertheless, the results demonstrate that the novel dCas9-HP1 α and dCas9-HDAC1 epigenetic modifier tools are capable of interfering with HS2 enhancer function, when coupled with appropriate gRNA design. Data from the globin gene locus clearly show that when a distal (>10 kb) enhancer is targeted by a single, specific gRNA, the expression of multiple target genes can be repressed.

Figure 22 Quantification of globin gene expression in K562 cells transfected with HBB gRNA. RNA from erythroid cells K562 were purified and quantified using RT-qPCR against the Succinate Dehydrogenase Complex Subunit A (SDHA) reference gene. (A) The $\Delta\Delta\text{Ct}$ method was utilized for relative quantification of *HBE* and (B) *HBG* globin subunit genes in the transfected cells against untreated cells. Fold change was then calculated relative to the respective non-gRNA transduced control for each epigenetic effector protein. Error bars indicate $\pm\text{SD}$ for two technical replicates. $^{ns}\text{P} > 0.05$, $^*\text{P} < 0.05$, $^{**}\text{P} < 0.01$, $^{***}\text{P} < 0.001$ and $^{****}\text{P} < 0.0001$ as compared to the respective controls. (C) The ΔCt method was utilized for relative quantification of *dCas9* and (D) *gRNA* expression in the transfected cells. Error bars indicate $\pm\text{SD}$ for $n=2$ biological replicates, each with two technical replicates. UT and MT represent untransfected samples and mock transfection respectively.

Figure 23 Quantification of SCL locus gene expression in K562 cells transfected with TAL1 gRNA. RNA from erythroid cells K562 were purified and quantified using RT-qPCR against the Succinate Dehydrogenase Complex Subunit A (SDHA) reference gene. (A) The $\Delta\Delta Ct$ method was utilized for relative quantification of *TAL1* and neighboring gene (B) *STIL* expression in the transfected cells against untreated cells. Fold change was then calculated relative to the respective non-gRNA transduced control for each epigenetic effector protein. (C) The ΔCt method was utilized for relative quantification of *dCas9* and (D) *sgTAL1* expression in the transfected cells. In all data sets, error bars indicate $\pm$ SD for technical replicates. ^{ns}P > 0.05 as compared to the respective controls. UT and MT represent untransfected samples and mock transfection respectively.

Specific aim III: Examine efficacies of CRISPRi constructs in repressing transcription of novel variant enhancer-driven genes differentially expressed in a chemoresistant model

Preliminary proof-of-concept experimentation on the well-characterized globin gene locus validated the repressive potency of newly synthesized dCas9-HP1 α and dCas9-HDAC1 fusion proteins. We then moved on to determine the performance of these CRISPRi repressor tools when embedded in novel gene regulatory networks. Cells were transfected with these constructs to target 'candidate' enhancer regions of genes involved in drug resistance using SCLC cells with varying drug sensitivity. Similarly, RT-qPCR was utilized in assessing the repression of the target *ABCC1* and *INO80C* genes, both of which were overexpressed when cells acquire the drug resistance phenotype and their expressions were driven by variant enhancers. In addition, neighboring genes potentially under the regulatory control of the same enhancer were also quantified for their transcript levels. This includes gene *FOPNL* (60kb upstream of *ABCC1*) and gene *GALNT1* (84kb downstream of *INO80C*).

Quantifying knockdown activity of CRISPRi constructs: Dox^{low} tolerant state

Two *ABCC1* gRNA were designed to target separate regions within the locus, each postulated to be critical in enhancer function. Knockdown of enhancer-driven gene expression was quantified using a series of gene-specific primers designed to detect downregulation of the respective mRNA transcripts.

When the CRISPRi complex was targeted upstream of the *ABCC1* gene in Dox^{low} tolerant cells, only the dCas9-HP1 α tool demonstrated significant repressive potency. While the steric hindrance effect of dCas9 alone did cause repression of *ABCC1* transporter gene, the effect was found to be insignificant after correcting for multiple t test. The only significant activity observed was from HP1 α , with a 50% knockdown of *ABCC1* transcript (**Fig. 24A**) and 40% reduction in neighboring gene *FOPNL* transcripts (**Fig. 24B**). On top of monitoring dCas9/sgRNA expression levels, RT-PCR with primers against puromycin marker were also used as an indicator of plasmid uptake. The observed repressive activity stemmed from modest levels of plasmid uptake (**Fig. 24E**) and dCas9/sgRNA expression (**Fig. 24C & 24D**) in the HP1 α transfected cells.

The CRISPRi complex was also directed to *ABCC1* intragenic enhancer regions in DoxI^{ow} tolerant cells. Transcript quantification of *ABCC1* and *FOPNL* demonstrate similar repression by the dCas9-HP1 α tool (**Fig. 25A & 25B**). The transcriptional inhibition by dCas9 alone through steric occlusion mechanisms is also more pronounced when directed to *ABCC1* intragenic regions instead of upstream regulatory elements. Additionally, the KRAB domain also revealed moderate repression of 60% in both *ABCC1* and *FOPNL* when coupled with *ABCC1*-Int gRNA. This elevated repression potency could be attributed to astoundingly high levels of dCas9/sgRNA and recombinant plasmid present in KRAB transfected cells (**Fig. 25C-25E**). Different sized constructs are known to exhibit differential nuclear delivery efficiency, despite using equivalent molar mass concentrations of plasmids during transfection (Lukacs et al., 2000). Small plasmids are postulated to have shorter transit times in the cytoplasm, thereby evading degradation and show higher nuclear delivery efficacy. In contrast, large plasmids are limited by inefficient cytoplasmic trafficking that lead to lower successful nuclear deliveries (McLenachan et al., 2007). As the recombinant plasmid encoding dCas9-KRAB is 1kb and 2kb smaller than the dCas9-HDAC1 and dCas9-LSD1 plasmids respectively, it would be reasonable to expect a higher plasmid uptake efficiency in KRAB transfected cells.

Regardless of gRNA targeting to upstream or intragenic enhancers, dCas9 expression levels in HDAC1 transfected cells were found to be low in both data. This is owing to rampant cell death observed a few days after dCas9-HDAC1 plasmids were introduced. Without a gRNA to orientate the CRISPRi complex, widespread transcriptional repression may have occurred in these cells, resulting in cellular apoptosis.

Taken together, this result is in line with previous work that have identified the *gRNA* complementary region in *ABCC1* locus as a crucial enhancer region capable of augmenting target gene expression in resistant cells. Thus, coupling dCas9-epigenetic effector domains with *gRNA* to the aforementioned target can deactivate gene regulatory networks and result in repression of multiple genes.

Figure 24 Quantification of ABCC1 locus gene expression in low Doxorubicin tolerant cells transfected with ABCC1-Ups gRNA. RNA from SCLC H69-150DT were purified and quantified using RT-qPCR against the Succinate Dehydrogenase Complex Subunit A (SDHA) reference gene. (A) The $\Delta\Delta C_t$ method was utilized for relative quantification of *ABCC1* and neighboring (B) *FOPNL* gene expression in the transfected cells against untreated cells. Fold change was then calculated relative to the respective non-gRNA transduced control for each epigenetic effector protein. Error bars indicate $\pm$ SD for two technical replicates. ^{ns}P > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001 as compared to the respective controls. (C) The ΔC_t method was utilized for relative quantification of *dCas9*, (D) *gRNA* and (E) Puromycin marker expression in the transfected cells. Error bars indicate $\pm$ SD for n=2 biological replicates, each with two technical replicates. UT and MT represent untransfected samples and mock transfection respectively.

Figure 25 Quantification of ABCC1 locus gene expression in low Doxorubicin tolerant cells transfected with ABCC1-Int gRNA. RNA from SCLC H69-150DT were purified and quantified using RT-qPCR against the Succinate Dehydrogenase Complex Subunit A (SDHA) reference gene. (A) The $\Delta\Delta Ct$ method was utilized for relative quantification of *ABCC1* and neighboring (B) *FOPNL* gene expression in the transfected cells against untreated cells. Fold change was then calculated relative to the respective non-gRNA transduced control for each epigenetic effector protein. Error bars indicate $\pm$ SD for two technical replicates. $^{ns}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$ as compared to the respective controls. (C) The ΔCt method was utilized for relative quantification of *dCas9*, (D) *gRNA* and (E) Puromycin marker expression in the transfected cells. Error bars indicate $\pm$ SD for $n=2$ biological replicates, each with two technical replicates. UT and MT represent untransfected samples and mock transfection respectively.

Quantifying knockdown activity of CRISPRi constructs: Dox^{high} tolerant state

The CRISPRi constructs were also assessed for their repressive potential in Dox^{high} tolerant SCLC cells. Two gRNAs were directed to the proposed *ABCC1* enhancer at intragenic and intergenic regions, as well as a single gRNA designed to orient dCas9 binding to a postulated enhancer within the *INO80C* locus.

In Dox^{high} tolerant cells, targeting the epigenetic effectors to both upstream and intragenic regions did not instigate significant transcriptional repression in *ABCC1* (**Fig. 26A & 27A**). When dCas9 alone was targeted to *ABCC1* upstream and intragenic regions, no significant changes in transcripts was noticed, possibly due to limited steric occlusion of RNA polymerase. As the *ABCC1* transcripts are highly amplified in the Dox^{high} tolerant cells (**Fig. 4A**), enhancers in such instances may experience crowding of transcription factors needed to facilitate this amplification (Matsuda et al., 2014). Consequently, chromatin remodeling machinery recruited by the epigenetic modifiers may be spatially prevented from acting on the enhancer sequence. With much lower initial expression levels, modest knockdown activity of approximately 20-30% was observed in *FOPNL* transcripts by HDAC1 and LSD1 domains (**Fig. 26B & 27B**). Consistent with the quantification results in Dox^{low} tolerant cells, dCas9 and gRNA expression remained high for smaller plasmids. Widespread cell death was also observed in the dCas9-HDAC1 transfected cells, resulting in a low dCas9 expression measurement.

Interestingly, HDAC1 and KRAB effector domains were revealed to contribute slight *ABCC1* transcript upregulation when coupled with gRNA (**Fig. 26A & 27A**). This supports the notion of compensatory mechanisms present to counter the dCas9-epigenetic modifier activity (Stojic et al., 2018). Moreover, there is extensive amplification of the *ABCC1* locus, including the gene and associated enhancers. Our study targeted only a single enhancer in closest proximity to the *ABCC1* gene, amongst the clusters of enhancers (super-enhancer) in the locus. Thus, the remaining untargeted enhancers could continue to sustain robust *ABCC1* expression. Improved understanding of the molecular mechanism that reinforce the strong activation of *ABCC1* gene in Dox^{high} tolerant cells is required for improved CRISPRi complex design.

On the contrary, directing the CRISPRi complex to the postulated enhancer downstream of *INO80C* in Dox^{high} tolerant cells yielded considerable repression in *INO80C* but not in neighboring gene *GALNT1* transcripts. The inhibitory action of HDAC1 produced a small reduction of 20-30% in *INO80C* mRNA (**Fig. 28A**) but did not result in significant repression on *GALNT1* (**Fig. 28B**). Consistent with what was observed in the *ABCC1* locus, the remaining epigenetic domains KRAB, LSD1 and HP1α caused slight upregulation of *INO80C* mRNA, likely attributed to compensatory gene activation mechanisms. Alternative enhancers in the *INO80C* locus may have to be concurrently inactivated for a compelling transcriptional repression to ensue. No significant changes in *GALNT1* transcript were noticed across the various epigenetic effectors, indicating that its expression may not be under the control of the same enhancer as *INO80C*.

Thus, the propensity for CRISPRi complexes to efficiently repress multiple genes may not be confined to cell-type specificity. The effectiveness of each epigenetic effector type in instigating histone modifications consequential on the cell's transcriptional program may be locus and enhancer-specific.

Figure 26 Quantification of ABCC1 locus gene expression in high Doxorubicin tolerant cells transfected with ABCC1-Ups gRNA. RNA from SCLC H69AR were purified and quantified using RT-qPCR against the Succinate Dehydrogenase Complex Subunit A (SDHA) reference gene. (A) The $\Delta\Delta Ct$ method was utilized for relative quantification of *ABCC1* and neighboring (B) *FOPNL* gene expression in the transfected cells against untreated cells. Fold change was then calculated relative to the respective non-gRNA transduced control for each epigenetic effector protein. (C) The ΔCt method was utilized for relative quantification of *dCas9*, (D) *gRNA* and (E) Puromycin marker expression in the transfected cells. In all data sets, error bars indicate $\pm$ SD for two technical replicates. nsP >0.05, **P <0.01 and ***P <0.001 as compared to the respective controls. UT and MT represent untransfected samples and mock transfection respectively.

Figure 27 Quantification of ABCC1 locus gene expression in high Doxorubicin tolerant cells transfected with ABCC1-Int gRNA. RNA from SCLC H69AR were purified and quantified using RT-qPCR against the Succinate Dehydrogenase Complex Subunit A (SDHA) reference gene. (A) The $\Delta\Delta C_t$ method was utilized for relative quantification of *ABCC1* and neighboring (B) *FOPNL* gene expression in the transfected cells against untreated cells. Fold change was then calculated relative to the respective non-gRNA transduced control for each epigenetic effector protein. (C) The ΔC_t method was utilized for relative quantification of *dCas9*, (D) *gRNA* and (E) Puromycin marker expression in the transfected cells. In all data sets, error bars indicate $\pm$ SD for two technical replicates. ^{ns}P > 0.05, *P < 0.05 and **P < 0.01 as compared to the respective controls. UT and MT represent untransfected samples and mock transfection respectively.

Figure 28 Quantification of INO80C locus gene expression in high Doxorubicin tolerant cells transfected with INO80C-Enh gRNA. RNA from SCLC H69AR were purified and quantified using RT-qPCR against the Succinate Dehydrogenase Complex Subunit A (SDHA) reference gene. (A) The $\Delta\Delta C_t$ method was utilized for relative quantification of *INO80C* and neighboring (B) *GALNT1* gene expression in the transfected cells against untreated cells. Fold change was then calculated relative to the respective non-gRNA transduced control for each epigenetic effector protein. Error bars indicate $\pm$ SD for two technical replicates. ^{ns}P > 0.05, *P < 0.05 and **P < 0.01 as compared to the respective controls. (C) The ΔC_t method was utilized for relative quantification of *dCas9*, (D) *gRNA* and (E) Puromycin marker expression in the transfected cells. Error bars indicate $\pm$ SD for n=2 biological replicates, each with two technical replicates. UT and MT represent untransfected samples and mock transfection respectively.

DISCUSSION

Novel dCas9-HP1 α and dCas9-HDAC1 mediated repression

In this work, we have harnessed the repressive properties of epigenetic modulators, coupled with CRISPR/Cas9's ability to direct site-specific binding to downregulate enhancer-driven gene expression. Each of these dCas9-epigenetic modulator systems was assessed for their ability to decommission well-characterized cis-regulatory networks. We assembled 25 CRISPR interference recombinant plasmids, targeted at 5 distinct enhancer regions. By measuring the level of enhancer-driven transcript expression, we found that dCas9-LSD1 and dCas9-KRAB complex demonstrate strong repression of the HS2 enhancer activity and consequently, globin gene knockdown. Although this was found to be in line with previous studies (Thakore et al., 2015), it is important to note that the repressive property and efficiency are organism, cell type, mutation type and target dependent (Xu et al., 2014; Kearns et al., 2015). Therefore, we next interrogated the repressive properties of the dCas9/epigenetic modulator complexes on genes implicated in chemoresistance.

We targeted the CRISPRi constructs to a variant enhancer at the *ABCC1* locus, postulated to drive *ABCC1* drug transporter gene expression due to its close proximity. In the chemoresistant SCLC model, we found that the dCas9-HP1 α complex is capable of repressing enhancer driven *ABCC1* gene expression in Dox^{low} tolerant cells, while INO80C enhancer activity can be suppressed by dCas9-HDAC1 in Dox^{high} tolerant cells. Heterogeneity in response towards various epigenetic modulator types suggest that the repressive complex composition may vary for different genes. Hence, this reinforces the need for an assortment of CRISPR based epigenetic modifying tools. In the context where dCas9-KRAB or dCas9-LSD1 fails to exert repressive functions, dCas9-HP1 α and dCas9-HDAC1 may serve to fill in these loopholes.

Recent studies have found evidence of CRISPRi complex's markedly improved binding specificity when targeting dCas9 to endogenous promoters and enhancers (Thakore et al., 2015; Kearns et al., 2015). To ascertain if fusion of dCas9 with HP1 α and HDAC1 will alter binding specificity, ChIP-Seq using the FLAG epitope can be performed to analyze the genome-wide off target binding. Additionally, RNA-seq for a global transcription profiling should also be performed to determine if the observed repression is specific to the targeted enhancer. Conducting a ChIP-qPCR analysis,

using H3K9me3 or H3K27me3 specific antibody would assess the degree of chromatin condensation at the targeted site, which will ensure CRISPRi complex formation. This would substantiate the conclusion that observed repression is a direct effect of the transfected CRISPRi constructs. Most importantly, the functional consequence of this CRISPRi mediated knockdown of cancer drug resistance genes must be determined. The transfected SCLC cells demonstrating effective enhancer-driven gene repression can be evaluated for their re-sensitization to Doxorubicin through quantifying cell viability differences upon drug exposure.

Limiting factors to CRISPR-based epigenetic modulating tools

The epigenetic manipulation of cellular transcriptional program with no direct modification of genomic sequences has been a long-standing goal in chromatin studies. While existing studies provide the tangible support for the potential applications of CRISPR-based technologies, they also underscore the fact that it has yet to come to full term. CRISPR-based tools have important caveats that need to be addressed to enhance its validity. The factors that affect efficiency of CRISPR interference complex can be broadly classified into 4 types: Cas9 component, gRNA design, Cas9 to gRNA ratio and epigenetic modulator promiscuity. The efficiency and specificity of CRISPRi can be optimized by regulating each component's expression and improving the design criteria.

Firstly, as the *Streptococcus pyogenes* Cas9 protein is prokaryotic derived, it must be tethered to a nuclear localization signal (NLS) on the amino or carboxyl terminus in order to direct protein translocation to the nucleus (Qi et al., 2013). However, the NLS may get buried during Cas9 protein folding, resulting in inaccurate subcellular localization. Recent studies reveal enhanced DNA editing activity through the addition of a 32-amino acid linker between Cas9 and the NLS peptide (Shen et al., 2013). This could be due to an exposed NLS peptide that contributes to improved nuclear localization.

Secondly, the degree of complementarity between gRNA and target protospacer sequence is a crucial aspect of Cas9 based transcriptional control. While the entire gRNA contributes to its overall specificity, Cas9 was found to be sensitive to the position and number of nucleotide mismatch. Mismatch in the 3' half of the target recognition sequence is not well-tolerated and severely impair or even abolish Cas9 cleavage functioning (Jiang et al., 2013). Interestingly, only a fraction of complementary sequence is required for the Cas9/gRNA to be semi-transiently

bound to multiple DNA sites with moderate homology (Wu et al., 2014). Despite binding at multiple off-target sites, Cas9/gRNA complex rarely cleaves these unintended sites, giving specificity to Cas9-based genome editing. However, this off-target binding would present as a problem in the use of catalytically null dCas9 fusion proteins, which may remodel chromatin architecture at unsolicited sites. However, other studies have also demonstrated that CRISPRa and CRISPRi complexes that do not bind within a narrow range around the transcription start site fail to impact transcription (Gilbert et al., 2014a; Wu et al., 2014). This greatly restricts the number of off-target binding events that will result in meaningful activity. Off-target effect is a common challenge for site-directed genome editing tools. The potential off-target events in our study were computationally profiled by the CRISPR-ERA web server tool, which optimized the gRNA selection rules (Liu et al., 2015). However, such programs should also consider incorporating thermodynamics characterization of gRNA and target DNA binding to allow for a more quantitative-based assessment and understanding of binding degeneracy (Hsu et al., 2014). Additionally, the dCas9 based transcriptional blocking was found to display strand specificity, as gRNAs targeting the non-template DNA strand showed higher repression efficiency than those against the template strand (Qi et al., 2013). Greater understanding of the Cas9/gRNA recognition features and functional dependency on the PAM sequence will also help to fine tune target selection criteria in these computational tools.

Thirdly, studies have found that altering the Cas9 to gRNA molecular ratio has consequential effects on genome editing. By supplementing increased gRNA concentration, more Cas9/gRNA complexes can be formed and hence, a higher cleavage efficiency can be expected (Wang et al., 2013). The performance of dCas9 effector dependent transcriptional regulation also shows gRNA concentration accumulation effects. Certain DNA delivery systems may allow for greater control of the Cas9/gRNA molecular ratio introduced into cells, such as lenti-viral system. In several CRISPRa studies, co-transfecting multiple gRNAs tiled across the target gene sequence induced multifold transcriptional enhancements (Cheng et al., 2013; Mali et al., 2013a). Unfortunately, this would also be accompanied by a concomitant increase in off-target effects due to inevitable low complementarity of non-specific genomic regions (Fu et al., 2013). While multiple gRNAs binding to distinct sites of the same gene results in an augmentative silencing effect, cautious design is

necessary to ensure no overlap in binding sites, which leads to suppressed repression (Qi et al., 2013; Mali et al., 2013a; Xu et al., 2014).

Lastly, contrary to conventional Cas9 tools, the off-target effect of CRISPR-directed epigenome editing technologies is not exclusive to the DNA binding activity of the gRNA. The innately promiscuous nature of epigenetic effectors can cause them to catalyze a series of off-target reactions near the intended target. Limited understanding of these epigenetic effector complexes and their secondary effects will continue to be a key obstacle that undermines the validity of their use. Studies on epigenetic domains such as LSD1 (Forneris et al., 2005) and histone acetyltransferases (Hilton et al., 2015) have unveiled cross-talk between post-translational modifications to orchestrate the complex transcriptional program. Inserting an epigenetic modifier will affect this cross-talk dynamics, thereby changing the 3-dimensional chromatin structure in ways more than merely a single cluster of histone marks alteration. This cross-talk is more pronounced for KRAB and HP1 α domains, which are functionally less straightforward than LSD1 and HDAC1 enzymes that directly catalyze repressive marks on histones. KRAB and HP1 α proteins serve as scaffolding platforms to further recruit heterochromatin-forming complexes. Thus, the control over transcriptional outcome is certainly more challenging, as it involves extensive understanding of the mechanistic contribution by each recruited protein component. In addition, the epigenetic enzymes such as histone demethylases and deacetylases have varying degrees of substrate specificity (Huang et al., 2007; Cho et al., 2011; Wang et al., 2009; Haberland et al., 2009). This low fidelity characteristic could elicit an unforeseen deleterious secondary effect accompanying the desired silencing event. Taken together, even if CRISPRi repressive complexes are accurately directed to the target locus, the cascade of secondary events and varying substrate specificity may drastically change the local chromatin architecture by unknown processes. Thus, functional characterization of epigenetic effectors and relevant catalytic partners in their isolated states as well as in the complexes they form will be an important step towards addressing the off-target event obstacles. Resolving the structural information and careful protein engineering could also help to improve substrate specificity.

CONCLUSION

These Cas9 based tools embody many crucial advantages that previous targeted tools cannot match. Since its inception, CRISPR systems have revolutionized genome editing studies and are in a good position to exert a similar influence in the budding field of epigenome editing. Overcoming the identified shortcomings, optimizing protocol design, and continued forecasting of future research opportunities beyond CRISPR directed systems could go a long way to building the next generation of epigenome editing tools. Our study expands on and combines novel epigenetic modifier with the CRISPR system to improve regulatory control of endogenous gene control. Armed with a broadened library of advanced epigenome editing tools, we believe that our study opens alternative areas of pursuit for understanding the relationship shared between regulatory elements, complex chromatin architecture and cellular phenotypes. Thus, the functional diversity of CRISPR-dCas9 systems will continue to be deeply valued in both basic research and translational medicine.

REFERENCES

- Abrink, M., Ortiz, J., Mark, C., Sanchez, C., Looman, C., Hellman, L., Chambon, P. and Losson, R. (2001). Conserved interaction between distinct Kruppel-associated box domains and the transcriptional intermediary factor 1. *Proceedings of the National Academy of Sciences*, 98(4), pp.1422-1426.
- Anand, R. and Marmorstein, R. (2007). Structure and Mechanism of Lysine-specific Demethylase Enzymes. *Journal of Biological Chemistry*, 282(49), pp.35425-35429.
- Barde, I., Laurenti, E., Verp, S., Groner, A., Towne, C., Padrun, V., Aebischer, P., Trumpp, A. and Trono, D. (2009). Regulation of Episomal Gene Expression by KRAB/KAP1-Mediated Histone Modifications. *Journal of Virology*, 83(11), pp.5574-5580.
- Beerli, R., Dreier, B. and Barbas, C. (2000). Positive and negative regulation of endogenous genes by designed transcription factors. *Proceedings of the National Academy of Sciences*, 97(4), pp.1495-1500.
- Bellefroid, E., Poncelet, D., Lecocq, P., Revelant, O. and Martial, J. (1991). The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proceedings of the National Academy of Sciences*, 88(9), pp.3608-3612.
- Bitinaite, J., Wah, D., Aggarwal, A. and Schildkraut, I. (1998). FokI dimerization is required for DNA cleavage. *Proceedings of the National Academy of Sciences*, 95(18), pp.10570-10575.
- Cano-Rodriguez, D., Gjaltema, R., Jilderda, L., Jellema, P., Dokter-Fokkens, J., Ruiters, M. and Rots, M. (2016). Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. *Nature Communications*, 7, p.12284.
- Cardarelli, F., Digiacomo, L., Marchini, C., Amici, A., Salomone, F., Fiume, G., Rossetta, A., Gratton, E., Pozzi, D. and Caracciolo, G. (2016). The intracellular trafficking mechanism of Lipofectamine-based transfection reagents and its implication for gene delivery. *Scientific Reports*, 6(1).
- Chavez, A., Tuttle, M., Pruitt, B., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S., Cecchi, R., Kowal, E., Buchthal, J., Housden, B., Perrimon, N., Collins, J. and Church, G. (2016). Comparison of Cas9 activators in multiple species. *Nature Methods*, 13(7), pp.563-567.
- Chen, Y., Feng, B. and Chen, L. (2012). Update of Research on Drug Resistance in Small Cell Lung Cancer Chemotherapy. *Asian Pacific Journal of Cancer Prevention*, 13(8), pp.3577-3581.
- Chen, Y., Yang, Y., Wang, F., Wan, K., Yamane, K., Zhang, Y., Lei, M. (2006). Crystal structure of human histone lysine-specific demethylase 1 (LSD1). *Proc Natl Acad Sci USA*, 103(38), pp. 13956–13961.
- Cheng, A., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T., Rangarajan, S., Shivalila, C., Dadon, D. and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Research*, 23(10), pp.1163-1171.
- Cho, H., Suzuki, T., Dohmae, N., Hayami, S., Unoki, M., Yoshimatsu, M., Toyokawa, G., Takawa, M., Chen, T., Kurash, J., Field, H., Ponder, B., Nakamura, Y. and Hamamoto, R. (2010). Demethylation of RB Regulator MYPT1 by Histone Demethylase LSD1 Promotes Cell Cycle Progression in Cancer Cells. *Cancer Research*, 71(3), pp.655-660.
- Cole, S., Bhardwaj, G., Gerlach, J., Mackie, J., Grant, C., Almquist, K., Stewart, A., Kurz, E., Duncan, A. and Deeley, R. (1992). Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science*, 258(5088), pp.1650-1654.
- Cong, L., Ran, F., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P., Wu, X., Jiang, W., Marraffini, L. and Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121), pp.819-823.

- Covello, G., Siva, K., Adami, V. and Denti, M. (2013). An electroporation protocol for efficient DNA transfection in PC12 cells. *Cytotechnology*, 66(4), pp.543-553.
- Dalby, B., Cates, S., Harris, A., Ohki, E., Tilkins, M., Price, P. and Cirraccione, V. (2004). Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods*, 33(2), pp.95-103.
- Dawson, M. and Kouzarides, T. (2012). Cancer Epigenetics: From Mechanism to Therapy. *Cell*, 150(1), pp.12-27.
- Dean, A., Ley, T. J., Humphries, R. K., Fordis, M., & Schechter, A. N. (1983). Inducible transcription of five globin genes in K562 human leukemia cells. *Proceedings of the National Academy of Sciences of the United States of America*, 80(18), 5515–5519.
- Delcuve, G., Khan, D. and Davie, J. (2012). Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. *Clinical Epigenetics*, 4(1), p.5.
- Dhami, P., Bruce, A., Jim, J., Dillon, S., Hall, A., Cooper, J., Bonhoure, N., Chiang, K., Ellis, P., Langford, C., Andrews, R. and Vetrie, D. (2010). Genomic Approaches Uncover Increasing Complexities in the Regulatory Landscape at the Human SCL (TAL1) Locus. *PLoS ONE*, 5(2), p.e9059.
- Eissenberg, J. and Elgin, S. (2014). HP1a: a structural chromosomal protein regulating transcription. *Trends in Genetics*, 30(3), pp.103-110.
- Eissenberg, J. and Reuter, G. (2009). Chapter 1 Cellular Mechanism for Targeting Heterochromatin Formation in *Drosophila*. *International Review of Cell and Molecular Biology*, pp.1-47.
- ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), pp.57-74.
- Enríquez, P. (2016). CRISPR-Mediated Epigenome Editing. *The Yale Journal of Biology and Medicine*, 89(4), pp.471–486.
- Esvelt, K. and Wang, H. (2014). Genome-scale engineering for systems and synthetic biology. *Molecular Systems Biology*, 9(1), pp.641-641.
- Forneris, F., Binda, C., Vanoni, M., Mattevi, A. and Battaglioli, E. (2005). Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Letters*, 579(10), pp.2203-2207.
- Fu, Y., Foden, J., Khayter, C., Maeder, M., Reyon, D., Joung, J. and Sander, J. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, 31(9), pp.822-826.
- Gaj, T., Charles A. Gersbach, and Carlos F. Barbas. (2013) ZFN, TALEN and CRISPR/Cas-Based Methods for Genome Engineering. *Trends in biotechnology*, 31(7), pp.397–405.
- Gao, X., Tsang, J. C. H., Gaba, F., Wu, D., Lu, L., & Liu, P. (2014). Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Research*, 42(20), pp. e155.
- Garcia-Bassets, I., Kwon, Y., Telese, F., Prefontaine, G., Hutt, K., Cheng, C., Ju, B., Ohgi, K., Wang, J., Escoubet-Lozach, L., Rose, D., Glass, C., Fu, X. and Rosenfeld, M. (2007). Histone Methylation-Dependent Mechanisms Impose Ligand Dependency for Gene Activation by Nuclear Receptors. *Cell*, 128(3), pp.505-518.
- Garneau, J., Dupuis, M., Villion, M., Romero, D., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A. and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 468(7320), pp.67-71.
- Gilbert, L., Horlbeck, M., Adamson, B., Villalta, J., Chen, Y., Whitehead, E., Guimaraes, C., Panning, B., Ploegh, H., Bassik, M., Qi, L., Kampmann, M. and Weissman, J. (2014a). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, 159(3), pp.647-661.
- Gilbert, L., Larson, M., Morsut, L., Liu, Z., Brar, G., Torres, S., Stern-Ginossar, N., Brandman, O., Whitehead, E., Doudna, J., Lim, W., Weissman, J. and Qi, L.

- (2014b). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, 154(2), pp.442-451.
- Gillette, T. and Hill, J. (2015). Readers, Writers, and Erasers. *Circulation Research*, 116(7), pp.1245-1253.
- Goldberg, A., Allis, C. and Bernstein, E. (2007). Epigenetics: A Landscape Takes Shape. *Cell*, 128(4), pp.635-638.
- Gottesman, M. (2002a). Mechanisms of Cancer Drug Resistance. *Annual Review of Medicine*, 53(1), pp.615-627.
- Gottesman, M., Fojo, T. and Bates, S. (2002b). Multidrug resistance in cancer: role of ATP-dependent transporters. *Nature Reviews Cancer*, 2(1), pp.48-58.
- Groner, A., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, N., Bucher, P. and Trono, D. (2010). KRAB-Zinc Finger Proteins and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading. *PLoS Genetics*, 6(3), p.e1000869.
- Haberland, M., Montgomery, R. and Olson, E. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nature Reviews Genetics*, 10(1), pp.32-42.
- Hahn, M., Dambacher, S., Dulev, S., Kuznetsova, A., Eck, S., Worz, S., Sadic, D., Schulte, M., Mallm, J., Maiser, A., Debs, P., von Melchner, H., Leonhardt, H., Schermelleh, L., Rohr, K., Rippe, K., Storchova, Z. and Schotta, G. (2013). Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes & Development*, 27(8), pp.859-872.
- Hall, M., Curtis, D., Metcalf, D., Elefanti, A., Sourris, K., Robb, L., Gothert, J., Jane, S. and Begley, C. (2003). The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12. *Proceedings of the National Academy of Sciences*, 100(3), pp.992-997.
- Hamilton, G. and Rath, B. (2015). Smoking, inflammation and small cell lung cancer: recent developments. *Wiener Medizinische Wochenschrift*, 165(19-20), pp.379-386.
- Hanna-Rose, W. and Hansen, U. (1996). Active repression mechanisms of eukaryotic transcription repressors. *Trends in Genetics*, 12(6), pp.229-234.
- Hilton, I., D'Ippolito, A., Vockley, C., Thakore, P., Crawford, G., Reddy, T. and Gersbach, C. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology*, 33(5), pp.510-517.
- Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157(6), pp.1262-1278.
- Hsu, P., Scott, D., Weinstein, J., Ran, F., Konermann, S., Agarwala, V., Li, Y., Fine, E., Wu, X., Shalem, O., Cradick, T., Marraffini, L., Bao, G. and Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, 31(9), pp.827-832.
- Huang, J., Sengupta, R., Espejo, A., Lee, M., Dorsey, J., Richter, M., Opravil, S., Shiekhata, R., Bedford, M., Jenuwein, T. and Berger, S. (2007). p53 is regulated by the lysine demethylase LSD1. *Nature*, 449(7158), pp.105-108.
- Jiang, W., Bikard, D., Cox, D., Zhang, F. and Marraffini, L. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology*, 31(3), pp.233-239.
- Johnstone, R. (2002). Histone-deacetylase inhibitors: novel drugs for the treatment of cancer. *Nature Reviews Drug Discovery*, 1(4), pp.287-299.
- Kahnert, K., Kauffmann-Guerrero, D. and Huber, R. (2016). SCLC—State of the Art and What Does the Future Have in Store?. *Clinical Lung Cancer*, 17(5), pp.325-333.
- Kearns, N., Pham, H., Tabak, B., Genga, R., Silverstein, N., Garber, M. and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nature Methods*, 12(5), pp.401-403.
- Kerbel, R., Kobayashi, H. and Graham, C. (1994). Intrinsic or acquired drug resistance and metastasis: Are they linked phenotypes?. *Journal of Cellular Biochemistry*, 56(1), pp.37-47.

- Konermann, S., Brigham, M., Trevino, A., Joung, J., Abudayyeh, O., Barcena, C., Hsu, P., Habib, N., Gootenberg, J., Nishimasu, H., Nureki, O. and Zhang, F. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 517(7536), pp.583-588.
- Krichevsky, A., Gutgarts, H., Kozlovsky, S., Tzfira, T., Sutton, A., Sternglanz, R., Mandel, G. and Citovsky, V. (2007). C2H2 zinc finger-SET histone methyltransferase is a plant-specific chromatin modifier. *Developmental Biology*, 303(1), pp.259-269.
- Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., Carte, J., Chen, W., Roark, N., Ranganathan, S., Ravinder, N. and Chesnut, J. (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *Journal of Biotechnology*, 208, pp.44-53.
- Liu, H., Wei, Z., Dominguez, A., Li, Y., Wang, X. and Qi, L. (2015). CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation: Fig. 1. *Bioinformatics*, 31(22), pp.3676-3678.
- Lo, A. and Qi, L. (2017). Genetic and epigenetic control of gene expression by CRISPR–Cas systems. *F1000Research*, 6, p.747.
- Lukacs, G., Haggie, P., Seksek, O., Lechardeur, D., Freedman, N. and Verkman, A. (2000). Size-dependent DNA Mobility in Cytoplasm and Nucleus. *Journal of Biological Chemistry*, 275(3), pp.1625-1629.
- Maison, C. and Almouzni, G. (2004). HP1 and the dynamics of heterochromatin maintenance. *Nature Reviews Molecular Cell Biology*, 5(4), pp.296-305.
- Mali, P., Aach, J., Stranges, P., Esvelt, K., Moosburner, M., Kosuri, S., Yang, L. and Church, G. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*, 31(9), pp.833-838.
- Mali, P., Yang, L., Esvelt, K., Aach, J., Guell, M., DiCarlo, J., Norville, J. and Church, G. (2013b). RNA-Guided Human Genome Engineering via Cas9. *Science*, 339(6121), pp.823-826.
- Marraffini, L. and Sontheimer, E. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics*, 11(3), pp.181-190.
- Matsuda, H., Putzel, G., Backman, V. and Szleifer, I. (2014). Macromolecular Crowding as a Regulator of Gene Transcription. *Biophysical Journal*, 106(8), pp.1801-1810.
- McDowell, J. and Dean, A. (1999). Structural and Functional Cross-Talk between a Distant Enhancer and the ϵ -Globin Gene Promoter Shows Interdependence of the Two Elements in Chromatin. *Molecular and Cellular Biology*, 19(11), pp.7600-7609.
- McLenachan, S., Sarsero, J. and Ioannou, P. (2007). Flow-cytometric analysis of mouse embryonic stem cell lipofection using small and large DNA constructs. *Genomics*, 89(6), pp.708-720.
- Melissa Conrad, C. (2018). Lung Cancer Symptoms, Signs, Stages, Treatment & Survival Rate. [online] eMedicineHealth. Available at: https://www.emedicinehealth.com/lung_cancer/article_em.htm.
- Mendenhall, E., Williamson, K., Reyon, D., Zou, J., Ram, O., Joung, J. and Bernstein, B. (2013). Locus-specific editing of histone modifications at endogenous enhancers. *Nature Biotechnology*, 31(12), pp.1133-1136.
- Mihaylova, M. and Shaw, R. (2013). Metabolic reprogramming by class I and II histone deacetylases. *Trends in Endocrinology & Metabolism*, 24(1), pp.48-57.
- Minucci, S. and Pelicci, P. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nature Reviews Cancer*, 6(1), pp.38-51.
- Mojica, F., Díez-Villaseñor, C., García-Martínez, J. and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 155(3), pp.733-740.
- myVMC (2015). *Small cell lung cancer (carcinoma of the lung) information | myVMC*. [online] myVMC. Available at: <https://www.myvmc.com/diseases/small-cell-lung-cancer-carcinoma-of-the-lung/>.

- Pabo, C., Peisach, E. and Grant, R. (2001). Design and Selection of Novel Cys2His2 Zinc Finger Proteins. *Annual Review of Biochemistry*, 70(1), pp.313-340.
- Pott, S. and Lieb, J. (2014). What are super-enhancers?. *Nature Genetics*, 47(1), pp.8-12.
- Qi, L., Larson, M., Gilbert, L., Doudna, J., Weissman, J., Arkin, A. and Lim, W. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, 152(5), pp.1173-1183.
- Ran, F., Hsu, P., Lin, C., Gootenberg, J., Konermann, S., Trevino, A., Scott, D., Inoue, A., Matoba, S., Zhang, Y. and Zhang, F. (2013). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell*, 155(2), pp.479-480.
- Ruijter, A., Gennip, A., Caron, H., Kemp, S. and Kuilenburg, A. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochemical Journal*, 370(3), pp.737-749.
- Schultz, D., Ayyanathan, K., Negorev, B., Maul, G. and Rauscher, F. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & Development*, 16(8), pp.919-932.
- Shen, B., Zhang, J., Wu, H., Wang, J., Ma, K., Li, Z., Zhang, X., Zhang, P. and Huang, X. (2013). Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Research*, 23(5), pp.720-723.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J., Cole, P., Casero, R. and Shi, Y. (2004). Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell*, 119(7), pp.941-953.
- Shi, Y., Sawada, J., Sui, G., Affar, E., Whetstine, J., Lan, F., Ogawa, H., Po-Shan Luke, M., Nakatani, Y. and Shi, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature*, 422(6933), pp.735-738.
- Siegel, R., Miller, K. and Jemal, A. (2017). Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians*, 67(1), pp.7-30.
- Snowden, A., Gregory, P., Case, C. and Pabo, C. (2002). Gene-Specific Targeting of H3K9 Methylation Is Sufficient for Initiating Repression In Vivo. *Current Biology*, 12(24), pp.2159-2166.
- Sripathy, S., Stevens, J. and Schultz, D. (2006). The KAP1 Corepressor Functions To Coordinate the Assembly of De Novo HP1-Demarcated Microenvironments of Heterochromatin Required for KRAB Zinc Finger Protein-Mediated Transcriptional Repression. *Molecular and Cellular Biology*, 26(22), pp.8623-8638.
- Sternberg, S., Redding, S., Jinek, M., Greene, E. and Doudna, J. (2014). DNA Interrogation by the CRISPR RNA-Guided Endonuclease Cas9. *Biophysical Journal*, 106(2), p.695a.
- Stojic, L., Lun, A., Mangei, J., Mascalchi, P., Quarantotti, V., Barr, A., Bakal, C., Marioni, J., Gergely, F. and Odom, D. (2018). Specificity of RNAi, LNA and CRISPRi as loss-of-function methods in transcriptional analysis. *Nucleic Acids Research*.
- Taplick, J., Kurtev, V., Kroboth, K., Posch, M., Lechner, T. and Seiser, C. (2001). Homooligomerisation and nuclear localisation of mouse histone deacetylase 111 Edited by J. Karn. *Journal of Molecular Biology*, 308(1), pp.27-38.
- Thakore, P., D'Ippolito, A., Song, L., Safi, A., Shivakumar, N., Kabadi, A., Reddy, T., Crawford, G. and Gersbach, C. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods*, 12(12), pp.1143-1149.
- Torres, I. and Fujimori, D. (2015). Functional coupling between writers, erasers and readers of histone and DNA methylation. *Current Opinion in Structural Biology*, 35, pp.68-75.
- Tuan, D., Kong, S. and Hu, K. (1992). Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proceedings of the National Academy of Sciences*, 89(23), pp.11219-11223.
- Vakoc, C., Letting, D., Gheldof, N., Sawado, T., Bender, M., Groudine, M., Weiss, M., Dekker, J. and Blobel, G. (2005). Proximity among Distant Regulatory Elements at

- the β -Globin Locus Requires GATA-1 and FOG-1. *Molecular Cell*, 17(3), pp.453-462.
- Villemejeane, J., & Mir, L. M. (2009). Physical methods of nucleic acid transfer: general concepts and applications. *British Journal of Pharmacology*, 157(2), pp.207–219.
- Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F., & Jaenisch, R. (2013). One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell*, 153(4), pp.910–918.
- Wang, J., Hevi, S., Kurash, J., Lei, H., Gay, F., Bajko, J., Su, H., Sun, W., Chang, H., Xu, G., Gaudet, F., Li, E. and Chen, T. (2009). The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature Genetics*, 41(1), pp.125-129.
- Wang, S., Tang, J., Sun, T., Zheng, X., Li, J., Sun, H., Zhou, X., Zhou, C., Zhang, H., Cheng, Z., Ma, H. and Sun, H. (2017). Survival changes in patients with small cell lung cancer and disparities between different sexes, socioeconomic statuses and ages. *Scientific Reports*, 7(1).
- Wiedenheft, B., Sternberg, S. and Doudna, J. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, 482(7385), pp.331-338.
- Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D. and Bonventre, J. (1994). The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proceedings of the National Academy of Sciences*, 91(10), pp.4514-4518.
- Workman, J. and Kingston, R. (1998). Alteration of Nucleosome Structure as a Mechanism of Transcriptional Regulation. *Annual Review of Biochemistry*, 67(1), pp.545-579.
- Wu, X., Scott, D., Kriz, A., Chiu, A., Hsu, P., Dadon, D., Cheng, A., Trevino, A., Konermann, S., Chen, S., Jaenisch, R., Zhang, F. and Sharp, P. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnology*, 32(7), pp.670-676.
- Xu, T., Li, Y., Van Nostrand, J., He, Z. and Zhou, J. (2014). Cas9-Based Tools for Targeted Genome Editing and Transcriptional Control. *Applied and Environmental Microbiology*, 80(5), pp.1544-1552.
- Young, I., Verbeek, P. and Mayall, B. (1986). Characterization of chromatin distribution in cell nuclei. *Cytometry*, 7(5), pp.467-474.
- Yu, X., Liang, X., Xie, H., Kumar, S., Ravinder, N., Potter, J., de Mollerat du Jeu, X. and Chesnut, J. (2016). Improved delivery of Cas9 protein/gRNA complexes using lipofectamine CRISPRMAX. *Biotechnology Letters*, 38(6), pp.919-929.
- Yun, M., Wu, J., Workman, J. and Li, B. (2011). Readers of histone modifications. *Cell Research*, 21(4), pp.564-578.
- Yun, W., Kim, Y., Kang, Y., Lee, J., Dean, A. and Kim, A. (2014). The hematopoietic regulator TAL1 is required for chromatin looping between the β -globin LCR and human γ -globin genes to activate transcription. *Nucleic Acids Research*, 42(7), pp.4283-4293.
- Zhou, Y., Kurukuti, S., Saffrey, P., Vukovic, M., Michie, A., Strogantsev, R., West, A. and Vetrie, D. (2013). Chromatin looping defines expression of TAL1, its flanking genes, and regulation in T-ALL. *Blood*, 122(26), pp.4199-4209.