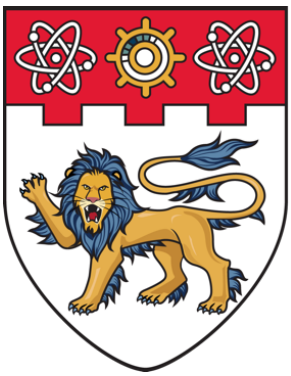

NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

Investigating Super-Enhancer-associated Chromatin Interactions at the *HOXA9-MEIS1* Oncogene Signalling Axis in Myeloid Leukaemia

WANG ZHENG JIE BENNY

SCHOOL OF BIOLOGICAL SCIENCES

2020

Investigating Super-Enhancer-associated Chromatin Interactions at the *HOXA9-MEIS1* Oncogene Signalling Axis in Myeloid Leukaemia

WANG ZHENG JIE BENNY

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2020

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

6th August 2020

. . . . . .

Date

. . . . . .

WANG ZHENG JIE BENNY

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

6th August 2020

. . . . . .

Date

. . . . . .

Dr Melissa Jane Fullwood

Authorship Attribution Statement

This thesis contains material from two papers published/in revision in the following peer-reviewed journal(s) *Trends in Genetics* & *Blood Cancer Discovery* in which I am listed as an author.

Chapter 1 is published as:

See, Y. X., Wang, B. Z. & Fullwood, M. J. 2019. Chromatin interactions and regulatory elements in cancer: from bench to bedside. *Trends in Genetics,* 35**,** 145-158. (Review Paper, Published).
doi: https://doi.org/10.1016/j.tig.2018.11.007

The contributions of the co-authors are as follows:

- The review paper was prepared and written as a co-first author with Mr See Yi Xiang. The diagrams and work within were completed in equal levels of contribution.
- Dr Melissa Fullwood vetted through and edited the manuscript.

Chapter 2 to 5 are submitted and is in revision as:

Wang, B., Kong, L., Babu, D., Choudhary, R., Fam, W., Tng, J. Q., Goh, Y., Liu, X., Song, F. F, Chia, P, Ming Chun Chan, Omer An, Cheng Yong Tham, Touati Benoukraf, Henry Yang, Wilson Wang, Wee Joo Chng, Daniel Tenen, Melissa Jane Fullwood. 2020. Three-dimensional Genome Organization Maps in Normal Haematopoietic Stem Cells and Acute Myeloid Leukemia. *bioRxiv*. (In revision at Blood Cancer Discovery).
doi**:** https://doi.org/10.1101/2020.04.18.047738

The contributions of the co-authors are as follows:

- All the following procedures: circularised chromosome conformation capture (4C), quantitative polymerase-chain-reaction, droplet digital polyermerase-chain-reaction. RNA extraction, reverse-transcription process, growth assays, chromatin immunoprecipitation-quantitative polymerase-chain-reaction.CRISPR design and experimental excision and genotyping were performed by me.
- ChIP-seq analyses were performed by me and super-enhancer calling from clinical samples and cell lines were performed by me with part of it contributed by Ms Ruchi Choudhary.
- I prepared the manuscript with Ms Kong Lingshi and Dr Deepak Babu.

- Hi-C and RNA-seq datasets that are dervied from the clinical CD34+ AML and CD34+ Knee samples were analysed by Ms Kong Lingshi.

- Hi-C and RNA-seq libraries were prepared by Dr Deepak Babu, Ms Goh Yufen, Fam Wee Nih and Ms Tng Jiaqi.

- Clinical acute myeloid leukaemia and healthy knee derived bone marrow samples were proceesed by the Cancer Science Institute Epigenomics Core (Ms Tng Jiaqi, Ms Fam Wee Nih & Ms Goh Yufen).

- The 4C and RNA-seq analysis were performed using the Cancer Science Institute Web Portal pipelines that were modified by Dr Omer An and Professor Henry Yang.

- Clinical acute myeloid leukaemia and healthy knee derived bone marrow samples were obtained through Professor Chng Wee Joo and Professor Wilson Wang from the National Univeristy Hospital. The patient consent for these clinical samples were approved by Dr Ming Chun Chan and coordinated by Ms Xin Liu, Ms Fang Fang Song and Ms Priscella Chia.

- Dr Melissa Fullwood provided guidance towards the project direction and edited the manuscript.

6th August 2020

. . . . . .

Date

……..

WANG ZHENG JIE BENNY

Table of Contents

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my PhD supervisor Dr Melissa Fullwood for the opportunity to work on this project and for her guidance and support through my PhD journey. I would also like to thank Ms Kong Lingshi for her work on the analyses of the Hi-C and RNA-seq data from the clinical knee and acute myeloid leukaemia samples as well as Dr Deepak Babu, Ms Tng Jiaqi and Ms Fam Wee Nih for their work in preparing the Hi-C libraries and clinical acute myeloid leukaemia and knee samples respectively. Furthermore, I like to specifically thank Mr See Yi Xiang and Ms Zhang Ying for their guidance and insightful discussions on bioinformatics and wet lab experimental procedures. I also thank the rest of the Fullwood lab members for their friendship and support through my PhD journey.

I would also like to thank the School of Biological Sciences, Nanyang Technological (NTU) for the scholarship that supported me through my PhD and the Cancer Science Institute of Singapore (CSI) at the National University of Singapore that I am affiliated with, for supporting and providing me with the support and resources to complete my PhD.

I like to take this opportunity to thank the National Research Foundation (NRF), NTU start-up funds and the RNA Biology Centre at the Cancer Science Institute of Singapore (Funded by the Singapore Ministry of Education Academic Research Fund Tier 3) for the support throughout my PhD work.

Finally, I like to express my gratitude to my family members and partner Ms Ong Hui Min for their unwavering and constant support and encouragement, without which this thesis will be impossible to complete.

Last but not least I like to express my deep appreciation to my mentor in life Dr Daisaku Ikeda and the members of the Singapore Soka Association for their constant encouragement and support through these years.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

3C	Chromatin Conformation Capture
4C	Circularised Chromosome Conformation Capture
5C	Carbon Copy Chromosome Conformation Capture
ALL	Acute Lymphoblastic Leukaemia
AML	Acute Myeloid Leukaemia
ATG16L2	Autophagy Related 16 Like 2
BASP1	Brain acid soluble protein
BAMBI	BMP and Activin Membrane Bound Inhibitor
BCR	Breakpoint Cluster Region
BET	Bromodomain and Extraterminal Domain
BIM	BCL2-interacting mediator
CDK7	Cyclin-dependent Kinase 7
CDKN2B	Cyclin-dependent Kinase Inhibitor 2B
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag
ChIP-seq	Chromatin Immunoprecipitation-Sequencing
CML	Chronic Myeloid Leukaemia
CTCF	CCCTC-Binding Factor
DDPCR	Droplet Digital Polymerase Chain Reaction
DMNT	DNA Methyltransferase
DOT1L	Disruptor Of Telomeric Silencing 1
eRNA	Enhancer RNA
EV	Empty Vector
EZH2	Enhancer of Zeste Homolog 2
FAB	French-American-British

FIRE	Frequently Interacting Region
FLT3	FMS-like tyrosine kinase 3
GEO	Gene Expression Omnibus
GSEA	Gene Set Enrichment Analysis
H3K27ac	Histone 3 Lysine 27 Acetylation
H3K27me3	Histone 3 Lysine 27 Tri-Methylation
H3K4me3	Histone 3 Lysine 4 Tri-Methylation
HAT	Histone Acetyl Transferase
HDAC	Histone Deacetylase
HOTTIP	HOXA Transcript At The Distal Tip
HOXA9	Homeobox A9
IDH	Isocitrate Dehydrogenase
lncRNA	Long Non-coding RNA
KRAB	Kruppel Associated Box
MEIS1	Myeloid Ectropic Viral Integration Site 1
miRNA	Micro RNA
MLL	Mixed Lineage Leuekamia
MNC	Mononuclear cells
mRNA	Messenger RNA
NFKB	Nuclear Factor Kappa-Light-Chain-Enhancer of Activated B Cell
NGS	Next Generation Sequencing
NPM1	Nucleophosmin 1
NPM1c	Nucleophosmin 1 Cytoplasmic
NUP98-HOXA9	Nucleoporin 98-Homeobox A9
OSCP1	Organic Solute Carrier Partner 1

PBS	Phosphate-Buffered-Saline
PBX	Pre-B Cell Leukaemia Transcription Factor
PGRA	Progesterone Receptor A
PGRB	Progesterone Receptor B
PML-RARa	Promyelocytic Leukaemia-Retinoic Acid Receptor Alpha
PRC2	Polycomb Repressive Complex 2
ROSE	Ranking of Super-enhancers
RT-qPCR	Reverse Transcriptase-Quantitative Polymerase Chain Reaction
RUNX1	Runt-Related Transcription Factor 1
SE	Super-enhancer
SNP	Single Nucleotide Polymorphism
SRA	Sequence Read Archive
TAD	Topologically Associating Domain
TALE	Three-Amino-Acid-Loop
TET	Ten-Eleven Translocation Methylcytosine Dioxygenase
TFAP2E	Transcription Factor AP-2 Epsilon
TLE4	Transducing-like-enhancer protein 4
WHO	World Health Organisation
YY1	Yin Yang 1

SUMMARY

Myeloid leukaemia is a form of blood cancer that is characterised by a block in haematopoietic stem cell differentiation and uncontrolled cell proliferation. *HOXA9* and *MEIS1* are two commonly overexpressed oncogenes that are associated with poor prognoses in myeloid leukaemia patients. However, our understanding of how 3D genome organisations including Topologically Associating Domains (TADs) and chromatin interactions regulating *HOXA9* and *MEIS1* expression remains limited.

In this study, I suggested the presence of a type of sub-TAD or "Frequently Interacting Region" (FIRE) at the *MEIS1* gene locus and the presence of chromatin loops between the super-enhancers and the *MEIS1* promoter within the FIRE. The excision of the CTCF boundary at the 3' end of the *MEIS1* FIRE in K562 myeloid leukaemia cells resulted in the disruption of the FIRE as well as led to the loss of chromatin loops to the *MEIS1* promoter and the subsequent downregulation of *MEIS1* expression. The reorganisation of chromatin loops that are linked to the *MEIS1* promoter is also seen at the distal upstream and downstream regions to the *MEIS1* FIRE. Analysis of H3K27 acetylation Chromatin-Immunoprecipitation-sequencing (ChIP-seq) derived super-enhancer regions in mononuclear acute myeloid leukaemia (AML) and normal knee CD34+ haematopoietic stem cells revealed the presence of novel super-enhancers at the *HOXA9* locus that are specifically present in the AML patients, thus suggesting the acquisition of novel super-enhancers in AML.

Circularised chromosome conformation capture (4C) technique of myeloid leukaemia cell lines THP-1, HL-60, K562 and Hi-C sequencing of three CD34+ AML and three CD34+ normal knee samples, also showed the presence of chromatin loops between these super-enhancer regions and the *HOXA9* promoter. Interestingly, these chromatin loops were present at the regions that acquired super-enhancers in both the Hi-C heatmaps of the AML and normal CD34+ clinical samples.

These results suggest that the acquired super-enhancers in myeloid leukaemia could hijack the pre-existing chromatin loops with the *HOXA9* promoter, thereby promoting the overexpression of *HOXA9* in myeloid leukaemia. Finally, I also demonstrated that *HOXA9* and *MEIS1* proteins can bind to their own and each other's promoters in the AML cell line THP-1, suggesting that they can cross-regulate each other's transcription in myeloid leukaemia. HOXA9 proteins were also shown to be enriched at the super-enhancer near the *SNX10* gene (*SNX10* SE) that forms chromatin loops with the *HOXA9* promoter, thus indicating that HOXA9 proteins have a role in regulating its associated super-enhancers to promote its transcription.

Taken together, my work provides new insight into the mechanism of FIREs, chromatin interactions, super-enhancers and their relationship with key oncogenes regulation such as with *HOXA9* and *MEIS1*in myeloid leukaemia.

1 Introduction

1.1 Myeloid Leukaemia

Myeloid leukaemia is a type of blood cancer that is characterized by a block in cell differentiation and the abnormal proliferation of hematopoietic progenitor cells (Estey and Döhner, 2006). Myeloid leukaemia is categorized into two groups, namely the fast progressing acute myeloid leukaemia (AML) and the slower developing chronic myeloid leukaemia (CML)(National Cancer Institute).

1.2 Acute Myeloid Leukaemia

AML occurs predominantly in the elderly population, with poorer prognosis in patients who are above 60 years of age (National Cancer Institute). AML diagnosis in patients typically presents blood count results with more than 20% of the blast consisting of myeloblast, monoblasts and megakaryoblasts (Döhner et al., 2016). Symptoms of AML typically manifest as fever, asthenia, anaemia and pallor (De Lima et al., 2016).

Statistics from the National Cancer Institute of the United States reported an estimated 19 940 new cases of AML with 11 180 estimated deaths in 2020 (National Cancer Institute). On a global scale, total AML incidence was reported to be 351 965 cases in 2012 with a higher onset rate in males than in females (Acute myelogenous leukemia and acute promyelocytic leukemia - In Union for International Cancer Control, 2014). Although complete remissions are generally achieved in AML patients, the risk of relapse was observed to be high and eventually fatal, with only 28.7% of patients surviving past 5 years (National Cancer Institute).

Traditionally, AML diagnosis involves the grading of disease phenotype through the French-American-British (FAB) classification (**Table 1**)(Acute myelogenous leukemia and acute promyelocytic leukemia - In Union for International Cancer Control, 2014). The FAB classification divides AML into eight different subtypes; M0 to M7, according to the levels of cell differentiation. The World Health Organisation (WHO) later added new criteria for the classification of genetic abnormalities observed in AML patients (**Table 2**) (Acute myelogenous

leukemia and acute promyelocytic leukemia - In Union for International Cancer Control, 2014). The revised additions on genetic abnormalities indicate different prognoses according to the types of genetic abnormalities found. For instance, the translocation occurring between chromosomes 8 and 21 are deemed as a favourable prognosis, while translocations between chromosomes 6 and 9 are considered as unfavourable prognosis.

Table 1. The French-American-British (FAB) classification. From '2014 Review of Cancer Medicines on the WHO List of Essential Medicines' Union for International Cancer Control. Adapted from (Acute myelogenous leukemia and acute promyelocytic leukemia - In Union for International Cancer Control, 2014)

FAB subtype	Description
M0	Undifferentiated AML
M1	AML without maturation/cell differentiation
M2	AML with maturation/cell differentiation
M3	Acute promyelocytic leukaemia
M4	Acute myelomonocytic leukaemia
M5	Acute monocytic leukaemia
M6	Acute erythroblastic leukaemia
M7	Acute megakaryoblastic (platelets precursor) leukaemia

Table 2. AML patient prognosis based on a revised list of genetic abnormalities. From '2014 Review of Cancer Medicines on the WHO List of Essential Medicines' Union for International Cancer Control. Adapted from (Acute myelogenous leukemia and acute promyelocytic leukemia - In Union for International Cancer Control, 2014)

Favourable prognostic abnormalities	Intermediate prognostic abnormalities	Unfavourable prognostic abnormalities
t (8;21) (AML M2)	Normal karyotype	Deletion/loss of chromosome 5 or 7
Inversion of chromosome 16 (16;16)		Translocation or inversion of chromosome 3
t (15;17) (APML M3)		t (6;9)
		t (9:22) Transformed CML or *de novo* AML or ALL

1.3 Chronic Myeloid Leukaemia

A large fraction of chronic myeloid leukaemia (CML) develops due to the t9;22 chromosomal translocation which is also known as the Philadelphia Chromosome (Rowley, 1973). This translocation occurs in hematopoietic stem cell and results in the formation of a fusion *BCR-ABL1* oncogene with increased aberrant fusion tyrosine kinase activity which leads to CML development (Sabath, 2013). CML makes up approximately 15% of adult leukaemias and carries an incidence rate of 1.6 per 100 000 people (Seer.cancer.gov, 2017). Up to 90% of CML patients are diagnosed during the chronic phase of the disease but if left untreated progress to the advanced blast phase of the disease that carries poor prognosis which is similar to those of AML patients (Giralt et al., 1995). Similar to AML genetic profiles, CML also possess features of altered DNA methylation at genes such as *BCR* (Koh et al., 2010), *ATG16L2* (Dunwell et al., 2010), *EZH2* (Enhancer of zeste homolog 2) (Schmidt et al., 2014) and *P53* (Bodoor et al., 2014), which are suggested to contribute to disease progression.

1.4 Current treatment options for acute myeloid leukaemia

The main treatment option for AML are chemotherapy and stem cell transplantation (Ferrara and Schiffer, 2013). Chemotherapy are often applied to eliminate the cancerous cells before stem cell transplantation is performed to restore the bone marrows in AML patients (American Cancer Society, 2020). Standard chemotherapy drugs used in AML include the combination of cytarabine and anthracyclines such as Daunorubicin and Idarubicin (Mayer et al., 1994). The common routine of this combination of drugs involves treatment with cytarabine at the dosage of 100-200mg/m^2 and 60mg/m^2 of daunorubicin in the first three days of the treatment cycle (Dombret and Gardin, 2016). Induction of cytarabine at the same dosage then continues for four following days (Döhner et al., 2016, Dombret and Gardin, 2016). Each cycle of this treatment averages a total of seven days (7 days with cytarabine+ 3 days of daunorubicin induction plan) with these two classes of chemotherapy drugs.

Anthracyclines function as chemotherapeutic agents by intercalating DNA base pairs and inhibiting the activities of DNA topoisomerases I and II (Malhotra and Perry, 2003). Cytarabine works by inhibiting DNA synthesis in S phase cancer cells (Momparler, 2013, Martindale, 1993). However, chemotherapy treatments such as those involving drugs like anthracyclines have not seen significant improvements in long-term patient outcomes with fatal relapses and serious side effects often reported in AML patients.

For example, while these anthracyclines and cytarabine regimes can achieve complete remission in up to 70% of AML patients, less than 40% of these AML patients exhibit long-term survival beyond five years (Fathi and Karp, 2009, Rowe et al., 2005). Furthermore, intermediate forms of the anthracycline such as Doxorubicin can be reduced by endogenous enzymes such as nitric oxide synthase 3 to produce free radicals that are suspected to be associated with cardiotoxicity (Malhotra and Perry, 2003, Volkova and Russell, 2011). Such side effects could be detrimental towards elderly AML patients with the potential occurrence of heart failures and other health complications (Hershman et al., 2008, Druhan et al., 2015). It was reported that the combination of confounding risk factors in elderly patients has often prevented them from undergoing

intensive chemotherapy treatment, with up to two-thirds of elderly patients being unsuitable for such treatments (Deschler et al., 2006). As the risk often outweighs the benefit of chemotherapy in elderly AML patients, physicians often avoid applying such treatments for this group of patients, thus leaving them with a lack of alternative therapy options. Although a small subset of elderly AML patients such as those who are below the age of 80 and possess no unfavourable chromosomal translocations like the t(5;7) showed slight benefits from intensive chemotherapy, most of the elderly AML patients who are above the age of 70 showed poor overall survival rates of just 4.6 months (Kantarjian et al., 2010). Furthermore, the chemoresistance of cancer cells has been widely reported in patients (Galmarini et al., 2002). Galmarini et al. (2002) reported that AML relapses often occur due to drug resistance and that only 30-50% of relapse patients achieved complete remission, with overall patient survival rates being low (Galmarini et al., 2002, Ciurea et al., 2015). Therefore, there is an urgent need to identify new therapeutic targets in the treatment of AML.

Increasingly, epigenetics regulation of oncogenes in AML is considered as a potential area for therapy in the clinical setting. This is because a large portion of AML onset was found to be a results of dysregulated epigenetics modulation (Guillamot et al., 2016, Sun et al., 2018a). Epigenetics dysregulation in AML mainly occurs as a consequence of gene mutations and chromosomal translocations and that these abnormal genetic mutations and chromosomal translocations were observed in a number of epigenetic modulating proteins (**Figure 1 & Table 3)**.

Table 3. List of selected aberrant epigenetic regulators in AML patients.
Table is adapted from (Wouters and Delwel, 2016)

Gene	Epigenetic Function	Abnormalities & their percentage occurrence in AML
DNMT3A	De novo DNA Methylation	Frameshift/Missense/Non-sense mutations (6-36%)
TET2	Conversion of 5-methylcytosine to 5-hydroxymethylcytosine (De-methylation process)	Frameshift/Missense/Non-sense mutations (8-27%)
IDH1/2	Converts isocitrate to α-ketoglutarate, a co-factor for TET2	Missense Mutations (*IDH1*- 5-16%, *IDH2* – 6-19%)
KMT2A (MLL)	H3K4 methyltransferase	MLL-Fusion genes. (1-10%), partial tandem duplications (4-7%)
EZH2	H3K27 methyltransferase	Gene mutations (2%)

In a large AML genome study by *The Cancer Genome Atlas,* AML genomes were associated with fewer mutations as compared to the genomes of other cancers (Network, 2013). However, approximately 40% of AML mutations tend to develop at epigenetic regulating proteins such as DNMT3A (DNA Methyltransferase 3 Alpha), TET1/2 (Ten-eleven-translocation Methylcytosine Dioxygenase) and IDH1/2 (Isocitrate Dehydrogenase) (Guillamot et al., 2016, Sun et al., 2018a). *DNMT3A*, *IDH1/2*, *TET2* and *MLL*-fusion genes formed the

highest percentages of mutated epigenetic regulators in AML at 22%, 20%, 14% and 10% occurrence frequencies in AML respectively (Sun et al., 2018a).

A study that analysed a cohort of AML patients for their genome methylation patterns revealed that patients who possess higher degrees of aberrant methylation patterns within their genomes suffered from quicker relapses (Li et al., 2016). Furthermore, the frameshift mutation in the *DNMT3A* gene results in the aberrant activity of the DNMT3A protein that is responsible for DNA methylation at the genome (Ribeiro et al., 2012, Network, 2013). The DNMT3A protein was also reported to work with the fusion protein PML-RARa to induce hypermethylation at PML-RARa targeted genes in acute promyelocytic leukaemia (Di Croce et al., 2002). Subramanyam et al. (2010) demonstrated that both DNMT3A and PML-RARa presence are required for the manifestation of the hypermethylated phenotype in mice. In another example, Zhou et al. (2018) suggested that DNA methyltransferases work with AML1-ETO proteins that are expressed in AML groups with the t(8;21) chromosomal translocation, to methylate the promoter region of the cell cycle regulating gene *BASP1* (brain acid-soluble protein 1), which consequently result in the downregulation of *BASP1* expression and increased cell proliferation. The inhibition of DNMNT activity via the drug decitabine resulted in the upregulation of *BASP1* and arrests cellular proliferation. Collectively, these results highlighted the role that epigenetic modulators such as DNMTs plays in contributing to genetic dysregulation in myeloid leukaemia and highlights the potential for the development of myeloid leukaemia therapies that are targeted towards them.

In another example, the aberrant expression of chromatin-modifying enzymes such as those like DMNT3A, IDH1/2 (Wouters and Delwel, 2016) were also observed in cases of initial AML diagnoses as well as in AML relapse cases, suggesting that epigenetic dysregulation could be responsible for AML development at all phases of AML development (Walter and Appelbaum, 2012). Since, these enzymes drive methylation at important tumour suppressor genes, their dysregulation can therefore result in the eventual onset of cancer (Wouters and Delwel, 2016).

On the other hand, a chromosomal rearrangement involving the Mixed-Lineage-Leukaemia (*MLL*) gene leads to the production of *MLL*-fusion genes (Meyer et al., 2018). These MLL-fusion proteins are able to deposit abnormal tri-methylation at the H3K4 histone location that results in aberrant oncogenes expression (Dawson et al., 2011).

In addition to the role of enzymes such as DNMTs, oncogenic transcription factors and fusion proteins also play a part in influencing epigenetic modulation. For instance, the c-MYC protein is a transcription factor that influences AML progression via epigenetically associated modulation (Salvatori et al., 2011). As c-MYC protein is involved in the regulation of cell growth, its aberration expression can lead to cancer (Salvatori et al., 2011). In their study Salvatori et al. (2011), demonstrated that in AML, c-MYC protein promotes *EZH2* transcription by directly binding to its promoter leading to subsequent dysregulated gene silencing and cancer progression.

The methyltransferase subunit of the poly-comb repressive complex 2 (PRC2) enzyme EZH2 (Enhancer of zeste homolog 2) is also seen to be dysregulated in AML. For instance, the Ezh2 depleted MLL-AF9 transformed AML cells exhibited reduced cell proliferation and hindered cell cycle progression as compared to wildtype Ezh2 *MLL-AF9* transformed AML cells (Tanaka et al., 2012). Additionally, Ezh2 deletion also reactivates myeloid differentiation cells that are originally repressed under wildtype Ezh2 presence (Tanaka et al., 2012, Thiel et al., 2013). Crucially, the loss of Ezh2 function also resulted in the reduction of AML aggressiveness in the mouse model (Tanaka et al., 2012).

In order to target these epigenetic modulations, specific drugs such as Azacitidine and Decitabine, DOT1L (Disruptor of telomeric silencing 1-like) and EZH2 (Enhancer of zeste homolog 2) inhibitors, which targets the DNA methyltransferases (e.g. DMNT3A), DOT1L and EZH2 enzymes respectively are being developed (Wouters and Delwel, 2016). Each class of these drugs are applied in different groups of AML patients according to their AML subtypes and genetic mutations and chromosomal translocations (**Table 4**). Although these epigenetic targets compounds are in development as therapeutic options for the treatment of AML, the understanding of their effects on modulating

chromatin interactions and their associated AML oncogenes transcriptions remains very limited.

Table 4. Overview of selected class of epigenetic targeted compounds for the treatment of AML and their current phase of developments. Table is adapted from (Wouters and Delwel, 2016).

Class of epigenetic regulator	Target	Compound	Phase of Development
DNA Methyltransferase	DNMTs	▪ Azacitidine ▪ Decitabine	Approved
Methylation Regulators	IDH1/2	▪ IDH inhibitors	Clinical Trials Phases I to III
Histone Lysine methyltransferase	EZH2 MLL-Fusion complexes	▪ EZH2 Inhibitors ▪ DOT1L Inhibitors ▪ MLL-Menin Inhibitors	Preclinical & Clinical Trials Phases I & II
Histone acetyl reader	Bromodomain containing (BET) proteins	▪ BET Inhibitors	Clinical Trials Phase I

Recently, alternative treatments such as immunotherapy for AML has been gaining increased attention. In the review by Walter and Appelbaum (2012), they discussed the potential therapeutic impact of a CD33+ antibody targeting drug Gemtuzumab Ozogamicin (Mylotarg) that was recently approved by the United States Food & Drug Administration. During phase 3 clinical trial, the combination of conventional chemotherapy and Gentuzumab Ozagamicin induction was found to improve events free survival by approximately 23% with around 27% increase of relapse-free outcomes in AML patients within a two-year period (Castaigne et al., 2012). This drug was also reported to be effective when used for the treatment of acute promyelocytic leukaemia and is also usable

in combination with chemotherapy for AML patients with favourable prognoses (Walter and Appelbaum, 2012).

1.5 Current treatment strategies for chronic myeloid leukaemia

The current treatments for CML also involve the use of chemotherapy or tyrosine kinase inhibitors such as dasatinib and imatinib (Hochhaus et al., 2007, Jönsson et al., 2012). Patients who do not respond to the treatment with tyrosine kinase inhibitors often experience primary or secondary drug resistance. Treatment with imatinib shows a variable degree of success with approximately 50% of CML developing drug resistance (Ottmann et al., 2018). The reasons for these observed drug resistances are suggested to be due to point mutations at the *BCR-ABL1* fusion gene as well as the occurrence of other chromosomal translocations in CML patients (Roche-Lestienne et al., 2002, Roumiantsev et al., 2002).

CML resistance to tyrosine kinase inhibitors is measured via the expression levels of pro-apoptotic genes such as the *BCL2*-interacting mediator (*BIM*). The BIM protein functions by antagonising the other anti-apoptotic proteins within the BCL-2 family (Kuribara et al., 2004). BIM expression levels were established to be lowered in CML cells as compared to normal non-cancer cells (Aichberger et al., 2005) and that the reduction in *BIM* expression in tyrosine kinase inhibitor-treated CML patients are associated with the onset of drug resistance (San José-Eneriz et al., 2009, Kuroda et al., 2006). The study by San José-Eneriz et al. (2009) in 100 CML patients showed that BIM reduction was associated with the DNA hypermethylation and that 36% of the CML patients who possessed higher DNA methylation had lower *BIM* expressions which are associated with decreased imatinib sensitivity. In another study, imatinib treatment in 120 CML patients showed mixed outcomes with approximately 41% of these patients experiencing resistance to imatinib (Jelinek et al., 2011). Intriguingly, the authors observed a significant increase in methylated genes in the imatinib-resistant group as compared to those who were responsive to the treatment (Jelinek et al., 2011). They also revealed that five out of ten genes (*CDKN2B, OSCP1, PGRA, PGRB, TFAP2E*) that they analysed for DNA methylation showed a positive correlation of CpG methylation as the disease progressed in severity from the chronic phase to the accelerated phase and blast phase (Jelinek et al., 2011).

In addition the dysregulation of epigenetic modulators such as the histone acetyltransferase (HAT) and histone deacetylase (HDAC) in *BCR-ABL* type CML was suggested to affect the translocation of the transcription factor and tumour suppressive protein TP53 from the nucleus to the cytoplasm (Kusio-Kobialka et al., 2012). This inhibition of TP53 translocation to the cytoplasm leads to the block in apoptosis response during DNA damage, thereby contributing to CML pathogenesis.

The EZH2 protein was also demonstrated to play a contributing role towards CML development. The PRC2-EZH2 complex functions by depositing H3K27 trimethylation (H3K27me3) marks at a certain genomic region to induce repression of gene transcription (Margueron et al., 2008). Aberrant regulation of this complex results in hypermethylation and the promotion of leukemic stem cell survival as well as used as a biomarker for prostate and breast cancers (Xu et al., 2012, Kleer et al., 2003). Furthermore, mutations to *EZH2* was suggested to promote the conversion of di-methylation to tri-methylation in non-Hodgkin lymphoma (Wigle et al., 2011). Separate studies have observed that the loss of *EZH2* strongly reduced leukemic stem cell survival in both the CML mouse model and in chronic phase CML patients (Xie et al., 2016, Scott et al., 2016).

Altogether, these studies strongly imply that aberrant epigenetic regulation plays an important role in myeloid leukaemia development and progression. Thus, targeting aberrant epigenetic regulation may have therapeutic potential for myeloid leukaemia treatment.

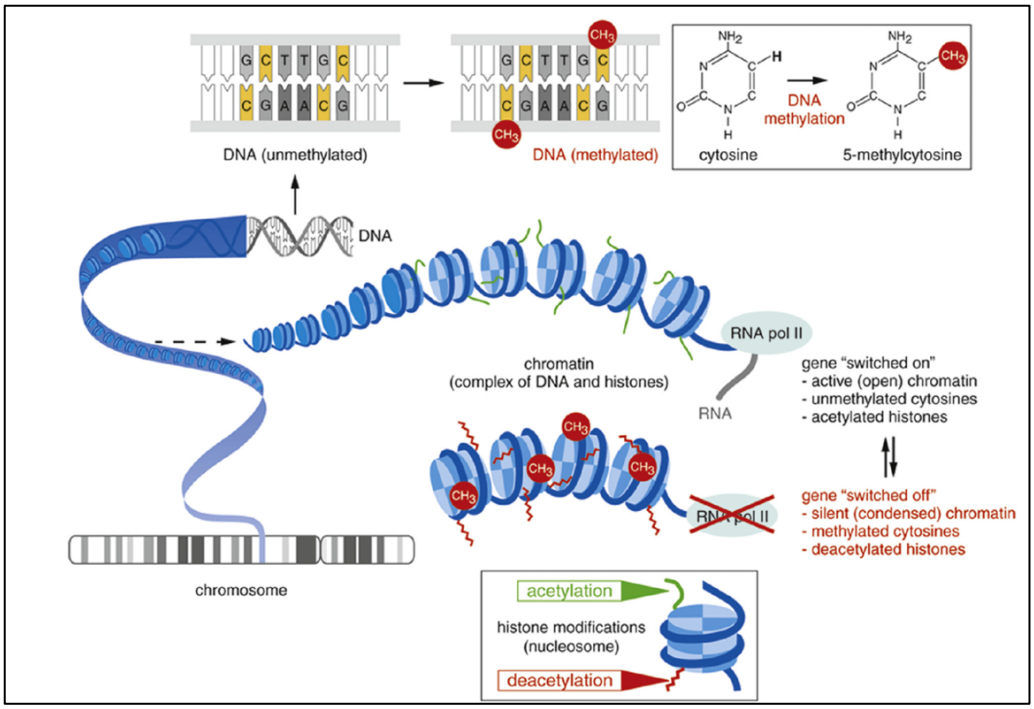

Figure 1. Schematic of epigenetic regulations on chromatin fibres and the effects of methylation and acetylation on chromatin and gene expression. The chromatin methylated and unmethylated states are indicated in the schematic above. Acetylation and deacetylation of the histones are depicted in the lower half of the figure. Figure adapted from (Mukherjee et al., 2015).

1.6 Three-Dimensional (3D) Genome Organization

Since epigenetic factors function by acting on chromatin and genomic DNA regions, it is important to understand their activity with respect to how these chromatin loops are organized and regulated. The three-dimensional (3D) genome of a cell is compartmentalized in a nonrandom and highly organized fashion within the nucleus. These organized genome folding is suggested to be relevant to various cellular mechanisms such as DNA replication and gene transcription (Szabo et al., 2019). For instance, active genes are found to be commonly located in the interior compartments of the nucleus, while genes that are repressed are seen to be residing in the periphery regions (Bickmore and van Steensel, 2013). The 3D genome is generally arranged into two large compartments; A & B. Compartment A represents part of the genome with active chromatin and gene transcription while compartment B represents the genome with repressed regions and silenced gene transcriptions (Lieberman-Aiden et al., 2009, Wang et al., 2016). Furthermore, active gene clusters have been reported

to be spatially clustered and associated together in the same compartments (Hakim et al., 2011, Lieberman-Aiden et al., 2009, Simonis et al., 2006).

1.7 Topologically Associating Domains (TAD)

As we delve through the hierarchy of the 3D genome from the macro scale of the chromosome territories to the micro-scale of the compartments and beyond, additional layers of regulation are observed (**Figure 2**). An example of such regulation is the Topologically Associating Domains (TAD). TADs are insulated genomic regions that are suggested to promote the interactions of genes and enhancers that are located within them and at the same time function to prevent external genes outside its domains from interacting with the region within it (Symmons et al., 2014, Eser et al., 2017).

TADs are widely proposed to be conserved and stable across different cell types, with only a limited number of TAD being altered by cell specificity and function (Dixon et al., 2015). Also, genes that are residing within the same TAD were observed to be co-regulated during the cell differentiation process (Zhan et al., 2017). TAD formation was reported to occur via the formation of a chromatin loop extrusion that is mediated with two convergent interacting CCCTC binding factor (CTCF) boundaries (Nuebler et al., 2018, Sanborn et al., 2015). In addition, the protein cohesin was also reported to work in tandem with CTCF in the formation of the TAD boundaries (Hanssen et al., 2017a). The co-immunoprecipitation of CTCF proteins also showed the presence of all the subunits of the cohesin complex (Hansen et al., 2018).

The disruption of these CTCF-cohesin TAD boundaries results in the loss of insulation and consequently allowed aberrant enhancer-promoter interaction formation across two originally different domains (Lupiáñez et al., 2015, Guo et al., 2015). In another study, the loss of the CTCF protein in mouse embryonic stem cells also caused dysregulation of chromatin loops in the region (Nora et al., 2017). This important relationship between CTCF and cohesin proteins in maintaining TAD boundaries was also demonstrated by the work of Rao et al. (2017) that showed that the loss of cohesin proteins eradicated loop domains and that the reintroduction of cohesin proteins rescues the loop domains recovery.

However, Hansen et al. (2018) also suggested that CTCF, cohesin and chromatin loops are dynamic in regulation, with CTCF and cohesin proteins seen to diffuse and bind at different genomic regions at different times. This observation is in line with other observations that CTCF proteins serve other roles besides being insulators of TADs. In other studies, CTCF proteins were also reported to bind at enhancer-promoter loops within a certain TAD, forming complexes with the mediator and cohesin proteins (Rao et al., 2014, Phillips-Cremins et al., 2013).

1.8 Frequently Interacting Regions (FIREs)

Recently, an alternative architecture feature that resembles a sub-TAD was revealed to be present in the 3D genome. Through their Hi-C libraries analyses, Schmitt et al. (2016a) reported the presence of extraordinarily dense regions of chromatin interactions in 21 types of human cells/tissues. By applying their novel bioinformatics algorithm across the genomes of 21 types of different human cells/tissues, they revealed the presence of hotspots of dense chromatin interactions, which was later termed as "Frequently Interacting Region" (FIRE). In addition to their dense chromatin interactions, some key characteristics of FIREs include being highly enriched for enhancers/super-enhancers which are located near genes that are cell/tissue type specific (e.g. hematopoietic stem cells) as well as being partially reliant on CTCF and cohesin complex for maintenance (Schmitt et al., 2016a). For example, comparison of a specific region between GM12878 and IMR90 cell lines showed that a FIRE is present with a specific-super enhancer that supports B cell gene regulation in GM12878, but this FIRE and super-enhancer was otherwise absent in the IMR90 cell line at the same region (Schmitt et al., 2016a). This demonstrates that this FIRE and its associated super-enhancers are regulated in a cell type-specific manner.

All in all, our knowledge of TADs, FIREs and their association with epigenetic modulating proteins like CTCF and cohesin provides us with a broader idea of how the 3D genome may be organized in different cell types and differentiation stages. However, it is also crucial to improve our understanding of how specific chromatin interactions within this TAD and FIRE compartments and epigenetic proteins function together to contribute to genome-wide transcriptome landscape. This is especially important in the circumstances of disease onset and progression.

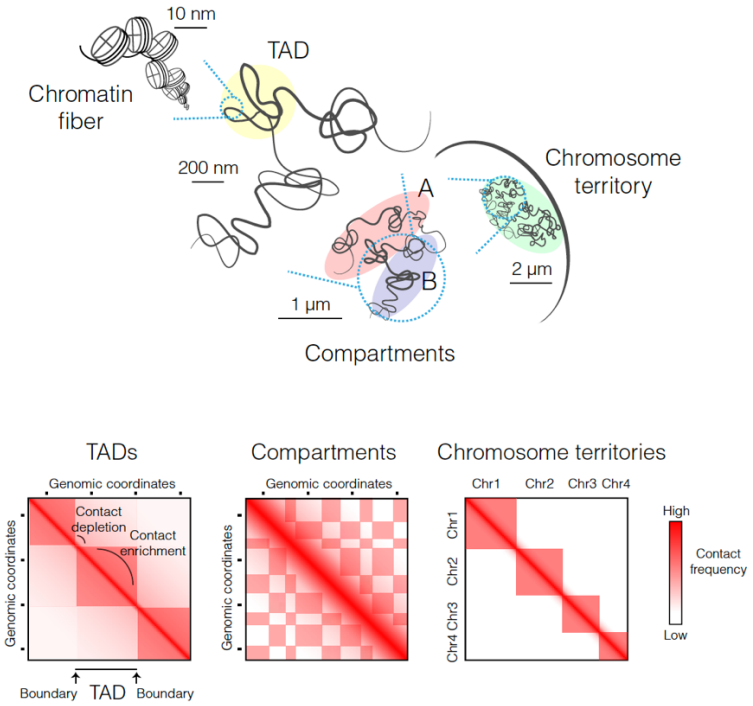

Figure 2. Overview of 3D genome organisation in a cell nucleus. Top image shows the unwrapping of chromosome packaging from the chromosome territories to the chromatin fibre. Lower schematic shows an example of Hi-C contact heatmaps of chromosome territories to the individual compartments within a chromosome and their associated topologically associated domains (TADs). Figure is adapted from (Szabo et al., 2019).

1.9 Chromatin Interactions

The interaction by two or more regions of the chromatin in *cis* or trans is referred to as a chromatin interaction (**Figure 3**). Chromatin interactions are often assisted by the binding of transcriptional proteins and epigenetic modulators. Chromatin interactions have been suggested to contribute towards the regulation of genes in cellular function and disease development (Nair and Kumar, 2012) and could occur as inter or intrachromosomal interactions (Zhao et al., 2006). Intra-chromosomal interactions occur when two or more separate genome loci come together in close spatial proximity (Babu and Fullwood, 2015, Peng et al., 2019). For example, an enhancer and promoter can interact, regulating gene transcription (Gorkin et al., 2014). Through such loops, regulatory proteins at the enhancer can then be associated with the promoter to initiate the transcription of a specific gene in the cell (Gorkin et al., 2014). Recently, Meddens et al. (2016) showed that 92 inflammatory bowel disease loci colocalise with non-coding

DNA regulatory elements interacts with candidates genes during disease onset. Martin et al. (2015) also identified long-range interactions between autoimmune diseases associated with loci and potential enhancers. Additionally, t hese genomic interactions were also observed to be T and B-cells specific (Martin et al., 2015).

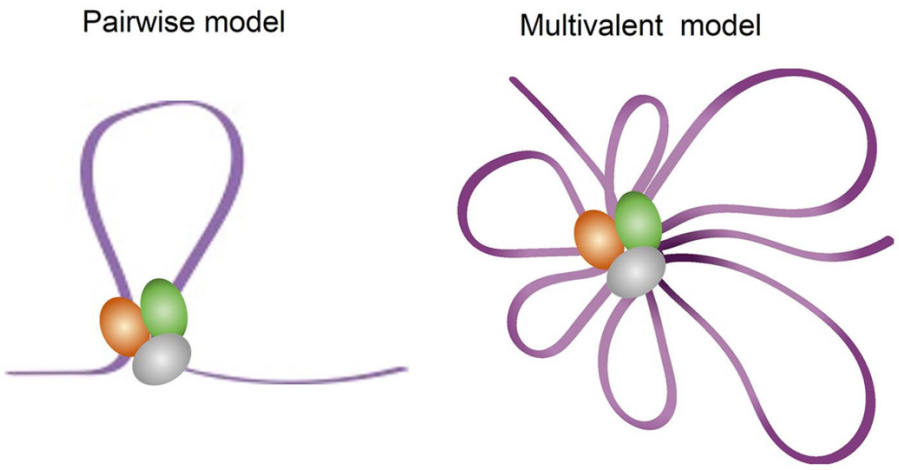

Figure 3. Schematic of chromatin interaction models. The left figure shows a pairwise one to one contact model. The right image shows a multiple loci chromatic contact model. The loops represent chromatin interactions. The colour ovals represent proteins such as CTCF and cohesins as well as transcriptional proteins such as HOXA9 & MEIS1. Schematic is adapted from (Wei et al., 2019).

Dysregulated chromatin interactions can influence cancer development through the silencing or overexpression of various important cellular genes. Several studies demonstrated the presence of aberrant chromatin interactions that contribute to specific cancers. Differences in intra-chromosomal interactions between cancer and non-cancerous cell lines suggest that chromatin interactions are largely dysregulated in breast cancer (Barutcu et al., 2015). Xiang et al. (2014) demonstrated the role of colorectal cancer-specific long noncoding RNA *CCAT1-L* in facilitating *MYC* transcription through long-range chromatin interactions and its importance in colorectal cancer development.

The formation of chromatin interactions loops are also regulated by proteins such as CTCF and cohesin (Zuin et al., 2014). These proteins interact to allow spatial segregation of chromatin and promote specific chromatin interactions (Hanssen et al., 2017b). Thus, the dysregulation of chromatin-modifying enzymes could

also prevent CTCF and cohesin mediated chromatin interactions, thereby leading to aberrant gene expressions.

1.10 Chromatin Conformation Capture Genomics Technologies

The study of three-dimensional (3D) genomics structures became possible with the development of the C technologies: chromosome conformation capture (3C), circularised chromosome conformation capture (4C), carbon copy chromosome conformation capture (5C), Hi-C and chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) (Wit and Laat, 2012, Fullwood et al., 2009) (**Figure 4**). These technologies allow for the examination of genomic regions that could potentially have chromatin interactions, thereby opening the doors to improved understanding of how specific genomic loci influence each other during gene regulation.

The basis of these C methods is dependent on the capture of the 3D organization of the genomic DNA. Therefore, the initial steps for these methods are always to preserve the chromatin structures with a fixative agent like formaldehyde (Dekker et al., 2002). The fixed genome then undergoes the procedure of restriction enzyme induced fragmentation, ligation and reverse cross-linking procedures to yield a ligated linear DNA strand that contains both interacting fragments from possible chromatin interactions. This is termed as the 3C library which is then quantified for interactions between two regions in the genome (Wit and Laat, 2012). The 3C method is referred to as a 'one versus one' method as it investigates two specific interacting loci.

Circularised chromosome conformation capture (4C) is known as the 'one versus all' method because it allows for the identification of potential sequences that contact one specific locus (Simonis et al., 2006). 4C libraries are constructed in a similar way to the 3C library, with the ligated 3C library being subjected to another restriction enzyme digestion to produce a shorter linear DNA, which can then ligate into a circular form. These circular DNA that are made up of two regions of interactions can then be sequenced using inverse primers that starts from a viewpoint of interest, thereby revealing other possible loci of interaction with the specific locus of interest (Simonis et al., 2006).

The carbon copy chromosome conformation capture (5C) builds on the 3C template by introducing oligonucleotides that anneal to the 3C library template. Primers that are annealed next to each other can then be ligated to form a 5C library (Dostie et al., 2006). The 5C library can then be sequenced through next-generation sequencing to identify the interacting loci along the sequence, thus the 5C library is regarded as a 'many versus many' methods (Dostie and Dekker, 2007).

The Hi-C method allows an “all versus all” identification of chromatin interactions that are present in the whole genome of interest. The Hi C library is prepared similarly to 3C libraries, with the exception that biotin labelled nucleotides are added to restriction enzyme digested ends of the DNA before blunt ligations are performed to link interacting the interacting fragments (Wit and Laat, 2012). These interacting regions can then be pulled down by biotin-binding action for next-generation genome sequencing to identify all the chromatin interactions that are occurring in the genome (Lieberman-Aiden et al., 2009).

Another “all versus all” method is the chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) that allows for the analysis of chromatin interactions through chromatin immunoprecipitation (ChIP) of specific DNA-protein complexes (Fullwood et al., 2009). The initial steps for the ChIA-PET procedure crosslink chromatin interactions by formaldehyde treatment, before sonication to obtain interaction specific fragments. The sonicated fragments are then isolated through ChIP and subsequently connected with DNA linkers through proximity ligation and the pair-end tags (PETs) are then sequenced (Fullwood et al., 2009).

Lastly in the past few years, there the emergence of another “all versus all” method called Hi-ChIP (Mumbach et al., 2016). As the name suggest, this new chromatin interactions analysis method combines both the Hi-C and ChIP procedures together to allow for the interaction of global genome chromatin interactions that are associated with a particular protein of interest. The Hi-ChIP preparation steps are similar to the Hi-C and ChIP protocols with the cells being fixed and subjected to nuclei lysis, restriction enzyme induced DNA fragmentation, biotin labelling and protein isolation steps before the sample is reverse crosslinked and DNA fragments of interest being captured using the biotin-streptavidin method (**Figure 5**) (Mumbach et al., 2016). Lastly, the captured Hi-C contacts are being prepared as a Hi-ChIP library using the Illumina Tn5 library generation kits, PCR amplified and sequenced. The Hi-C method is regarded as an improved version to the ChIA-PET method due to its requirement of up to 100 times less number of cells with a shorter experimental duration (Mumbach et al., 2016). This new method can thus allow for the investigation of chromatin-protein association loop interactions in clinical samples which often face the challenge of having low cell numbers.

The advent of these 'C' techniques along with another next-generation sequencing (NGS) methods such as ChIP-seq and RNA-seq allows for deeper investigation into the role of chromatin interactions and their associated transcription factors on regulating epigenetic patterns and the subsequent alterations in gene expressions in normal and diseased models.

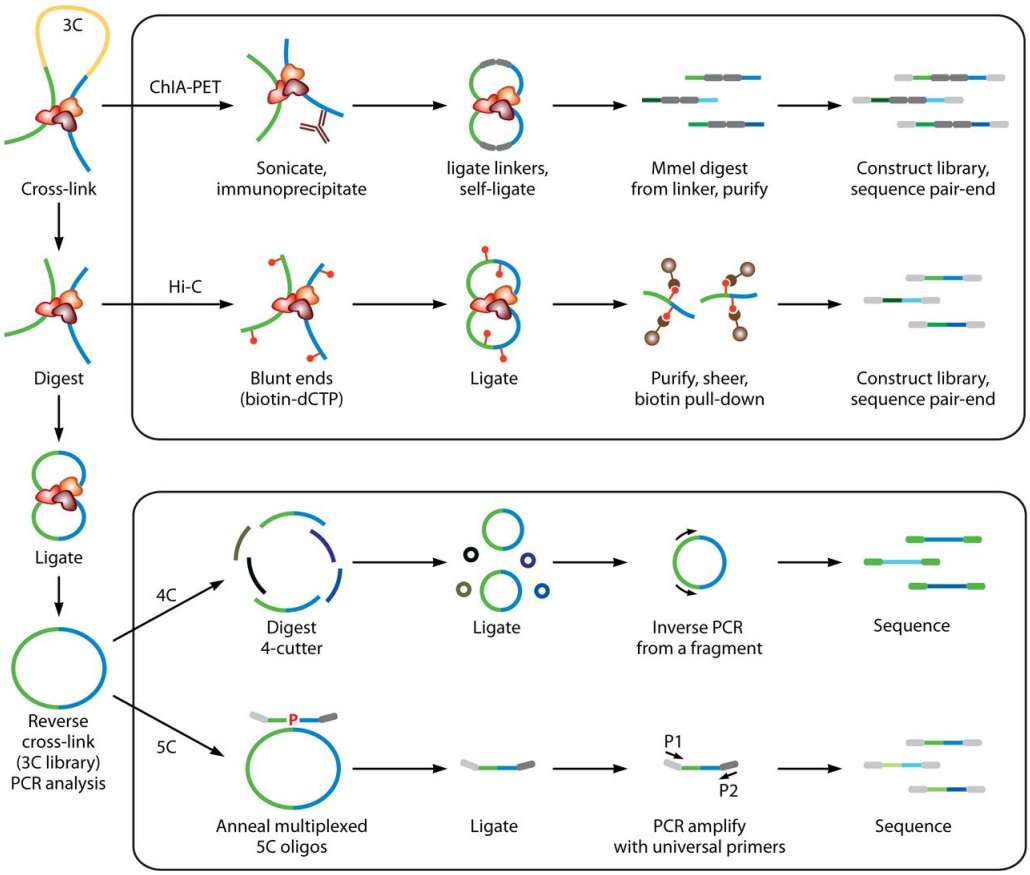

Figure 4. Chart of 'C' techniques used for chromatin interactions studies. All these methods begin with the same samples preparation steps used in generating the 3C library (left). The ChIA-PET and Hi-C methods are used for the study of global chromatin interactions. 4C is used for the study of one to many interactions with a specific region of interest. 5C is used for the study of many to many interactions within a region. The Y-shaped molecule indicates antibodies. The red dots indicate biotinylated nucleotides while the streptavidin beads are show in in brown. Figure is adapted from (Fraser et al., 2015).

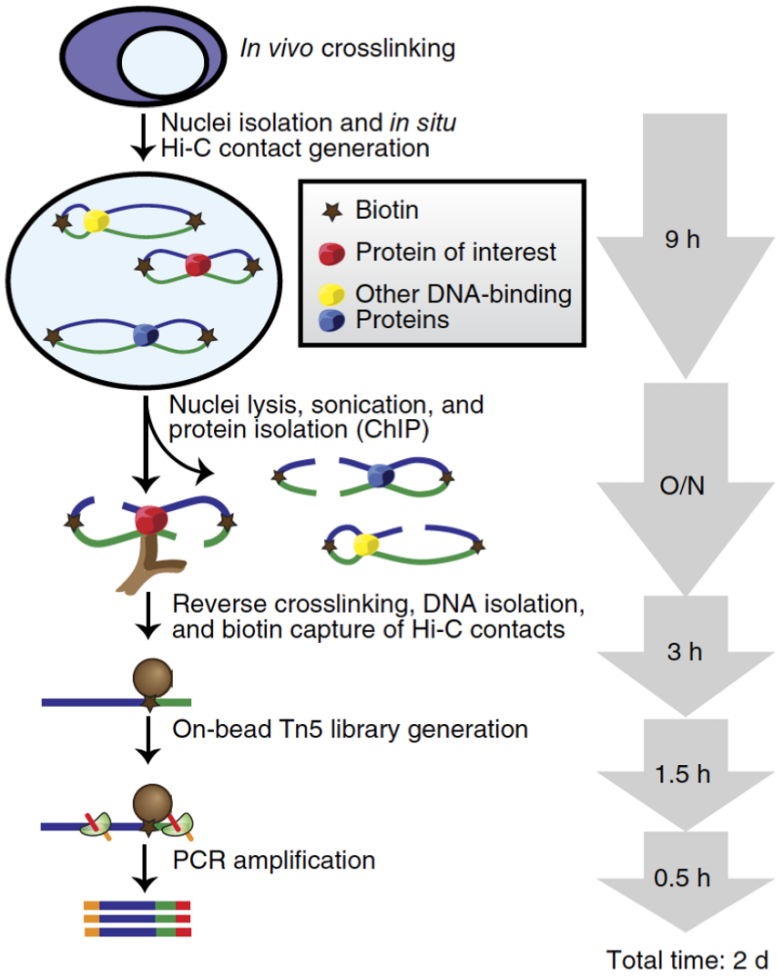

Figure 5. Schematic of Hi-ChIP preparation workflow. The left hand side shows the work flow of Hi-ChIP library preparation and the right depicts the time that is required for each step of the procedure. The star represent the biotin labels. The red circule represents the protein of interest for ChIP pulldown. The chromatin interactions are represented as loops of various colours. Figure adapted from (Mumbach et al., 2016).

1.11 Super-enhancers and Chromatin Interactions

Enhancers are regions that possess the capacity to regulate oncogene expression via chromatin loop formation. The effects of enhancers on gene transcriptions can also be amplified via the clustering of multiple enhancers within close genomic proximity. These densely clustered enhancers within a neighbouring genomic region are termed as a super-enhancer (Wang et al., 2019) (**Figure 6**). Interestingly, super-enhancers are reported to be involved in the control of cell type-specific gene transcription (Hnisz et al., 2013b). Hnisz et al. (2013b) analysed 86 human cell and tissue samples and revealed that the super-enhancers profiles between cells types were distinctly different, with each cell type

possessing super-enhancers that are associated with the expression of key transcription factors that are required for their specific biological uses. Furthermore, previous studies in murine embryonic, myotubes, pro-B, Th and macrophage cells also identified super-enhancers that are linked to the transcription of essential regulatory genes that are specific for each cell type (Whyte et al., 2013).

Furthermore, the application of 3C, 4C and Hi-C techniques coupled with the improved capability of bioinformatics pipelines have provided us with the means to examine how super-enhancers work to modulate gene expressions in normal and diseased conditions. For example, super-enhancers are identified by processing the H3K27ac ChIP-seq signals through bioinformatics pipelines that involve the use of algorithms such as "MACS2" and "Ranking of Super-enhancers" (ROSE) that isolates, merges and ranks the H3K27ac signal peaks in ascending order, with those above the cutoff limit determined as super-enhancers (**Figure 6**). The H3K27ac signal enrichment is a surrogate marker that indicates for enhancers and was also shown to be strongly associated with super-enhancers presence (Hnisz et al., 2013b), thus the mining of super-enhancers from cells' H3K27ac profiles is widely adopted. Through these bioinformatics analyses, we can obtain the locations of super-enhancers residency in each cell type. Since the 'C' techniques allow us to study global or individual regions of chromatin interactions, the mapping of chromatin loops profiles to those of the super-enhancers within a specific cell type can reveal insights into the how specific genes are modulated by individual super-enhancers.

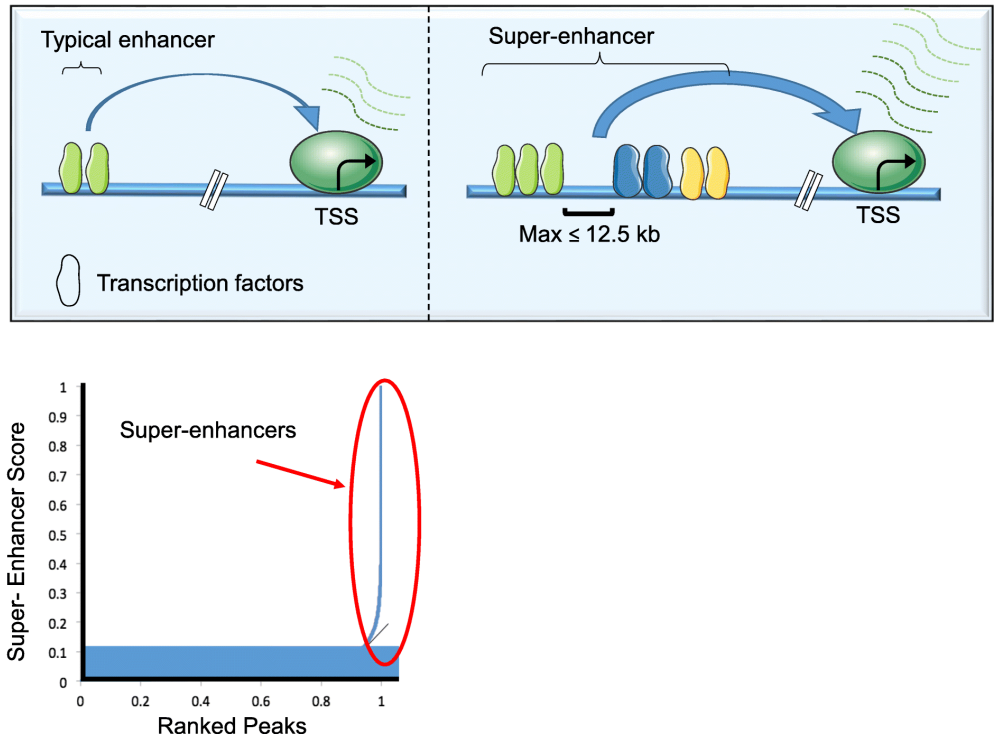

Figure 6. Schematic illustrating the mechanism of action for enhancer and super-enhancers. The top figure shows the interaction between an enhancer/super-enhancer and a transcription start site (TSS) in cis. Super-enhancers are regions of enhancers that are typically clustered together within a 12.5kilobase distance (left). Super-enhancers are able to induce higher gene expression as compared to the enhancer mediated transcription. Large numbers of transcription factors can bind to the super-enhancer as compared to the enhancer to modulate transcription via chromatin loops with the TSS. The lower image shows the image of H3K27ac ChIP-seq signals that are ranked in ascending order of signal strength. The peaks that fall under the scale of 1 are called as super-enhancers. Figure is adapted from. (Jia et al., 2019)

Using the data from the experiments of cohesin depletion and rescue in the HCT-116 human colorectal cancer cell lines (Rao et al., 2017), Ryu et al. (2019), analysed the changes in super-enhancer associated loops upon cohesin depletion and reintroduction. Their analysis demonstrated that chromatin loops that are associated with super-enhancers that are commonly present across 30 different human cell types appear to recovery rapidly upon the reintroduction of cohesin proteins. Distal super-enhancers to genes can also be transcribed into enhancer-RNAs (eRNA) that aids in the formation of super-enhancer-promoter chromatin interactions. A super-enhancer that is located in between the gene loci of the embryonic transcription factors Nanog and Dppa3 was shown to regulate their gene expression in embryonic stem cells through chromatin interaction with the

enhancer RNA that is transcribed from the super-enhancer region (Blinka et al., 2016). The importance of this particular distal super-enhancer derived RNA to gene expression was observed when the deletion of the super-enhancer region resulted in decreased *Nanog* and *Dppa3* expression levels. Furthermore, the CRISPR/Cas9 deletion of two super-enhancer regions upstream of the *Nanog* gene result in 40-50% in *Nanog* gene expression (Blinka et al., 2016). These suggest that multiple super-enhancer/enhancers may regulate specific gene expression under specific conditions such as in embryonic stem cells.

Intriguing, super-enhancers are shown to bind to high levels of transcription factors (Whyte et al., 2013). Hnisz et al. (2013b) analysed for transcription factors enrichments at super-enhancers in embryonic stem cells and observed that embryonic stem cells related transcription factors such as Oct4, Sox2, Nanog and Klf4 binds to the super-enhancer regions that are responsible for modulating their gene transcription. These patterns of transcription factors enrichment at super-enhancers are similar to previous reports that show the presence of transcription factors Oct4, Sox2 and Nanog at enhancer regions in mice (Chen et al., 2008). A feature of super-enhancers is that they can attract significantly higher numbers of transcriptional proteins at its binding motifs (Hnisz et al., 2013b, Wang et al., 2019). In addition, these evidence supports the proposed model by Boyer et al. (2005), which states that transcription factors bind to their own super-enhancers to drive its own transcription and that this develops into a cycle which promotes the auto-autoregulation of gene expression.

An interesting aspect of super-enhancers is that they tend to be acquired during disease state as compared to normal conditions. Single Nucleotide Polymorphisms (SNP)s are discovered to develop at super-enhancers that are associated with the driving of Alzheimer's and Type 1 diabetes-related genes expression (Hnisz et al., 2013b), which may thus become motifs for aberrant transcription factors binding to initiate gene overexpression.

Super-enhancers are also noted to be dynamically regulated. From a hair follicle stem cell differentiation model, super-enhancers were observed to be remodelled and shifting to different locations of the genome as the cell differentiates or with changes in their microenvironment such as during hair cell injury (Adam et al., 2015). These changes in super-enhancer regions also influence alterations in the areas and types of transcription factors binding (Adam et al., 2015). Furthermore, it was suggested that the protein SOX9 was responsible for the regulation of super-enhancers by preventing its silencing by H3K27me3 and promoting the deposition of activating histone marks such as H3K27ac (Adam et al., 2015).

Based on these collective understanding of chromatin interactions between enhancers/super-enhancers-promoter regions and the presence of similar transcription factors at both interacting regions, it is surmised that these chromatin loops may be involved in the transfer of transcriptional proteins and complexes between the super-enhancer and gene promoter region to promote transcription. However, these remain to be further understood and investigated.

Sabari et al. (2018) demonstrated that super-enhancers can interact with gene regions in a three-dimensional manner through the formation of phase-separated droplets in the nucleus. Transcriptional co-activating proteins such as MED1 and BRD4 which are strongly enriched at super-enhancers and gene promoters (Lovén et al., 2013) were found to be interacting in the phase condensates in the nucleus murine embryonic stem cells (Sabari et al., 2018). The disruption of MED1 and BRD4 proteins in the phase condensates with the use of 1,6 hexanediol resulted in the loss of MED1, BRD4 and RNA polymerase II binding at the super-enhancers and chromatin regions. Furthermore, they demonstrated that MED1 and BRD4 proteins can form weak protein to protein interactions in-vitro (Sabari et al., 2018), which supports the idea that super-enhancers and gene promoters can interact in three-dimensional fashion with chromatin interactions potentially contributing to it.

1.12 Super-enhancers are associated with aberrant oncogenes in myeloid leukaemia

As mentioned above, super-enhancers are potentially gained during disease development and help to drive aberrant genes expression in disease pathogenesis. The role of super-enhancers was also observed in myeloid leukaemia. Lovén et al. (2013), reported the presence of super-enhancers that are associated with oncogenes such as *MYC* that are highly relevant in multiple myeloma. While in the analysis of 18 human cancer cells, it was revealed that cancer cells gained super-enhancers in the proximity of oncogenes that are tumour specific in contrast to the absence of such super-enhancers in healthy cells (Hnisz et al., 2013b). For example, super-enhancers were observed to be present in the gene regions that were surrounding the *MYC* oncogene in multiple types of cancer cells (Hnisz et al., 2013b). Recently, our lab also found that super-enhancer regions have a higher association with chromatin interactions than normal enhancers and that super-enhancers interact with cancer-related oncogenes and tumour suppressor genes (Cao et al., 2017a).

An analysis with the H3K27ac ChIP-seq of 19 primary pediatric AML samples discovered a high frequency of super-enhancer presence at the *RARA* (retinoic acid receptor alpha) gene (Daramola et al., 2019, Fiore et al., 2016). Interestingly, the RARA targeting drug tamibarotene which was found to be effective against high *RARA* expressing high acute promyelocytic leukaemia was successful in reducing proliferation and increasing cellular apoptosis in pediatric AML samples which possess the super-enhancer at the *RARA* gene (Daramola et al., 2019). This result was further reiterated when tamibarotene was applied to another pediatric AML cell; Kasumi which does not possess a super-enhancer at *RARA* and showed no effects on cellular proliferation and apoptosis changes (Daramola et al., 2019).

As super-enhancers are regions that consist of clusters of enhancers that exhibits strong levels of transcription factors and co-factors like Mediator (Pott and Lieb, 2015), the approach of inhibiting accumulated transcription factors could result in the disruption of subsequent gene expression. Lovén et al. (2013), reported that super-enhancers are more sensitive and susceptible to disruption as

compared to normal enhancers. This was demonstrated when multiple myeloma cells were treated with the bromodomain and extra-terminal (BET) inhibitor JQ1. The results of the treatment with JQ1 led to the loss of BRD4 transcription factors that were recruited to the super-enhancer region (Lovén et al., 2013). Therefore, drug targeting of super-enhancers may be a probable option in the treatment of diseases like cancer that could be mediated by super-enhancer association. Since super-enhancers largely exist in a cell type-dependent manner, its role in upregulating oncogenic gene expression in myeloid leukaemia should be explored at a deeper level.

Besides the aberrant regulation of epigenetic genes, gene expression profiling in AML revealed that 48 messenger RNA (mRNA), 52 micro RNAs (miRNA) and 73 long non-coding RNAs (lncRNA) are among the top 100 genes that are associated with poor prognosis in patients (Li et al., 2020), thus suggesting that these genes are also dysregulated in AML. In addition, somatic gene mutations at AML associated genes like *RUNX1*, *TLE4* and *SHKBP1* are also seen (Greif et al., 2011). Furthermore, Gerber et al. (2013) identified 97 genes that are differentially expressed in CML as compared to normal stem and progenitor cells. The differentially expressed genes include those that regulate cell-surface proteins such as *DPP4* and *IL2RA* as well as cell proliferation promoting genes like *MYCN* (Gerber et al., 2013). Collectively, these studies strongly suggest that a large number of transcriptional and cellular processes associated genes are dysregulated in myeloid leukaemia. Therefore, the understanding of the mechanisms driving these aberrant gene expressions is crucial towards the development of effective treatments in myeloid leukaemia.

Examples among the growing list of oncogenes that are being studied are the *HOXA9* and *MEIS1* oncogenes. MEIS1 (The Myeloid Ectropic viral Integration Site) and HOXA9 are transcription co-factors that are reported to be significantly dysregulated in myeloid leukaemia, especially in AML. Earlier reports have demonstrated that *MEIS1* and *HOXA9* transduced mice rapidly developed aggressive AML (Morgado et al., 2007).

1.13 Dysregulated *HOXA9* expression in myeloid leukaemia

The *HOXA9* gene belongs to the cluster of genes termed as homeobox genes and was found to be overexpressed in 78% of AML cases (Lebert-Ghali et al., 2016, Casas et al., 2003). The encoded HOXA9 transcription factor has been revealed to participate in downstream genes regulation by binding to gene promoters and distal enhancers regions (Collins and Hess, 2016b). *HOXA9* is part of the *HOXA* locus and encodes the homeodomain transcription factors that are associated with the maintenance of hematopoietic stems cells (Li et al., 2013a). *HOXA9* is also suggested to be a predictor of AML prognosis, with AML patients who expressed a high level of *HOXA9* associated with a poorer prognosis (Tholouli et al., 2012, Collins et al., 2014).

The activation of *HOXA9* expression is dependent on the trimethylation of its promoters by the MLL protein (Milne et al., 2002) as well as epigenetic modifiers and lncRNA that are located upstream of the *HOXA* locus (Collins and Hess, 2016b). *HOX* genes expression levels is understood to be important in the maintenance of cell hematopoiesis and *HOX* genes expression levels are usually silenced as cells mature and lose CD34+ expression (Collins and Hess, 2016b). Overexpression of *Hoxa9* in mice led to the proliferation of hematopoietic stem cell and the eventual development of AML (Kroon et al., 1998, Thorsteinsdottir et al., 2002). While murine models with the homogenous knockout of *Hoxa9* showed reduced ability in the repopulation of their myeloid and lymphoid cells (Magnusson et al., 2007). In addition, the HOXA9 protein was demonstrated to recruit chromatin-modifying enzymes such as the histone acetyltransferase p330/CBP to initiate the activation of downstream targets (Dintilhac et al., 2004). The collaborative relationship of the HOXA9 protein and the histone methyltransferase G9a was also demonstrated to contribute to leukaemia progression in the mouse model (Lehnertz et al., 2014).

The transcription of *HOXA9* in myeloid leukaemia is an important factor for the development of mixed-lineage leukaemia (*MLL*) rearranged AML (Sun et al., 2018b). Faber et al. (2009) demonstrated that the depletion of the *HOXA9* oncogene in 17 human acute myeloid leukaemia and acute lymphoblastic leukaemia with germline or rearranged MLL resulted in the reduction of other

oncogenes expressions such as *HOXA10*, *MEIS1* and *PBX3.* Importantly, *MLL* rearranged leukaemia cells with depleted *HOXA9* expression showed increased proliferation arrest and apoptosis phenotypes (Faber et al., 2009).

The *HOXA9* gene was also reported to form various types of fusion genes in partnership with other oncogenes. For instance, *HOXA9* partners *NUP98* to form a *NUP98-HOXA9* fusion gene in AML, with its formation being attributed to the occurrence of the t(7;11) translocation (Nakamura et al., 1996). AML patients who possess the NUP98-HOXA9 fusion protein typically show a poor prognosis (Rio-Machin et al., 2017). Interestingly, the NUP98-HOXA9 protein interacts with the MLL proteins in leukemogenesis and that leukemogenesis is severely affected by the ablation of the MLL protein (Shima et al., 2017).

Till date, the mechanisms behind *HOXA9* gene regulation remains limited. In an example of one possible way of *HOXA9* regulation, Wang et al. (2011) demonstrated that the lncRNA *HOTTIP* regulated the *HOXA* locus via chromosomal looping that brings the lncRNA in close proximity to facilitate gene transcription. *HOTTIP* lncRNA was elucidated to bind to the WDR5/MLL (Mixed Lineage Leukaemia) complex for transcription activation to occur on the *HOXA* genes. *HOTTIP* lncRNA influence on *HOXA* genes was also suggested to decrease as the distance of locus increased from *HOTTIP* loci (Wang et al., 2011). The *HOTTIP* lncRNA was also revealed to be upregulated in AML patients as compared to normal bone marrows (Hao and Shao, 2015), suggesting its potential link with the upregulation of *HOXA9* in AML pathogenesis.

1.14 Myeloid Ectropic viral Integration Site (MEIS1) cooperates with HOXA9 in aggressive AML

The HOXA9 protein works in tandem with several co-factors in myeloid leukaemia pathogenesis. MEIS1 (The Myeloid Ectropic viral Integration Site) protein is one of the most widely studied co-factors of HOXA9 in myeloid leukaemia, with *HOXA9* and *MEIS1* expression uniquely seen in myeloid leukaemias with severely negative effects on disease prognosis (Lawrence et al., 1999) (**Figure 7**). The presence of *MEIS1* in additional to *HOXA9* overexpression often leads to the rapid development of myeloid leukaemia as compared to leukaemias with only high *HOXA9* expression (Kroon et al., 1998, Zeisig et al., 2004).

MEIS1 belongs to the three-amino-acid loop extension (TALE) family of homeodomain proteins (Moskow et al., 1995). *MEIS1* along with *HOXA9* are suggested downstream gene targets of MLL fusion proteins in *MLL* leukaemias (Imamura et al., 2002, Wong et al., 2007). Independently, *MEIS1* is also a predictor for poor prognosis in myeloid leukaemia (**Supplementary Figure 38**). Furthermore, *Meis1* was also shown to be a factor that inhibits the progression of *MLL*-leukaemia in mice (Imamura et al., 2002). Ferreira et al. (2016)'s work with *DNMT3A* mutated AML also suggested that mutations in *DMNT3A* results in the hypomethylation and increased *MEIS1* expression in the absence of *MLL* fusion occurrence in AML.

The functionality of the MEIS1 protein as a cofactor of the HOXA9 protein was suggested to be dependent on another TALE protein PBX3 (Garcia-Cuellar et al., 2015). The loss of Meis1-Pbx3 interactions resulted in the destabilization and degradation of the Meis1 protein, suggesting that Pbx3 to Meis1 protein-protein interaction is required for the stabilization of the Meis1 protein (**Figure 7**). Furthermore, the dimerization of Meis1 and Pbx3 proteins was noted to be needed for stable Meis1 and Hoxa9 partnership (Garcia-Cuellar et al., 2015). Pbx3 protein was also reported to contribute towards *Meis1* regulation as the overexpression of *Pbx3* resulted in subsequent upregulation of *Meis1* expression (Garcia-Cuellar et al., 2015).

Besides MLL and PBX3 modulation of *MEIS1* expression, enhancers and super-enhancers are also suggested to contribute to *MEIS1* transcription in cancer, in particular myeloid leukaemia. In a study, a novel enhancer that is associated with acute leukaemia was discovered at the *MEIS1* locus in a zebrafish embryonic hematopoiesis model (Wang et al., 2014). This enhancer was observed to possess high H3K4 monomethylation and H3K27ac deposition, which are predictors of an active enhancer (Wang et al., 2014). The authors also suggested that the MEIS1 and HOXA9 proteins can bind to this enhancer and the *MEIS1* promoter to initiate *MEIS1* transcription in an autoregulatory loop fashion. Interestingly, *MEIS1* expression was downregulated along with the decrease in H3K4 monomethylation and H3K27 acetylation disposition at the enhancer upon cell differentiation (Wang et al., 2014). This observation agrees with the widely seen pattern of *MEIS1* expression being specifically upregulated during cell hematopoiesis.

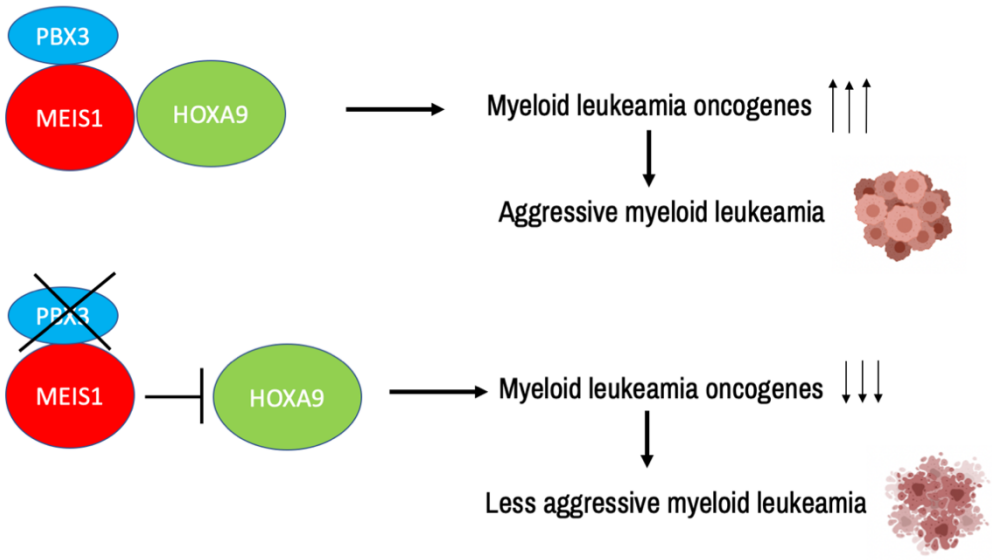

Figure 7. Schematic of MEIS1, HOXA9 and PBX3 relationship on myeloid leukaemia development. PBX3, HOXA9 and MEIS1 proteins are indicated as blue, red, green coloured ovals respectively. Arrow indicate up or downregulation of oncogenes expression. Top image represents growing leukaemia cells and lower image represents dying leukaemia cells. (Part of illustration in this figure was generated with the website Biorender.com).

Altogether, the current literature demonstrates the strong relationship between HOXA9 and MEIS1 proteins in aggressive myeloid leukaemia progression. While the MLL fusion proteins and enhancers have been suggested to initiate *HOXA9* and *MEIS1* transcription in myeloid leukaemia, further work is required to better understand the collective mechanisms of how these oncogenes are transcribed. In addition, enhancers & super-enhancers are demonstrated to be acquired to drive genes development under specific conditions such as in cancer. However, the role of 3D genome organization on *HOXA9* and *MEIS1* transcription and the coordinated expression of *HOXA9-MEIS1* oncogenes through altered TADS and acquired super-enhancers in myeloid leukaemia remains limited.

Therefore, in this study, I aim to investigate the molecular basis of coordinated dysregulation of the *HOXA9-MEIS1* signalling axis in myeloid leukaemia, with specific attention on the roles of TADs and enhancer/super-enhancer-promoter chromatin loops in regulating the aberrant expression of oncogenes such as *HOXA9* and *MEIS1* in myeloid leukaemia.

1.15 Objectives & Hypotheses

The specific objectives of this study are as follow:

1. To investigate potential dysregulated TADs/FIREs and loop alterations in myeloid leukaemia at *HOXA9* and *MEIS1*.

2. Characterize the interplay between super-enhancer acquisition, TAD and loop alterations in myeloid leukaemia at *HOXA9* and *MEIS1*.

3. To understand the molecular basis for coordinated *HOXA9* and *MEIS1* gene expression.

Hypotheses:

1. Changes in TAD/FIREs compartments and chromatin loops can lead to the increased transcription of *MEIS1 and HOXA9* oncogenes in myeloid leukaemia (**Figure 8**).

2. Acquired super-enhancers at regions of existing chromatin interactions with the *MEIS1* and *HOXA9* promoter leads to high *HOXA9* and *MEIS1* expression in myeloid leukaemia (**Figure 9**).

3. Transcriptional proteins such as the HOXA9 and MEIS1 proteins can bind to the HOXA9 and MEIS1 promoters to regulate their expression in myeloid leukaemia (**Figure 10**).

The central hypothesis of my thesis is that altered TADs and chromatin interactions as well as the acquisition of super-enhancers result in the dysregulation of *HOXA9* and *MEIS1* oncogenes in myeloid leukaemia which contributes to rapid disease progression. These dysregulated chromatin interactions at the *HOXA9* and *MEIS1* oncogenes are also mutually bound and regulated by the HOXA9 and MEIS1 proteins.

If my above hypotheses are true, then the loss of a CTCF boundary between two adjacent TADs at the *MEIS1* and *HOXA9* gene loci will result in the formation of new chromatin loops between the enhancers/super-enhancers and gene promoters such as *HOXA9* and *MEIS1* that are originally located in a separate TAD. Also, the disruption of these chromatin loops between super-enhancers and the *HOXA9* and *MEIS1* promoters as well as the binding of HOXA9 and MEIS1 proteins at each other's promoter will result in the upregulation of *HOXA9* and *MEIS1* gene expression in myeloid leukaemia.

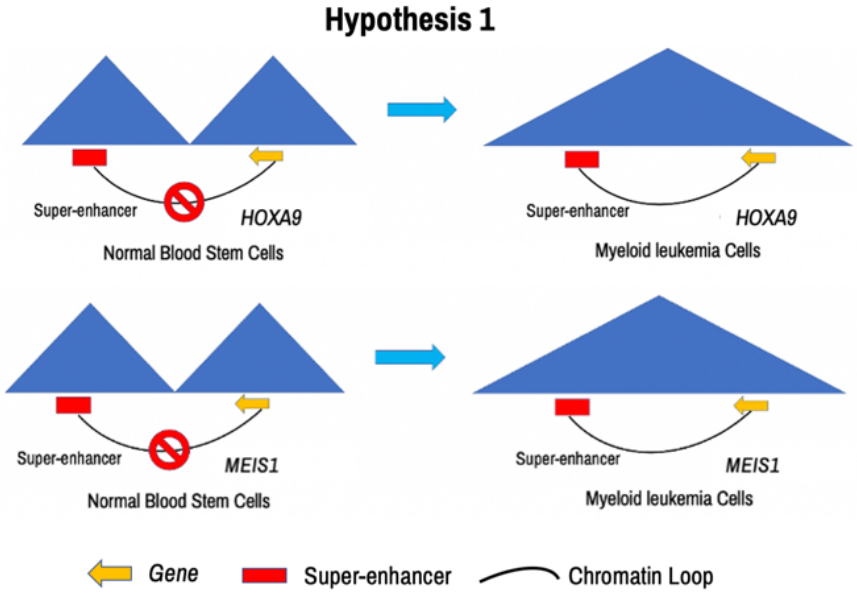

Figure 8. Schematic of hypothesis 1 showing that the hypothesized difference between normal blood stem cells and myeloid leukaemia cells. The loss of TAD/FIRE boundary may result in altered chromatin interactions between HOXA9 or MEIS1 and their associating super-enhancers region. The triangles represent a compartment like a TAD or FIRE. The green arrow represents the gene location and the red box represent a super-enhancer region. Chromatin loops are shown as curve lines.

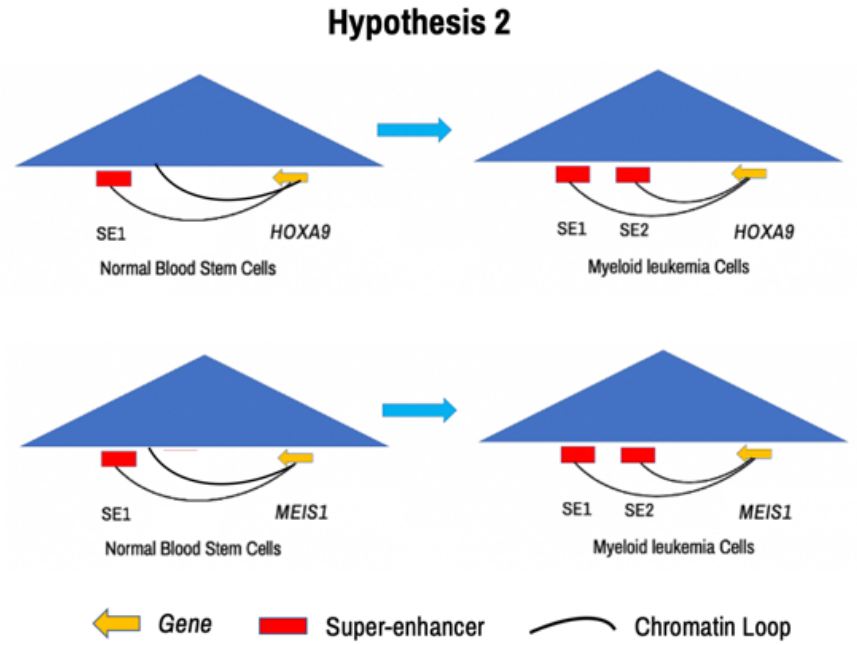

Figure 9. Schematic of hypothesis 2 showing the hypothesized difference in super-enhancers locations in normal blood stem cells and myeloid leukaemia cells. The gain of super-enhancers at regions of pre-existing chromatin loops can mediate overexpression of *HOXA9* and *MEIS1* in myeloid leukaemia. The triangles represent a compartment like a TAD or FIRE. The green arrow represents the gene location and the red box represent a super-enhancer region. Chromatin loops are shown as curve lines.

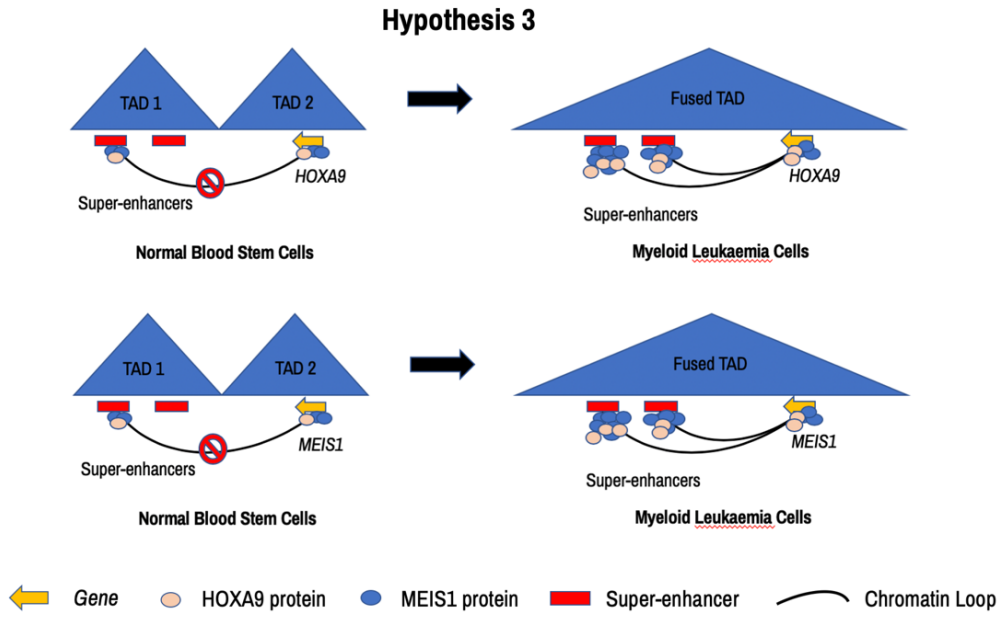

Figure 10. Schematic of hypothesis 3 showing increased binding of HOXA9 and MEIS1 proteins at their own gene promoters and their associated super-enhancers that are interacting via chromatin loops that may be formed from altered TADs/FIRE. The triangles represent a compartment like a TAD or FIRE. The green arrow represents the gene location and the red box represent a super-enhancer region. Chromatin loops are shown as curve lines.

2 Material & Methods

2.1 Cell Culture

Human Acute Myeloid Leukaemia cells THP-1 and HL-60, Chronic Myelogenous Leukaemia cells K562, and B Lymphoblastoid cells GM12878 were cultured at 5% CO_2 at 37^0C. THP-1, K562 and GM12878 were cultured with Roswell Park Memorial Institute (RPMI) 1640 media (Hyclone) supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS; Hyclone) and 1% penicillin/streptomycin (Hyclone). HL-60 cells were cultured using Iscove's Modified Dulbecco's Medium (IMDM; Gibco), supplemented with 20% heat-inactivated FBS (Hyclone) and 1% penicillin/streptomycin (Hyclone).

2.2 Clinical Samples Preparation

Bone marrow samples from AML patients were taken from the back of the pelvic (hip) bone (Domain Specific Review Board "DSRB" Reference Number: 2015/00804) while bone marrow from healthy counterparts were withdrawn during Total Knee Arthroplasty as part of a standard operative procedure (DSRB Reference Number: 2015/01037). All clinical samples were obtained from the National University Hospital Singapore and collected according to the requirements of the Human Biomedical Research Act. Informed consent was obtained for all clinical samples used in the study.

Knee bone aspirates collected from normal individuals were resuspended in 50 ml PBS containing 10% FCS, 0.4% Sodium Citrate and 3 mM EDTA, and pipetted to homogenize. Homogenized samples were passed through a 100 μm cell strainer and pelleted at 300 × g for 5 minutes. Samples were then lysed in 15 ml ACK Lysing Buffer (Cat No. A1049201, Gibco) following a 5-minute incubation at room temperature. Blood cells were pelleted at 300 × g for 10 minutes at room temperature and resuspended in 12 ml PBS containing 2% FBS and 2 mM EDTA. Samples were then layered on Ficoll-Paque Plus in 6 ml cell suspension: 4.5 ml Ficoll ratio and centrifuged at 400 × g for 30 minutes without acceleration and braking. Mononuclear layer was collected and washed twice with 13 ml PBS containing 2% FBS and 2 mM EDTA.

Bone marrow aspirates collected from AML patients were resuspended to a total of 12 ml PBS containing 2% FBS and 2mM EDTA and pipetted to homogenize. Homogenized samples were then layered on Ficoll-Paque Plus in 6 ml cell suspension: 4.5 ml Ficoll ratio and centrifuged at 400 × g for 30 minutes without acceleration and braking. The mononuclear layer was collected and washed twice with 13 mL PBS containing 2% FBS and 2 mM EDTA.

2.3 Isolation of CD34+ hematopoietic stem cells from total mononuclear cells (MNC)

Isolation of CD34+ cells from AML and normal sample MNCs was performed according to the manufacturer's instructions using CD34 MicroBead Kit UltraPure, human (Cat No. 130-100-453, Miltenyi Biotec, Germany). The purity of CD34+ cells was determined by flow cytometry using BD LSR II Flow Cytometer (BD Biosciences, San Jose, CA) at post-isolation with cell marker (CD34, clone 8G12, Cat No 348057, BD Biosciences, San Jose, CA) after exclusion of cell debris based on scatter signals and dead cells by DAPI fluorescent stain (Cat. No. 130-111-570, Miltenyi Biotec, Germany). Data analysis was performed by FACSDiva software.

We adapted the manufacturing instructions of Miltenyi Biotec to establish the isolation protocol for CD34+ cells from AML patient sample MNCs. The percentage of CD34+ cells was determined by flow cytometry at a BD LSR II Flow Cytometer pre- and post-isolation. The amount of CD34 microbeads added varied following the sample's percentage of CD34+ cells: if the percentage of CD34+ cells in the sample was under 20%, the manufacturer's protocol was followed. If the percentage of CD34+ cells was between 20-50%, twice the amount of CD34 Ultrapure microbeads was added; if the percentage was above 50%, thrice the amount was added.

2.4 RNA Extraction, Reverse Transcription (RT) and Quantitative Polymerase Chain Reaction (qPCR)

RNA of clinical samples was extracted using the DNA/RNA Allprep Kit (Qiagen). RNA of THP-1, HL-60, K562 & GM12878 cell lines were extracted using the RNeasy Mini Kit (Qiagen) and according to the manufacturer's instructions. Reverse transcription of RNA into cDNA was performed using the qScript cDNA Supermix (Quantabio). Quantitative polymerase chain reaction (qPCR) was performed with the GoTaq qPCR Mastermix (Promega) and QuantStudio 5 Real Time PCR (Applied Biosystems). GAPDH was selected as the endogenous control for qPCR. Alternatively, RNA-Seq analysis was performed (see the section on "RNA-Seq analysis")

2.5 Droplet Digital Polymerase Chain Reaction (ddPCR)

ddPCR experiments on cDNA were performed with the EvaGreen Mastermix (Biorad) and the QX200 Droplet Digital PCR system. Post analyses were done with the Quantasoft Analysis Pro Softward (Biorad).

2.6 Quantitative Polymerase Chain Reaction (qPCR)/ Droplet Digital (ddPCR)

Table 5. List of primers used in qPCR & ddPCR (5' to 3')

***GAPDH* Forward**	GCACCGTCAAGGCTGAGAAC
***GAPDH* Reverse**	GGATCTCGCTCCTGGAAGATG
***HOXA9* Forward**	TAAACCTGAACCGCTGTCGG
***HOXA9* Reverse**	GCCTTCGCTGGGTTGTTTTT
***MEIS1* Forward**	ACGGCATCTACTCGTTCAGG
***MEIS1* Reverse**	CCATCACCTTGCTCACTGCT
***MYC* Forward**	CACCAGCAGCGACTCTGA
***MYC* Reverse**	GATCCAGACTCTGACCTTTTGC

2.7 Circularised Chromosome Conformation Capture (4C)

4C-seq was performed as previously described with some modifications (Splinter et al., 2012). In brief, 4×10^7 cells were harvested and crosslinked with 1% formaldehyde for 10 min at room temperature with rotation. The crosslinking was quenched by glycine for 5 min at room temperature with rotation. Following SDS and Triton X-100 permeabilization, nuclei were digested with HindIII-HF (NEB) restriction enzyme overnight. Following proximity ligation, reverse cross-linking and DNA purification, the circular DNA was digested with DpnII (NEB) restriction enzyme at 37 °C overnight and circularized. The 4C-seq library was generated by performing nested inverse PCR using Phusion DNA polymerase (Thermo Scientific) with the primers listed below. 10% of the 1st PCR product was used for the 2nd PCR. The 4C-seq library was purified by 4–20% gradient TBE PAGE gel (ThermoFisher Scientific) and the smear band regions including the expected sizes were excised. The library was recovered by incubating the crushed gel slice with 200 µl of TE buffer overnight at 37 °C and the DNA in the supernatant was ethanol precipitated in the presence of GlycoBlue (ThermoFisher Scientific). The multiplex 4C-seq library was pooled in equal molar ratio and sequenced on MiSeq (Illumina) with 1X150 bp. At least 1,000,000 reads were produced for each library. BWA-Mem (0.4.17-r1188) was used to map to the human genome. The mapped 4C-seq data was analysed by r3CSeq (1.30.0) via the Cancer Science Institute web portal.

Table 6. 4C Outer Primers (5’ to 3’)

***MEIS1* Forward**	ATTGCCTGGGCTTTAATTCC
***MEIS1* Reverse**	GGCGCGAGGTTCTTACA
***HOXA9* Forward**	ATCTCTCCCTACTCCCAACA
***HOXA9* Reverse**	TGGGATTCTTTTTGCCCATT

Table 7. 4C Nested Primers (5' to 3')

***MEIS1* Forward**	TCTCTTTATTCTCCCGCTCT
***MEIS1* Reverse**	TCGGCGCCTCTAAGACA
***HOXA9* Forward**	GACCAGGGCATTGGATTTAT
***HOXA9* Reverse**	GGAAGCCACTAGTAAGCAAG

2.8 CRISPR-Cas9 Plasmid Cloning

CRISPR-Cas9 excision was performed with the All-in-One vector system as described previously (Sakuma et al., 2014). Two sgRNA oligonucleotides (Integrated DNA Technologies) targeting two regions of interest (chr2:66,802,343-66,802,362 & chr2:66,803,315-66,803,334) were designed using the Benchling web interface (Benchling [Biology Software] (2019). The strand and anti-strand oligonucleotides were then annealed. The annealed region was ligated into the pX330A-Cas9-2A-GFP or pX330S-Cas9-2A-GFP vector plasmid by Golden Gate Assembly (Sakuma et al., 2014). The Golden Gate Assembly was performed with the pX330A-Cas9-2A-GFP and pX330S-Cas9-2A-GFP plasmids containing each targeted cut side to yield a single fused pX330A-Cas9-2A-GFP plasmid containing two cut sides. Generated all-in-one plasmids were confirmed for the successful insertion of the sgRNAs by Sanger sequencing (1st base) using the CRISPR-step2-F and CRISPR–step-2-R primers.

Table 8. CRISPR-Cas9 Excision Primers (5' to 3')

Region One	AAGCCAAAAAACGTGCCTTG
Region Two	TGCCCCGAGAGGAAATCCAG

Table 9. CRISPR-Cas9 Sanger Sequencing Primers (5' to 3')

CRISPR-step2-F	GCCTTTTGCTGGCCTTTTGCTC
CRISPR-step2-R	CGGGCCATTTACCGTAAGTTATGTAACG

2.9 Transfection

Transfection of K562 cells was performed with the Neon Transfection System (ThermoFisher). Briefly, 5µg of pX330A-Cas9-2A-GFP with two cut sites was added to 1 million K562 cells and electroporated. Transfected cells were kept at 37^{o}C with 10% fetal bovine serum and 1% of penicillin-streptomycin RPMI1640 media for 48 hours before being Fluorescence-activated cell sorting (BD FACSAria Flow Cytometer) sorted for GFP fluorescing positive clones. Each positive clone was sorted into a single well of a 96 well-plate and cultured until a visible cell pellet could be observed.

2.10 Genotyping

Cell pellets of the positive CRISPR clones were resuspended and passage into one well of a 6 well plate containing 10% Fetal Bovine Serum and 1% of penicillin-streptomycin RPMI1640 media and further cultured for another five days. One million cells were subsequently harvested from each clone and their genomic DNA was extracted (Wizard SV Genomic DNA purification system, Promega). Genotyping of each clone was done with region-specific internal and flanking primers.

Table 10. Genotyping Primers (5' to 3')

Flanking Forward	CTGCAATTCATCCGCTGCTC
Flanking Reverse	TCCCAGGCTCCTGTAGTCTC
Internal Forward	CGACTCGGTAGGAAACGGAG
Internal Reverse	CACACAGCAACTAACCCCGA

2.11 Growth curve assay

10 000 cells/well were seeded in 96 well plates and measured for cell growth at 0, 24, 48, 72 hours using the Cell Titer Glo assay kit (Promega, G7571). Luminescence was measured using the Tecan M200 plate reader.

2.12 Chromatin Immunoprecipitation-Quantitative Polymerase Chain Reaction (ChIP-qPCR)

Cells were crosslinked with 1% formaldehyde (Thermo Scientific) for 15 minutes and quenched with glycine for 5 minutes at room temperature. Following this, the crosslinked cells were lysed with 1% SDS lysis buffer supplemented with protease inhibitor (Roche). Lysed cells were sonicated at 25 cycles with the Bioruptor Pico (Diagenode) and subsequently added to antibody-conjugated A/G beads (Invitrogen) and rotated overnight at 4^0C. Anti-HOXA9 (Sigma-HPA061982) and Anti-MEIS1 (abcam-ab19867) antibodies were used. The incubated beads are then washed in the following order: thrice with 0.1% SDS lysis buffer, once with high salt wash buffer, once with lithium chloride wash buffer and once with Tris-EDTA buffer. The beads were eluted in ChIP elution buffer before treatment with RNase A (Qiagen) and Proteinase K (Ambion) at 37^0C for 4 hours. The ChIP DNA was cleaned up with the QIAquick PCR

purification kit (Qiagen). The ChIP DNA was then used for performing of ChIP-qPCR. All ChIP-qPCR experiments were performed with four biological replicates of cells.

Table 11. ChIP-qPCR Primers (5' to 3')

***HOXA9* Region Forward**	GCCCGTCCAGCAGAACAATA
***HOXA9* Region Reverse**	ACAAACCCCATCGTAGAGCG
***MYC* Region Forward**	GCTGCAAACTCAACGGGTAAT
***MYC* Region Reverse**	CCTCCACCACCTCCAAAAGAG
***MEIS1* Region Forward**	GTGGATGTGAGGTTTCCAGC
***MEIS1* Region Reverse**	ATATTTCTGTCCGCTGGGCTC
***SNX10* SE Region Forward**	TGCTGCCTTTCAAGCTAGTTATT
***SNX10* SE Region Reverse**	CTTGACACAGACCTCCCTTCC
Negative Control Forward	CCCCTTTTCCAACTTCCCCA
Negative Control Reverse	AACTTGGACACAGGGCATGA

2.13 Dovetail Hi-C library preparation and sequencing

A Dovetail Hi-C library was similarly prepared for each sample in a similar manner as described previously (Lieberman-Aiden et al., 2009). Briefly, for each library, chromatin was fixed in place with formaldehyde in the nucleus and then extracted. Fixed chromatin was digested with DpnII restriction enzyme, the 5' overhangs filled in with biotinylated nucleotides, and then free blunt ends were ligated. After ligation, crosslinks were reversed, and the DNA purified from protein. Purified DNA was treated to remove biotin that was not internal to ligated fragments. The DNA was then sheared to ~350 bp mean fragment size and sequencing libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adapters. Biotin-containing fragments were isolated using streptavidin beads before PCR enrichment of each library. Clinical Hi-C libraries were prepared by Dr Deepak Babu with the clinical samples obtained from the Cancer Science Institute Epigenomic core.

2.14 Topologically Associated Domain (TAD) and Chromatin Loop Calling

Hi-C data were aligned and processed by Juicer (version 1.5). The reference genome was hg38. The reads with mapping quality under 30 were discarded. TADs were called by Arrowhead at 10kb resolution with Knight-Ruiz (KR) as the normalization method. We used HiCCUPS as a loop calling software, Knight-Ruiz (KR) as the normalization method to call loops from 3 resolutions: 5kb, 10kb and 25kb and merged them to get a total loop list. 10% False Discovery Rate (FDR) was used to select significant loops. For other parameters, we retained the default settings. All clinical samples were provided by the Cancer Science Institute Epigenomic core (Ms Tng Jia Qi, Ms Fam Wee Nih and Ms Goh Yufen) and TAD and loops calling were performed by Ms Kong Lingshi.

2.15 RNA-Seq samples preparation

RNA-Seq was performed on extracted RNA (see the section on "RNA Extraction, reverse transcription and quantitative polymerase chain reaction (qPCR)") through the ribosomal rRNA-depleted RNA-Seq approach (Illumina), either 2x150-151 or 2x101 bases.

2.16 RNA-seq analysis of clinical samples

All RNA-Seq alignment was done by STAR (v2.7.3a) (Dobin et al., 2013). We normalized the raw data reads numbers input by their read length and unique mapping ratio. We randomly selected reads after we calculated the normalized read numbers, to make sure each sample has the same genome coverage (7× of the genome). UCSC genome browser tracks for RNA-Seq were also prepared by STAR (v2.7.3a) (Dobin et al., 2013). We normalized the signals by reads per million (RPM). RNA-seq analyses of clinical samples were performed by Ms Kong Lingshi.

2.17 RNA-seq analysis of publicly obtained RNA-seq Data

In addition to our RNA-seq samples, published data from 58 AML and 2 normal CD34+ samples were downloaded from the Sequence Read Archive (SRA) under the accession number SRP103200 (McKeown et al., 2017a). I performed the RNA-seq analysis of cell lines and published data via the Cancer Science Institute web portal. Briefly, the RNA-seq data are similarly aligned by STAR (v2.7.3a) (Dobin et al., 2013). The aligned reads are processed by the R package HTSeq-count (Anders et al., 2015) before being compared for their differences with the DESeq2 R package (Love et al., 2014).

2.18 Gene Set Enrichment Analysis (GSEA)

Gene set enrichment analysis were conducted on normalized RNA-seq gene counts from each sample and analyzed against the MSigDB: Hallmark gene sets via the Cancer Science Institute web portal.

2.19 ChIP-Seq & Super-Enhancer Identification

Super-enhancers (SE) were called from 58 AML patients samples and 2 FACS-purified hematopoietic stem and progenitor cell (HSPC) samples (McKeown et al., 2017a). ChIP-Seq data for these clinical AML and HSPC samples were obtained via the Sequence Read Archive (SRA) database under accession numbers SRP103200 (primary samples) and SRP103029 (cell lines) (McKeown et al., 2017a). CTCF, HOXA9, H3K27ac ChIP-Seq data for the cell lines were obtained from the Gene Expression Omnibus (GEO) database under accession numbers GSE26320 (Ernst et al., 2011), GSE123947 (Mohaghegh et al., 2019), GSE106359 (Barbieri et al., 2018), GSE79899 (Prange et al., 2017), GSE76922 (Zhang et al., 2018), GSE129380 (Burr et al., 2019) and from the European Nucleotide Archive (ENA) under the accession number SAMEA4821174.

H3K27ac ChIP-seq sequences were aligned to the human genome using Bowtie2 (Langmead and Salzberg, 2012) with the default parameters. PCR duplicates were removed using ‘samtools markdup’. Blacklisted regions that fall within the ENCODE consensus were removed using ‘bedtools intersect’. After sorting and indexing sequences with ‘samtools’, narrow peaks were called using MACS2 (version 2.1.2) (Zhang et al., 2008b). Enhancer peaks within 12.5 kb distance were stitched together and identified as super-enhancers based on the ChIP-seq

signal using a custom script similar to ROSE package as previously described (Cao et al., 2017a, Hnisz et al., 2013a).

2.20 Statistical Analyses

Statistical analyses were performed using the GraphPad Prism 8 software. One or two tailed student T-tests were performed to evaluate the significance between two different groups of samples or genes. One-way ANOVA test was applied for the evaluation of the significant differences between the gene expression of four cell lines. Significance are indicated as: ns – $P > 0.05$, * $P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P<0.0001$.

3 A Frequently Interacting Region (FIRE) at the *MEIS1* locus is heterogeneously present in Acute Myeloid Leukaemia (AML) patients.

3.1 A compendium of clinical Hi-C datasets from Acute Myeloid Leukaemia (AML) CD34+ cells & normal CD34+ cells

3D genome organization is altered in cancers such as leukaemia (Li et al., 2018, Yang et al., 2019, Kloetgen et al., 2020). For instance, chromatin interactions were shown to be altered in mixed-lineage leukaemia (*MLL*) translocated leukaemia that results in the gain-of-function fusion proteins (Rousseau et al., 2014). Using the 5C chromatin interaction frequencies that are obtained from a mixture of *MLL* translocated acute myeloid (AML), lymphoblastic leukaemia (ALL) and normal non *MLL*-rearranged samples, Rousseau et al. (2014) developed a bioinformatic algorithm to differentiate between *MLL*-rearranged and wildtype *MLL* cells at the *HOXA* locus. They further tested their algorithm with various types of MLL-rearranged leukaemias and their results show that the *MLL*-rearranged leukaemias develop significantly contrasting chromatin interaction frequencies to the wildtype *MLL* samples at the *HOXA* locus (Rousseau et al., 2014).

Myeloid leukaemia is thought to arise from precursor CD34+ haematopoietic stem cells (Shlush et al., 2014, Bonnet and Dick, 1997, Chávez-González et al., 2013), with several fusion oncogenic proteins found to be aberrantly activated and upregulated during AML pathogenesis (Iijima-Yamashita et al., 2018, Huret and Senon, 1997, Kundu et al., 2002, Wang et al., 2017). A study by Tijchon et al. (2019) showed that AML-ETO protein expression in induced pluripotent stem (IPS) cells resulted in the block of granulocytic differentiation and the expression of leukemic characteristics. These observed phenotypes resemble that of AMLs with the t(8;21) chromosomal translocation which is responsible for producing the AML-ETO fusion protein (Tijchon et al., 2019). However, even with the growing pool of information on myeloid leukaemia disease features, the differences between chromatin interactions mediated regulation of myeloid leukaemia associated oncogenes in normal CD34+ and myeloid leukaemia CD34+ cells remain largely unknown.

Here I seek to understand whether there are differences in the patterns of chromatin interactions between CD34+ cells from AML samples as compared with CD34+ haematopoietic stem cells from healthy individuals. This information will be useful in understanding whether AML pathogenesis involves chromatin interactions alterations and for the development of therapeutic treatments that target dysregulated chromatin interactions in AML.

In initial work on this subject, my lab colleagues Lingshi Kong, Deepak Babu, Yufen Goh, Fam Wee Nih, Tng Jiaqi performed Hi-C sequencing of three CD34+ AML and three normal CD34+ clinical samples (derived from bone marrows of knee replacement surgeries) (**Table 12 & 13**). The AML samples were obtained from our collaborator Professor Chng Wee Joo, from the National University Hospital, while the normal knee derived bone marrow samples were obtained from Professor Wilson Wang, Dr Ming Chun Chan, Xin Liu, Fang Fang Song and Priscella Chia from the National University Hospital.

My colleague Lingshi Kong next analysed these Hi-C results from the six clinical samples by identifying their respective topologically associating domains (TADs), chromatin loops and contact heatmap profiles. The clinical samples libraries were sequenced to sufficient depth and complexity for TAD and loop calling. **(Table 14)**. TADs and loops were identified using the Arrowhead & HiCCUPs (Rao et al., 2014) software respectively, whereas Hi-C contact heatmap visualization was done with JuiceBox (Durand et al., 2016). The above analytical packages were chosen as they are commonly used in the analysis of Hi-C data (Rao et al., 2014).

MEIS1 is suggested to be essential in the maintenance of haematopoietic CD34+ stem cells (Ariki et al., 2014), that are also the progenitors for AML development (Jan et al., 2011). Co-expression of *MEIS1* and myeloid leukaemia associated transcription factor *HOXA9* was also reported to be sufficient for the rapid induction of AML (Lawrence et al., 1999, Kroon et al., 1998). Furthermore, *MEIS1* was reported to play a role in the induction of leukemic associated genes expression such as Fms-like tyrosine kinase 3 (*FLT3*) (Wang et al., 2005). Therefore, I approached this study by first examining the *MEIS1* gene locus for potential alterations of TADs and chromatin loops. Analysis of the six clinical

Hi-C contact heatmaps at the *MEIS1* locus shows the presence of a TAD-like structure, located at the *MEIS1* gene region **(Figure 11A)**. Interestingly, I observed that this TAD-like structure is present in all three normal CD34+ samples but present in only one of the three CD34+ AML samples **(Figure 11B)**. Samples AML 28 and 30 showed the lack of this TAD-like structure as compared to AML 29. This observation is striking as the lack of this TAD-like structure at *MEIS1* in two of the three AML samples in comparison against all three normal CD34+ samples, may suggest that chromatin loops are altered at the *MEIS1* locus in AML.

Since CTCF mediated formation of TADs have been suggested to be through the mechanism of loop extrusion (Guo et al., 2015, Zuin et al., 2014), I reasoned that the TAD-like structure may be called as a loop by methods of loop identification such as HiCCUPs (Blinka et al., 2016, Rao et al., 2014). In order to examine this region further, I analyzed the loops from the contact heatmaps of these six clinical samples (**Figure 12**). My observation of the loops from the six clinical Hi-C data and alignment of the Hi-C contact heatmap from AML 29, revealed the presence of a small loop that spans the length of the *MEIS1* gene in the three normal CD34+ and AML 29 samples (**Figure 12**). This small loop within the *MEIS1* region can be predominantly found in the normal CD34+ and AML 29 CD34+ samples, that also contains the TAD-like structure. The only exception observed was the presence of this small loop in the AML 30 sample which does not possess the TAD-like structure. Furthermore, additional inspection of the loops in AML 28 and 30, show that they both contain either the loss of the small loop within the gene region or the presence of loops that extends in the up and downstream directions beyond the *MEIS1* gene location (**Figure 12**). The difference in loops observed between the AML samples could be a result of the genetic heterogeneity between them or the deferring degree of loops disrupted at the *MEIS1* locus in these samples. For example, a larger loss of the TAD-like structure and its associated loops in AML28 can be seen as compared to AML 30, which show a less degree of TAD-like structure loss, with some minor loops being retained. Taken together, the lack of the TAD-like structure is associated with the presence of dysregulated loops at the *MEIS1* locus.

Table 12. Information about AML patient samples

	AML 28	AML 29	AML 30
Gender	Male	Female	Female
Age	52yr	35yr	28yr
Total MNC count	56.3 million	106 million	98.4million
Viability	91%	72%	93%
Percentage CD34+	80.50%	74.40%	57.20%
Total CD34+ count	22 million	22.4 million	20.2 million
Viability	95%	92%	97%
Karyotype	Normal Karyotype	Trisomy 8	45,X,-X,t(8;21)(q22;q22)/46,XX
FLT3	Negative	FLT3/ITD: Positive	FLT3/ITD: Positive
Relapse	relapse	relapse	No

This table includes the clinical details of acute myeloid leukaemia isolated CD34+ samples used for Hi-C sequencing. Column one show the different characteristics of the clinical samples. A common AML associated oncogene (*FLT3*) is included within. 'MNC' refers to the isolated mononuclear cells. Column two to four lists the details of each clinical sample. "N.D." indicates not done.

Table 13. Information about normal knee patient samples

	Knee 47	Knee 49	Knee 50
Gender	Male	Female	Female
Age	72yr	67yr	74yr
Total MNC count	329 million	567 million	690.2 million
Viability	88%	93%	88%
Percentage CD34+	7.80%	10.20%	10.30%
Total CD34+ count	3.02 million	6.28 million	6.9 million
Viability	77%	90%	85%
Karyotype	N.D	N.D	N.D
FLT3	N.D	N.D	N.D
Relapse	N.D	N.D	N.D

This table includes the clinical details of normal knee bone marrow derived CD34+ samples used for Hi-C sequencing. Column one show the different characteristics of the clinical samples. A common AML associated oncogene (*FLT3*) is included within. 'MNC' refers to the isolated mononuclear cells. Column two to four lists the details of each clinical sample. "N.D." indicates not done

Table 14. Table of number of TADs & loops identified from the six clinical CD34+ AML and normal samples

Sample	Reads	#TADS	#Loops
AML28	1,290,109,443	1,682	24,394
AML29	1,245,251,169	1,108	19,541
AML30	1,387,582,168	1,296	10,733
Knee 47	1,337,973,222	1,153	4,733
Knee 49	1,156,332,234	927	13,107
Knee 50	1,165,296,399	1,234	19,795

Column one indicates the sample labels, column two to four indicates the number of sequenced reads, identified topologically associating domains (TADs) and chromatin loops.

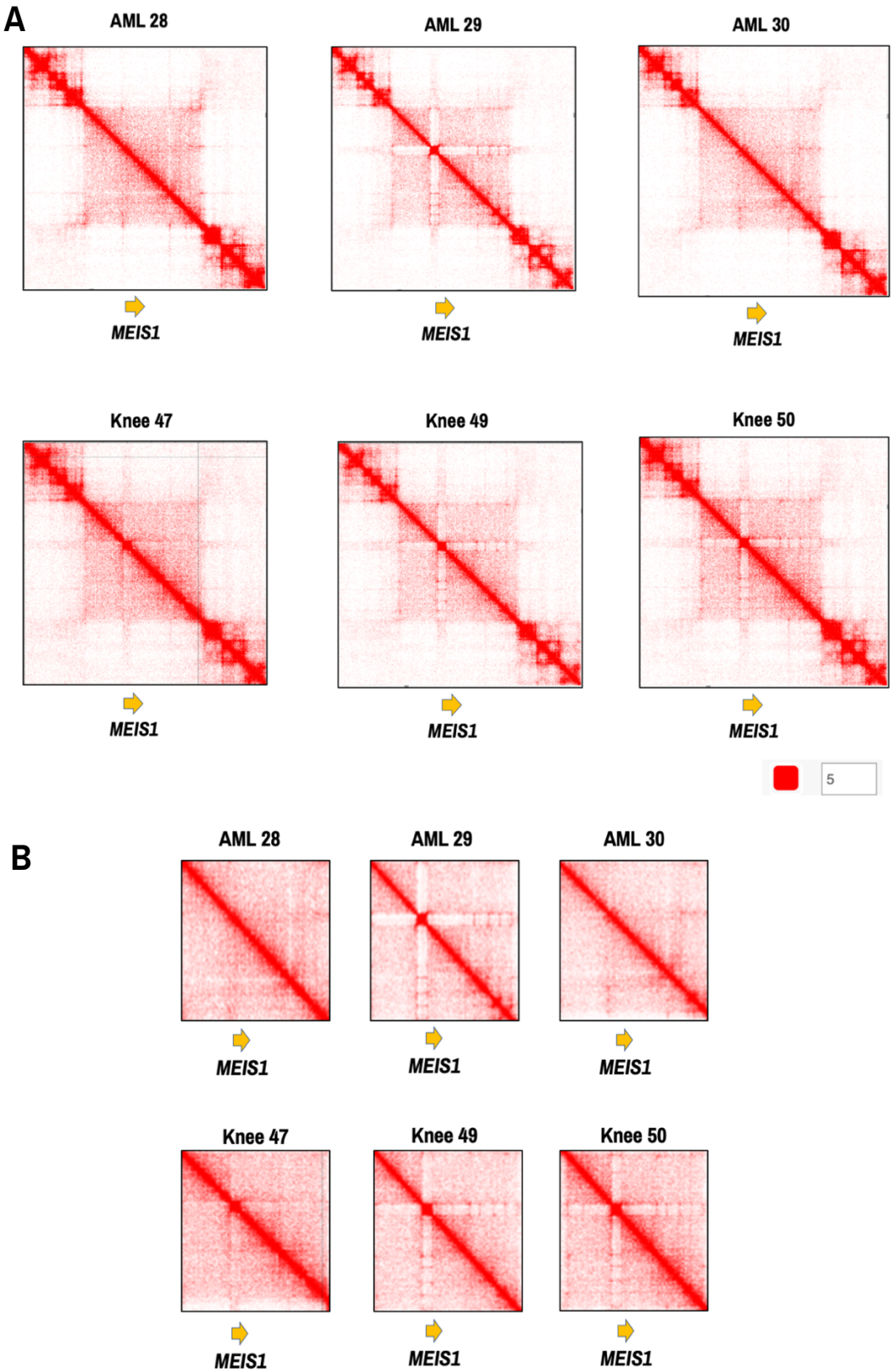

Figure 11.View of chromatin interactions at the *MEIS1* locus in six clinical samples A. Hi-C contact maps from six clinical CD34+ samples (three AML 28-30 & three normal knees 47,49-50) at the *MEIS1* at region. B. Zoomed in visual of the loops at *MEIS1* locus. The yellow arrow represents the location of the *MEIS1* gene.

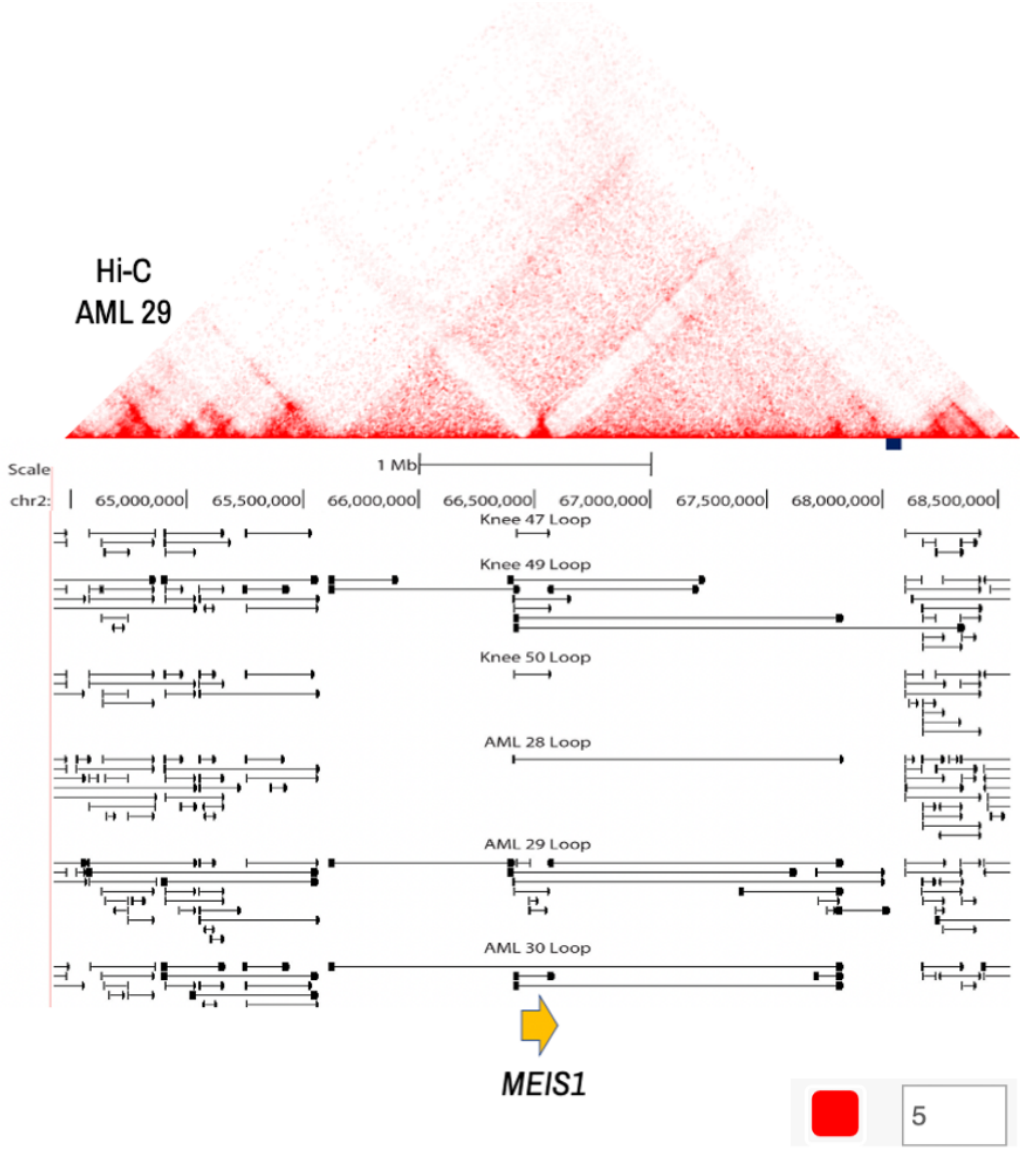

Figure 12. Results of Hi-C analyses from six clinical CD34+ samples (three AML 28, 29, 30 & three normal knees 47, 49, 50). HI-C loops from six clinical samples that are aligned to the AML 29 Hi-C contact heatmap at *MEIS1*. Scale of contact maps interaction is indicated as a numerical value beside the colour intensity. (Hi-C samples preparation by Ms, Fam Wee Nih, Ms Tng Jiaqi and Dr Deepak Babu and Hi-C analyses conducted by Ms Kong Lingshi).

3.2 Absence of TAD-like interactions in clinical AML patients is associated with low *MEIS1* levels

MEIS1 was reported to be involved in the maintenance of hematopoiesis in blood stem cells and is downregulated upon blood cell maturation (Collins and Hess, 2016a, Kumar et al., 2010). However, *MEIS1* expression was found to be continuously upregulated in groups of myeloid leukaemia patients, in particular those with the presence of *MLL* rearrangements (Zeisig et al., 2004), suggesting its role in supporting AML progression. To investigate if such expression pattern is similar in our clinical samples, I first performed droplet digital polymerase-chain-reaction (ddPCR) on 12 AML and 10 normal CD34+ patient samples to determine if their *MEIS1* expression levels differ **(Figure 13A)**. ddPCR was used as it is a method that detects for the absolute number of gene copies within a sample and therefore represents a more sensitive and accurate method for the quantification of gene expression in clinical samples as compared to the qPCR method. My ddPCR results show that *MEIS1* expression levels between AML samples were also of varying levels with some patients expressing lower levels of *MEIS1* as compared to the rest. Although *MEIS1* expression levels varied between each AML samples, the overall comparison shows that *MEIS1* levels were not significantly different between most of the AML and normal CD34+ groups, thus suggesting that *MEIS1* levels continue to be expressed or upregulated in AML. These suggest that mechanisms may be at play to enable the continuous high expression of *MEIS1* levels in AML cells.

To further examine the potential effects of the TAD-like structure on *MEIS1* expression, I performed ddPCR on another six normal CD34+ and AML 28, 29 & 30 CD34+ samples **(Figure 13B)**. In addition, I also conducted RNA-seq analyses with two normal CD34+ and AML 28, 29 & 30 CD34+ samples to validate the results of *MEIS1* expression. My ddPCR and RNA-seq results show that *MEIS1* expression levels were strongly downregulated in both AML 28 and 30, as compared to AML 29 and the normal CD34+ samples which presented higher expression levels **(Figure 13C)**. The comparison of the results from ddPCR, RNA-seq and Hi-C contact heatmaps from the AML and normal CD34+ samples, suggests that patients with the presence of the unique TAD-like

interaction at the *MEIS1* locus express higher *MEIS1* levels as compared to those without.

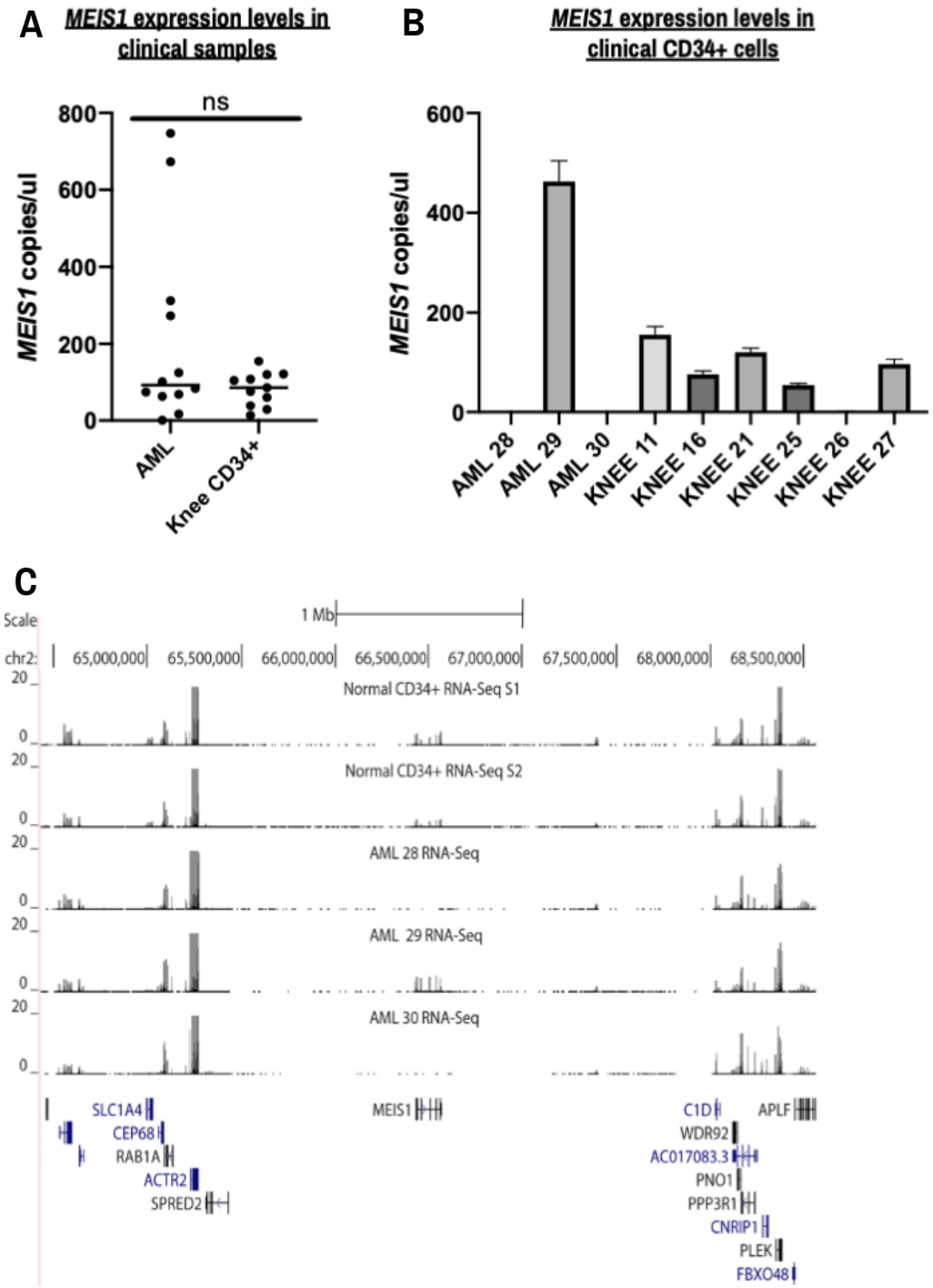

Figure 13. *MEIS1* expression levels in AML CD34+ and normal CD34+ cells. A. ddPCR of *MEIS1* expression levels in AML and normal CD34+ knee samples. Two-tailed t-test was performed on both groups of samples. B. ddPCR of *MEIS1* expression levels of nine CD34+ AML and CD34+ knee samples. Standard error of three technical replicates were calculated and indicated as bars. C. RNA-seq

expressions level of *MEIS1* in two normal CD34+ knee and three CD34+ AML samples. (Clinical samples were obtained from Ms, Fam Wee Nih, Ms Tng Jiaqi).

3.3 THP-1 & K562 are suitable cell line models for the study of chromatin interactions at the *MEIS1* locus

In order to determine the suitability of AML cell lines as models for the study of *MEIS1* regulation, I analyzed the Hi-C contact heatmaps of K562, THP-1 and GM12878 at the *MEIS1* locus where TAD-like structures are located. The GM12878 non-cancer B cell lymphoblastoid cell line was chosen as a control against the other two AML cell lines. Since there are presently no normal CD34+ or normal myeloid lineage cell lines available, I decided to use GM12878 as a control as it is a non-cancer cell line of blood lineage that originates from the same CD34+ hematopoietic stem cells as THP-1 and HL-60. My analyses of Hi-C contact heatmaps of K562 and THP-1 show the presence of the TAD-like interaction at the *MEIS1* gene region, which on the other hand appears to be absent in GM12878 **(Figure 14A)**.

Since the analyses of my clinical data suggest the potential association of *MEIS1* expression with the TAD-like interaction at the *MEIS1* locus, I performed ddPCR on four blood cell lines, K562, THP-1, HL-60 and GM12878 to evaluate their *MEIS1* expression levels and their suitability as a cell line model in this study. My ddPCR results show that *MEIS1* expression is specifically highly expressed in THP-1 and K562 cell lines, but absent or poorly expressed in HL-60 and GM12878 (**Figure 14B**). These trends of Hi-C and ddPCR results show consistency as compared to both the clinical AML CD34+ and normal CD34+ samples. Altogether, I observed that the presence of the TAD-like interaction at *MEIS1* locus is associated with higher *MEIS1* expression. Since K562 and THP-1 are myeloid leukaemia derived cell lines, while GM12878 belongs to a non-cancer B cell lineage, I speculate that the lower expression of *MEIS1* in GM12878 may be related to its cell maturation process that may otherwise be blocked in myeloid leukaemia cells like K562 and THP-1.

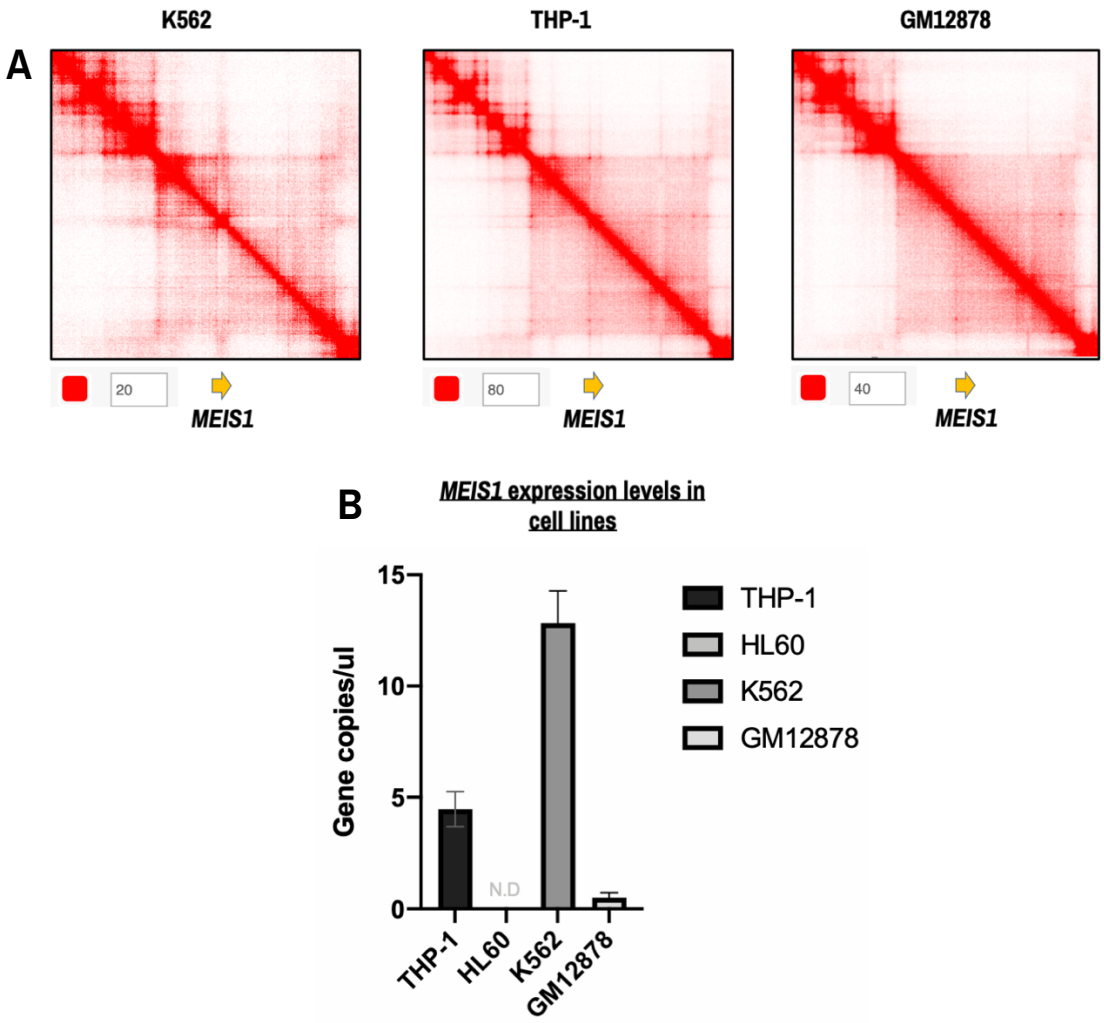

Figure 14. Hi-C contact map and ddPCR of *MEIS1* expression show that THP-1 & K562 are suitable cell line models. A. Hi-C contact map of blood cell lines K562, THP-1 and GM12878 at the MEIS1 locus. Scale of contact maps interaction is indicated as a numerical value beside the colour intensity. B. ddPCR of *MEIS1* expression levels in four blood cell lines.

3.4 Circularized chromosome conformation capture (4C) of THP-1 reveals multiple *MEIS1* promoter associated chromatin interactions within the *MEIS1* locus

To investigate the TAD-like interaction at *MEIS1*, I performed circularized chromosome conformation capture (4C) on the AML cell line THP-1, with the *MEIS1* promoter as the 4C bait. I then aligned the Hi-C contact heatmap with the 4C results for THP-1 at the *MEIS1* locus (**Figure 15**). From my 4C results of THP-1, I observed that strong chromatin loops are associated with the *MEIS1* promoter at the *MEIS1* locus. This pattern of chromatin interactions agrees with the TAD-like region observed in the THP-1 Hi-C contact heatmap. However, these chromatin interactions are not confined within the TAD-like region at the *MEIS1* gene but also occurs both upstream and downstream of the gene location (**Figure 15**). Furthermore, I also performed 4C on the K562 cell line, with the *MEIS1* promoter as the bait. The analysis of my 4C results for K562 show similar patterns of chromatin interactions that are strongest at the *MEIS1* gene with spreading towards both upstream and downstream of *MEIS1* (data not shown). Moreover, I observed the presence of CTCF binding sites within the region of the TAD-like structure as well as those of the *MEIS1* promoter-associated chromatin loops. Thus, these suggest that the chromatin interactions observed at the *MEIS1* locus may not be a TAD, but instead resemble a recently reported genomic sub-TAD structure termed 'Frequently Interacting Region" (FIRE) (Schmitt et al., 2016b).

FIRE is described as areas of unusually high level of chromatin interactions. It has the following characteristics of being expressed near cell-type specific genes, enriched in super-enhancers and to be partially dependent on CTCF and cohesin complex (Schmitt et al., 2016b). Upon further analysis of my 4C results as well as the clinical Hi-C data, I realised the characteristics of the chromatin interactions seen at *MEIS1* matches that of a FIRE. Based on my observations, there is the presence of a dense cluster of chromatin interactions at *MEIS1,* which is not confined to a region as usually seen with TADs. These interactions occurring at the *MEIS1* region is consistent with previous reports that FIREs occur near cell-type specific genes. As *MEIS1* is a cell-type specific gene that is widely reported to be involved in blood stem cell haematopoiesis and

downregulated upon cell maturation (Argiropoulos et al., 2007, Imamura et al., 2002, Pineault et al., 2002), the presence of this TAD-like structure at its locus and the results of my 4C data at the *MEIS1* locus, strongly suggests that the TAD-like structure seen at the *MEIS1* gene is likely to be a FIRE.

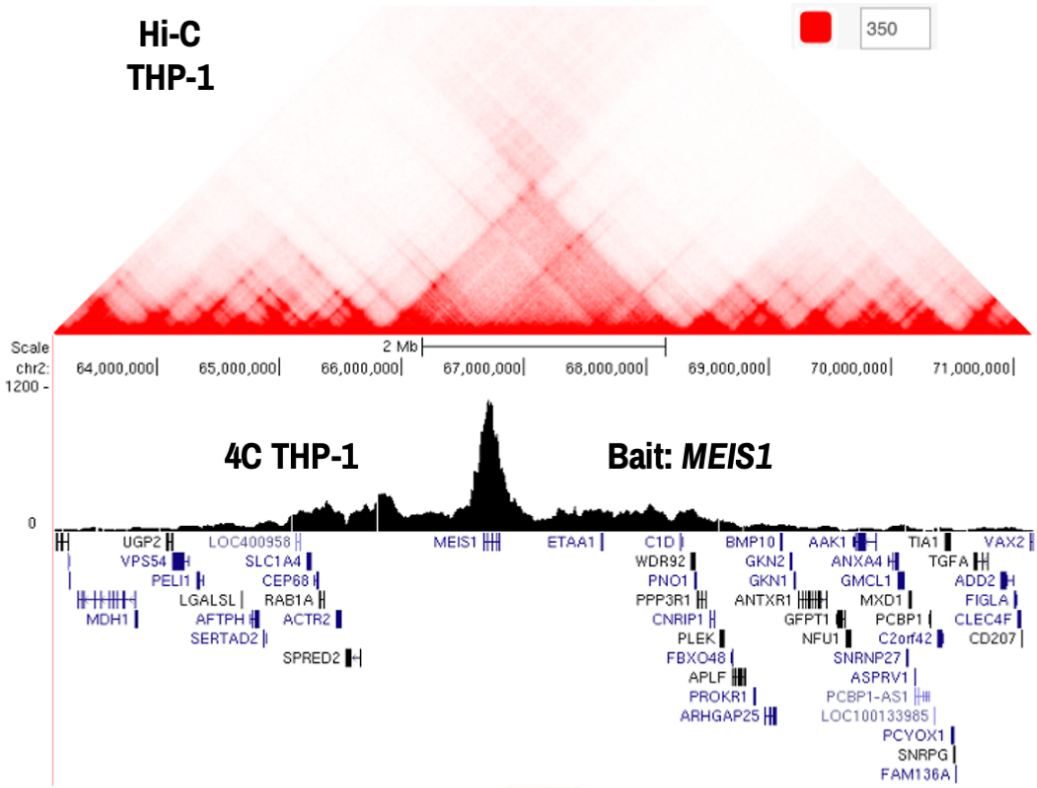

Figure 15. THP-1 Hi-C & 4C show the presence of a 'Frequently Interaction Region' (FIRE) at the *MEIS1* gene. THP-1 Hi-C contact heatmap is aligned with the 4C chromatin interactions that are associated with the *MEIS1* promoter.

3.5 Super-enhancers profiles from AML patients show occupancy at regions that interacts with *MEIS1* promoter

To further validate if a FIRE exists at the *MEIS1* locus, I analyzed H3K27 histone acetylation (H3K27ac) chromatin immunoprecipitation-sequencing (ChIP-seq) data of published cell lines (THP-1, HL-60, K562, GM12878) data (Burr et al., 2019, Mohaghegh et al., 2019, Barbieri et al., 2018, Zhang et al., 2018) and 60 clinical samples from an earlier study (Two normal CD34+ & 58 AML) (McKeown et al., 2017b).

First, I downloaded the raw FASTQ files of the H3K27ac and their corresponding input ChIP-seq files for the cell lines, THP-1, HL-60, K562 and GM12878 and for each of the 60 clinical samples from the Sequence Read Archive (SRA) database. In this case, it is important to note that the AML samples analyzed were not being enriched for CD34+ cells but taken as a total population of cells from each AML patient. I then mapped the H3K27ac ChIP-seq data of the four blood cell lines and 60 clinical samples to the human genome (hg19) using the BWA-MEM package (Li, 2013), before identifying the H3K27ac peaks with the bioinformatics package MACS2 (Zhang et al., 2008a). Since super-enhancers are made up of clusters of enhancers in proximity (Whyte et al., 2013), I analysed the peaks of H3K27ac according to the previously reported stitching distance of 2.5 kilobases and ranked them using the ROSE (ranking of super-enhancers) package (Cao et al., 2017b, Hnisz et al., 2013a).

Overall, my analyses of the three cell lines and 60 clinical samples show that super-enhancers are present at several locations around the *MEIS1* locus **(Figure 16**). Interestingly, I observed three super-enhancers (SE1,2 &3) to be present at the locations that interact with the *MEIS1* promoter in THP-1 (**Figure 16**). The number of patient samples that possess each super-enhancer is indicated in the boxes at the bottom of the tracks. Four cell lines - THP-1, HL-60, K562 and GM12878 show the presence of super-enhancers 1 & 2 with slight differences between them (**Figure 16**). I also observed that the AML cell lines THP-1, HL-60 and the non-cancerous GM12878 cell line showed the presence of super-enhancers 1 and 3 while, the chronic myeloid leukaemia cell line K562 showed the presence of super-enhancer 1 and 2. My analysis also revealed that super-

enhancer 1 is present in all four cell lines as well as in 54 out of 58 AML patient samples analyzed. These differences observed in each of the cell lines as well as those of the AML patient samples may be attributed to the fact that each of these samples belongs to a different subset of AML. However, more subtype-specific samples have to be analyzed for this to be concluded.

Since my analyses from the Hi-C contact heatmap and ddPCR of GM12878 show less defined TAD at the *MEIS1* gene region and low expression of *MEIS1* (**Figure 13**), the unchanged super-enhancer presence between THP-1 and GM12878 cell lines at the *MEIS1* locus suggest that the alteration of *MEIS1* promoter associated chromatin interactions affect *MEIS1* expression in AML. Furthermore, the presence of three clusters of super-enhancers (SE1,2&3) support my earlier findings that suggest that a FIRE occupies the *MEIS1* locus. Taken together, these findings suggest that the proposed *MEIS1* FIRE plays a contributive role by facilitating chromatin interactions between super-enhancers and the *MEIS1* promoter and *MEIS1* expression in AML (**Figure 17**).

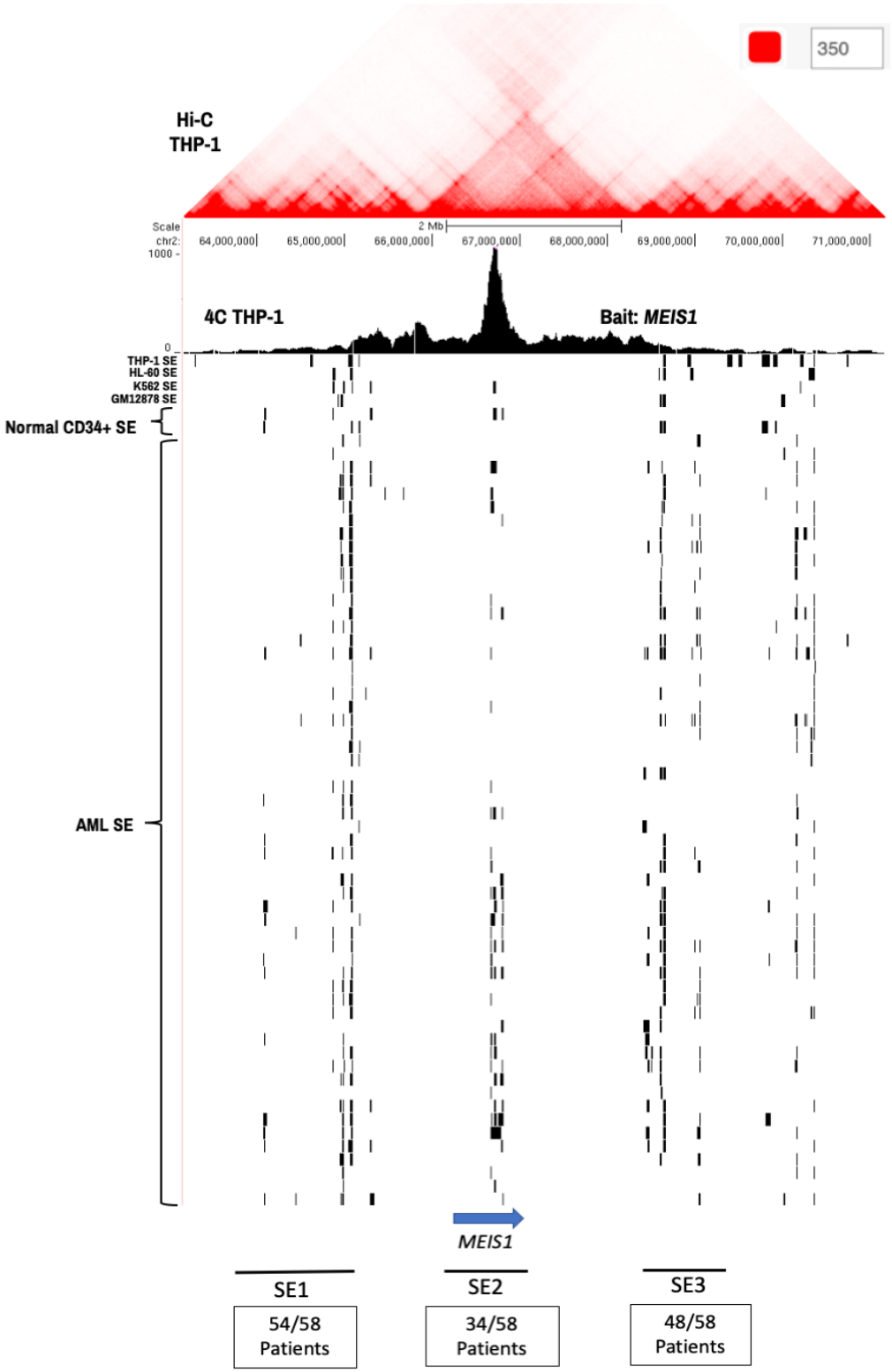

Figure 16. Three groups of super-enhancers are identified at the *MEIS1* FIRE region. THP-1 Hi-C contact map is aligned to THP-1 4C interactions with *MEIS1* as the bait. Super-enhancers (SE)s in blood cell lines THP-1, HL-60, K562, GM12878 and 60 clinical samples (Two normal CD34+ & 58 AML) (McKeown et al., 2017b) are aligned to Hi-C and 4C tracks. Three groups of super-enhancers and the number of AML patients possessing each of them are shown in boxes below. Colour intensity of heatmap is indicate as a value at the bottom left-hand corner.

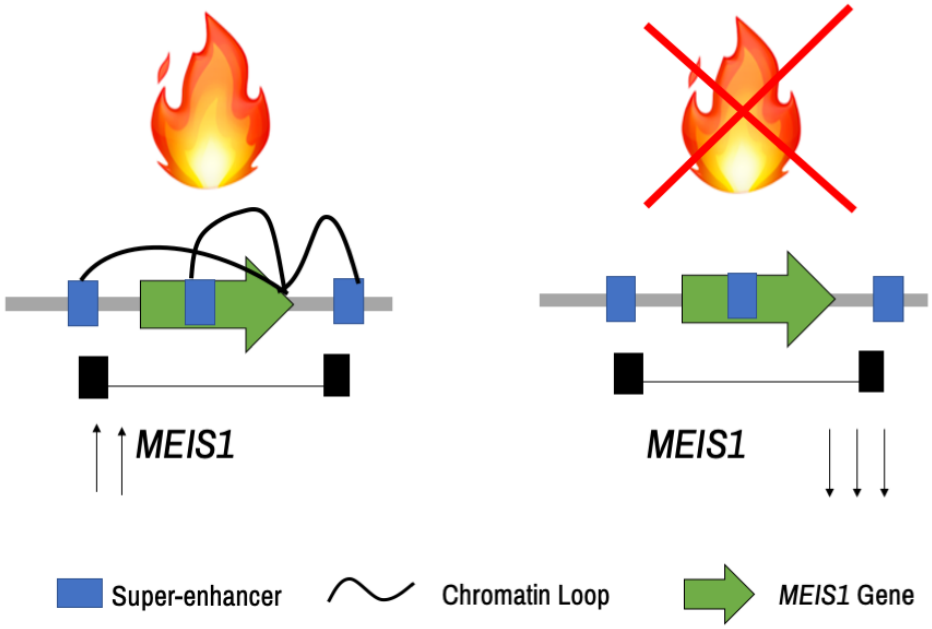

Figure 17. A Schematic of proposed relationship between super-enhancers in the *MEIS1* FIRE and *MEIS1* expression in AML. The green arrow represents the location of the *MEIS1* gene. The rectangular blue box indicates super-enhancers with black lines referring to chromatin loops. The black boxes and icon of a FIRE represent the location of the *MEIS1* FIRE. Arrows indicate the up or down regulation of *MEIS1* expression.

4 Maintenance of a CTCF binding site at the *MEIS1* FIRE boundary is required for *MEIS1* expression and cell growth in the K562 leukaemia cell line

4.1.1 CTCF-chromatin immunoprecipitation sequencing (ChIP-seq) analysis of publicly available data reveals the presence of CTCF proteins bound upstream and downstream of the *MEIS1* gene.

After validating the presence of the *MEIS1* FIRE in K562 and THP-1 cell lines, I seek to investigate the effects of the *MEIS1* FIRE on *MEIS1* expression in myeloid leukaemia. I set forth to design a CRISPR-Cas9 experiment to evaluate if deletion of CTCF boundaries will have an impact on *MEIS1* FIRE regulation and *MEIS1* expression. Based on my examination of normal Knee 50 CD34+ and AML 28 CD34+ samples, the chromatin loop at the *MEIS1* gene region in AML 28 appeared to be extended beyond the 3' end of the *MEIS1* gene which also fits in with the Hi-C heatmap of AML 28 that exhibited a loss of the *MEIS1* FIRE (**Figure 18A**). This alterations in chromatin loops and FIRE presence at the *MEIS1* FIRE of AML 28 as compared to the Knee CD34+ samples imply that *MEIS1* regulation via chromatin interaction may be different in AML 28 than the normal Knee CD34+ samples.

Since my prior analysis of Knee 50 and AML 28 contact heatmaps showed loss of *MEIS1* FIRE in AML 28, I hypothesize that the loss of CTCF downstream of *MEIS1* gene could have led to the formation of distal loops as observed in AML 28. To investigate this hypothesis, I first downloaded the raw FASTQ data of the CTCF ChIP-seq for the K562 cell line from the Sequence Read Archive (SRA) database. Subsequently, I mapped the CTCF ChIP-seq data against the reference human genome (hg19) using the BWA-MEM package (Li, 2013), before identifying the peaks of CTCF binding using the MACS2 package (Zhang et al., 2008a). My CTCF ChIP-seq analysis data (Ernst et al., 2011) revealed that CTCF proteins bind to proximal regions at the 5' and 3' end of the *MEIS1* gene which also contains nearby CpG islands (**Figure 15A**). These suggest that CTCF proteins that are bound to CTCF sites upstream and downstream of *MEIS1* are present in the *MEIS1* FIRE.

Since the formation of the distal loop was seen in AML 28 and 30, I designed a CRISPR-Cas9 experiment to target the specific CTCF binding site downstream of the *MEIS1* gene (**Figure 18A**). Based on my observations of the Hi-C contact heatmaps, the K562 cell line contains the *MEIS1* FIRE that is seen in AML 29 and the respective normal knee CD34+ samples as well as expresses high levels of *MEIS1*, therefore it was chosen as a suitable cell line model for this CRISPR experiment (**Figure 18B**).

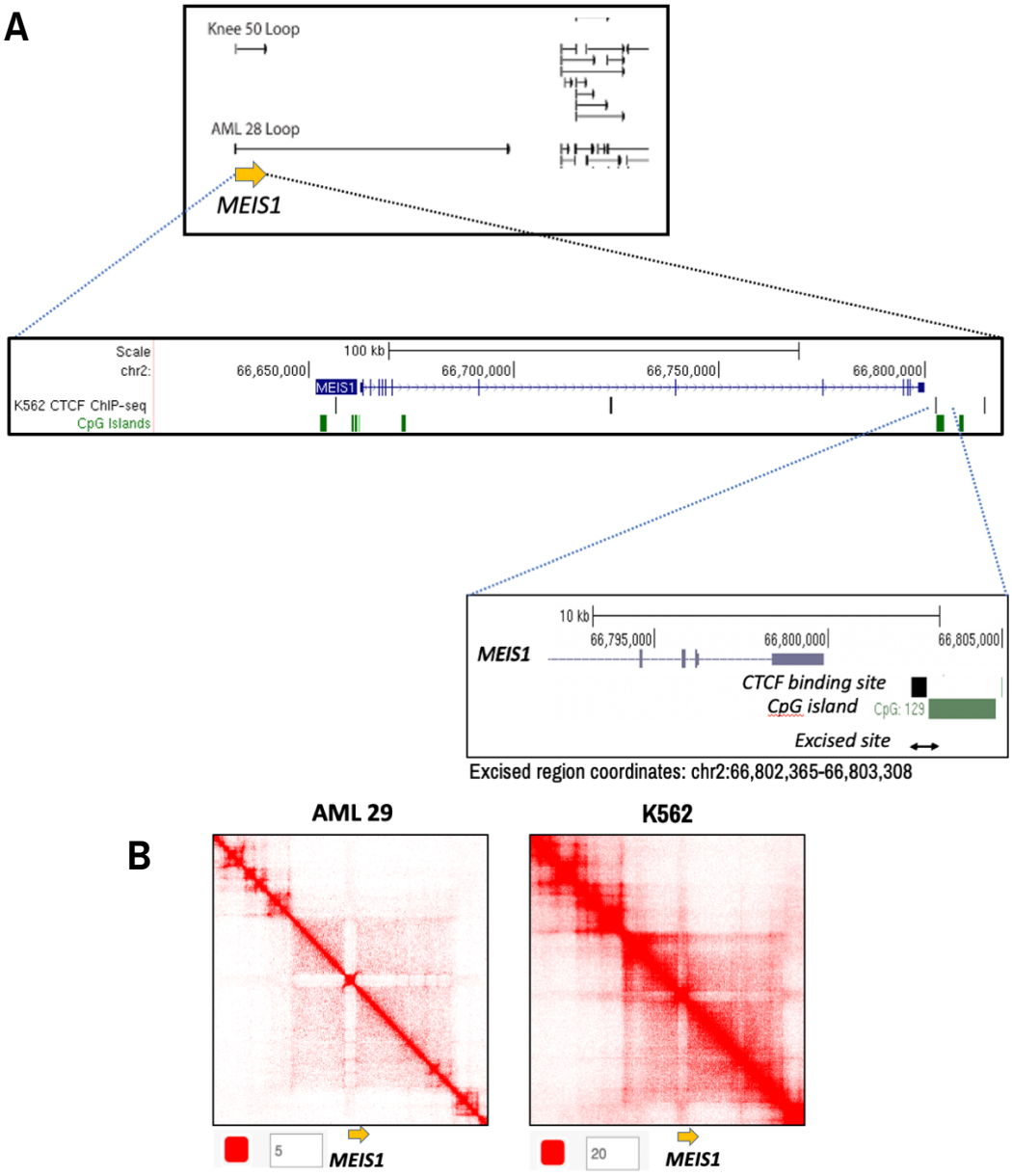

Figure 18. CTCF boundaries are present upstream and downstream of the *MEIS1* gene region. A. Hi-C loops of the normal Knee50 CD34+ cells and AML 28 CD34+ cells. Zoomed in view of CTCF binding sites upstream and downstream of the of the *MEIS1* gene. Zoomed in view of a strong CTCF binding site with a CpG island downstream of the *MEIS1* gene. CTCF ChIP-seq track

show the average of two biological-replicates from K562 (Ernst et al., 2011). Location for the excision of the downstream CTCF binding is indicated with a bidirectional arrow (chr2:66,802,365-66,803,308). B. Hi-C contact heatmap of AML 29 and K562 cell lines at the *MEIS1* region.

4.1.2 CRISPR-Cas9 experiment shows successful deletion of the CTCF binding region that is downstream of *MEIS1*

CRISPR-Cas9 deletion of the selected CTCF boundary that is downstream of the *MEIS1* gene showed successful deletion of the region in three clones of K562 cells (**Figure 19**). In addition, I also generated three clones of Cas9 empty vectors from K562 cells as controls to the CTCF excised clones. My gel images of the PCR products from the flanking regions of the deletion sites show that the CTCF boundary was successfully excised as compared to the empty vector (EV) clones (**Figure 19**). Moreover, the length of the deleted region is consistent with what we expect in a successfully excised clone.

I also designed primers within the location of the CTCF excised boundary to check for accurate excision location in each clone. From my gel images of the internal primers derived PCR products, I observed that the internal region where the CTCF boundary is located, are not present in all three CTCF site excised clones but are present within the EV control controls (**Figure 19**). This further indicates accurate CRISPR excision at the CTCF boundary.

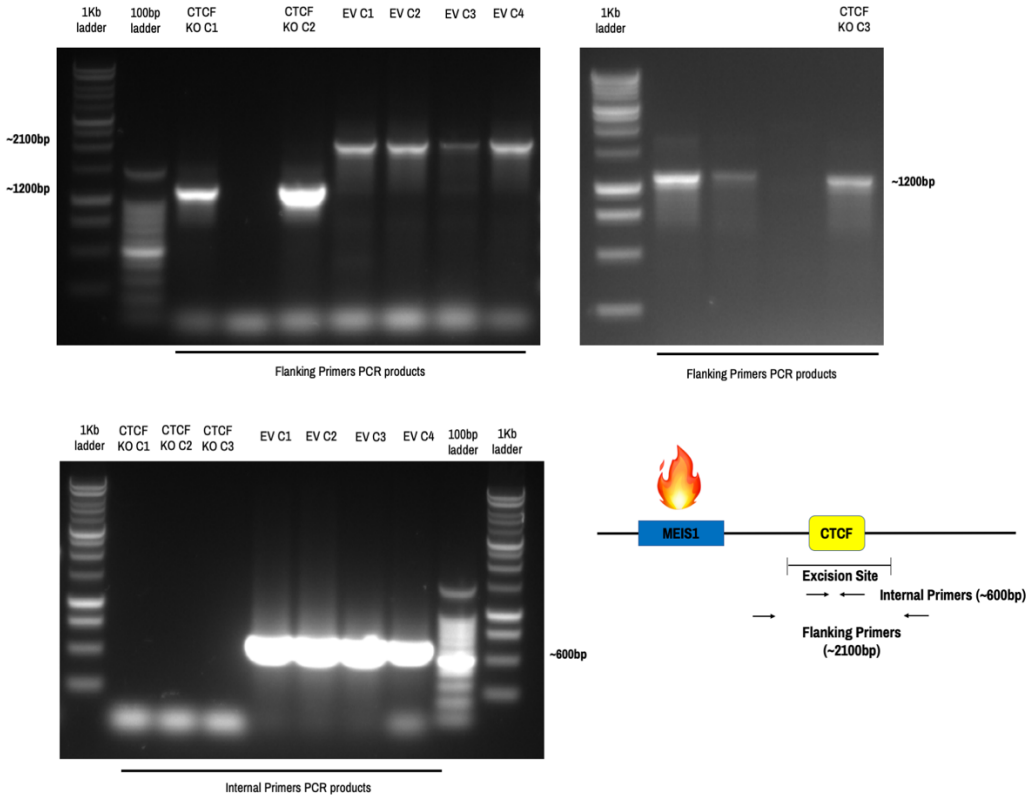

Figure 19. CRISPR excised K562 clones in *MEIS1* FIRE show loss of excised CTCF binding region. Each column is indicated at the top with the sample label. 1kb and100bp ladders are indicated at the left column of each gel image. CTCF KO represents the CTCF binding site that is excised from K562 cells. EV represents the empty vector K562 cells. The expected sizes of the bands are indicated at the side of each gel image. A schematic of the CTCF binding site excision location and the location of the internal and flanking primers with their estimated sizes is shown on the bottom right hand of the figure.

4.2 CRISPR-Cas9 excision of the CTCF binding site located downstream of *MEIS1* resulted in reorganization of chromatin loops within the FIRE.

Upon successful CRISPR deletion of the CTCF binding site that is located downstream of the *MEIS1* gene in K562 cells, I then conducted 4C, RT-qPCR, RNA-seq and growth assay experiments to evaluate the effects of the loss of the putative FIRE CTCF binding site on chromatin loops, gene expression and cell growth kinetics. CTCF deleted cells showed dramatic rewiring of *MEIS1* promoter-associated chromatin interactions **(Figure 20)**. I also observed that chromatin loops to the *MEIS1* promoter were clearly reduced in the CTCF deleted cells as compared to the empty vector (EV) cells. Interestingly, I saw that the loss of these chromatin loops occurs in the region of the *MEIS1* gene, where a super-enhancer has also been identified to be present in the K562 cell line and a subset of AML patients **(Figure 20)**. The loss of *MEIS1* promoter-super-enhancers interaction in CTCF binding site deleted cells suggest that the CTCF site that is located downstream of the *MEIS1* gene may play a role in regulating the super-enhancer to *MEIS1* promoter interactions.

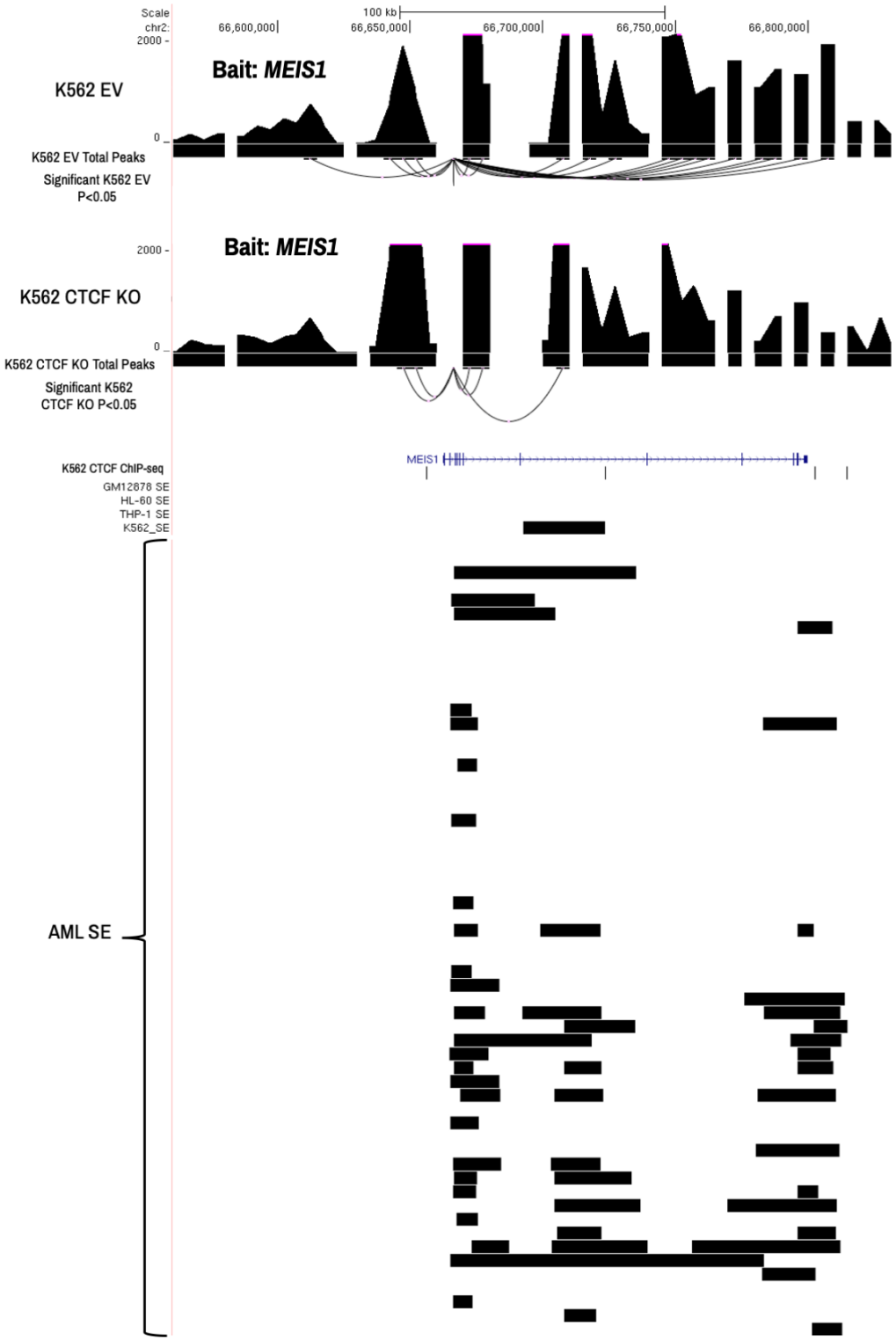

Figure 20. CTCF binding site deletion leads to the loss of chromatin interactions at the *MEIS1* FIRE. 4C tracks of K562 empty vector (EV) and CTCF binding site excised cells, with the *MEIS1* promoter as the bait. Total peaks tracks represent the view of total chromatin interactions in the region. Significant P<0.05 tracks represent the interactions with P <0.05 significance. CTCF ChIP-seq track of the average of two biological-replicates from K562 is shown at the *MEIS1* gene location. Super-enhancer tracks from K562, GM12878, HL-60 & THP-1 cell lines and 58 AML patients are aligned to the *MEIS1* FIRE.

4.3 CTCF binding site excision leads to the loss of *MEIS1* promoter interactions with genomic regions that show the presence of super-enhancers in AML clinical samples

Further inspection of chromatin loops revealed additional changes in both the upstream and downstream interactions to the *MEIS1* FIRE (**Figure 21**). This was unexpected as I hypothesized that the CRISPR excision of the downstream CTCF binding site to only affect loops downstream of the *MEIS1* gene. Based on other published studies, the removal of a CTCF boundary on one end of a TAD results in the formation of new chromatin interactions between the super-enhancers and a gene in the adjacent TAD to the removed CTCF bind site (Nora et al., 2017, Flavahan et al., 2019). Therefore, my observation of chromatin loop changes at both the upstream and downstream regions of the *MEIS1* FIRE led me to question if the loss of CTCF binding to the downstream region of *MEIS1* may have caused a cascade of chromatin interaction changes at the locus. I speculated that the loss of a CTCF downstream of *MEIS1* may have resulted in alteration and reorganization of chromatin interactions that may otherwise be kept intact by the *MEIS1* FIRE. This results also implies that the CTCF regulates TAD and FIREs in a slightly different manner as the loss of the downstream CTCF binding site to the *MEIS1* FIRE did not only altered the chromatin loops to the adjacent genomic areas to the CTCF binding site that was removed as was usually seen in CTCF removal in TADs.

Alignment of super-enhancer tracks from THP-1, HL60, K562, GM12878 and 58 AML patients showed significant loss of *MEIS1* promoter-associated interactions at the regions where three super-enhancers (SE1,2 &3) have been identified in both the blood cell lines and AML patients (**Figure 21**). Since the loss of *MEIS1* FIRE was observed to be associated with low *MEIS1* expression levels, I speculate that the deletion of CTCF boundaries within the *MEIS1* FIRE led to the loss of interactions between the *MEIS1* promoter and the super-enhancers in that region and consequently downregulate *MEIS1* expression levels in K562 cells. This implies that the CTCF binding site is essential for the maintenance of chromatin interactions in the FIRE and between the FIRE and other genomic regions.

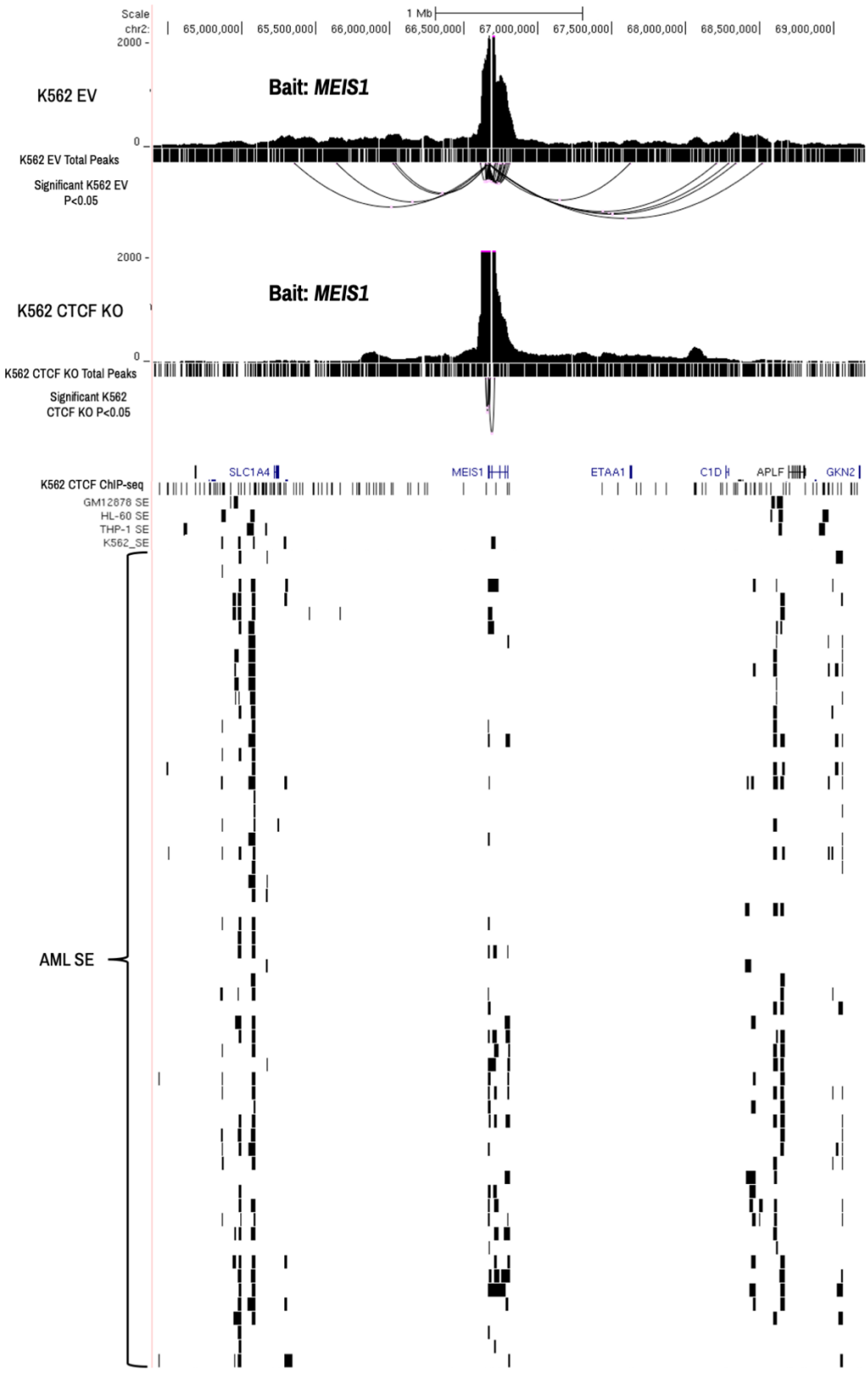

Figure 21. CTCF binding site excision leads to the loss of chromatin interactions between promoter and super-enhancers in the *MEIS1* FIRE. 4C tracks of K562 empty vector (EV) and CTCF binding site excision, with the *MEIS1* promoter as the bait. Total peaks tracks represent the view of total chromatin interactions in the region. Significant P<0.05 tracks represent the

interactions with P <0.05 significance. CTCF ChIP-seq track of the average of two biological-replicates from K562 is shown at the *MEIS1* gene location. Super-enhancer tracks from K562, GM12878, HL-60 & THP-1 cell lines and 58 AML patients are aligned to the *MEIS1* FIRE (McKeown et al., 2017b).

4.4 CTCF binding site deletion results in chromatin loop alterations at the *MEIS1* FIRE and its distal interacting sites

Following this, I surveyed the distal interacting sites to the *MEIS1* promoter by comparing the chromatin loops in the K562 empty vector control and CTCF binding site deleted cells (**Figure 22)**. Intriguingly, I saw large regions of altered chromatin interactions at the distal sites that are located upstream and downstream of the *MEIS1* FIRE (**Figure 22**). The loss of distal interactions is more apparent downstream of the *MEIS1* FIRE as compared to its upstream interaction. These loss of distal chromatin loops highlighted that the loss of the CTCF binding site within the *MEIS1* FIRE has an effect on *MEIS1* promoter associated chromatin interactions and potentially leads to alterations of gene transcription levels in the region with chromatin loop losses.

I also observed some gain of interactions, but these occur in fewer distal interacting sites of *MEIS1* (**Figure 22**). This could also signify that while interactions such as *MEIS1* promoter to super-enhancers loops may be lost, new loops with regions that were previously not accessible due to the *MEIS1* FIRE may now be formed through interactions with the *MEIS1* promoter. Collectively, these alterations in chromatin loops hint that the loss of the CTCF binding site may result in the disruption of the *MEIS1* FIRE and reorganization of chromatin loops. Since CTCF is reported to be a feature of FIREs (Schmitt et al., 2016b), I speculate that the presence of CTCF proteins in the FIRE could be involved in the confining of chromatin interactions within the *MEIS1* FIRE and that loss of the CTCF site led to the loss of CTCF binding and the abrogation of this confined space, resulting in the reorganization of loops. I also speculate that the alterations of chromatin loops that were originally associated with the *MEIS1* promoter will consequently lead to the promotion of new chromatin loops formation between the super-enhancers and other neighbouring genes. Taken together, the conservation of the *MEIS1* FIRE may be important in ensuring the transcription of *MEIS1* and the other genes within the *MEIS1* locus. In general, since FIRE are

observed to be located at cell-type specific genes, I also speculate that FIREs may serve to regulate the transcription of these cell-type specific gene via chromatin interactions.

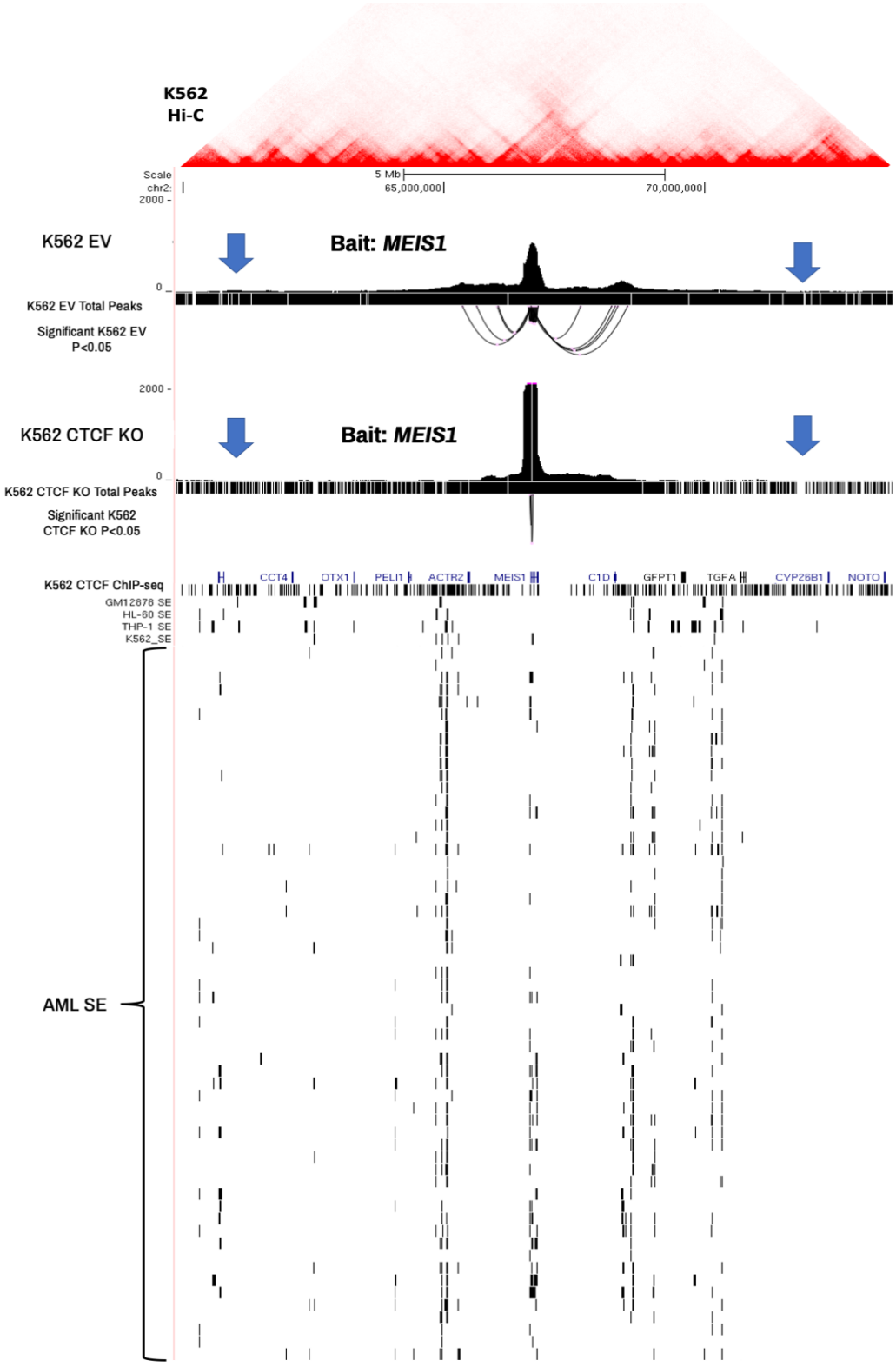

Figure 22. CTCF site deletion at *MEIS1* results in chromatin alterations at distal interacting sites to the *MEIS1* promoter. Hi-C contact heatmap of K562 is aligned with 4C tracks of the control and CTCF deleted K562 cells. 4C tracks of K562 empty vector (EV) and CTCF binding site excised cells, with the *MEIS1* promoter as the 4C bait. Total peaks tracks represent the view of total chromatin interactions in the region. Significant P<0.05 tracks represent the interactions

with P <0.05 significance. CTCF ChIP-seq track of the average of two biological-replicates from K562 is shown at the *MEIS1* gene location. Super-enhancer tracks from K562, GM12878, HL-60 & THP-1 cell lines and 58 AML patients are aligned to the *MEIS1* FIRE.

4.5 CTCF binding site excision downstream of *MEIS1* leads to the downregulation of *MEIS1* and *MYC* expression levels.

Next, I performed RT-qPCR on the CTCF binding site deleted cells to assess the impact of its binding site excision on associated myeloid leukaemia oncogene expressions. The RT-qPCR of oncogenes such as *MEIS1* and *MYC* collectively showed significant downregulation in expression in the CTCF binding site deleted cells (**Figure 23A**). The loss of *MEIS1* fire in CTCF binding site excised K562 cells led to the abrogation or downregulation of *MEIS1* expression levels, while its *MYC* expression is also seen to be lower as compared to the K562 EV control cells. This downregulation of *MYC* expression levels is likely due to the reduction of *MEIS1* expression. This is because the *MYC* promoter has been reported to be a downstream target of MEIS1 protein (Bessa et al., 2008).

Furthermore, my Gene Set Enrichment Analysis (GSEA) of RNA-seq results from the CTCF binding site deleted K562 and EV (empty vector) cells, also showed the downregulation of downstream MYC protein targeted pathways. This result supports my RT-qPCR findings which presented the downregulation of *MYC* expression levels after the CTCF binding site was excised (**Figure 23B**).

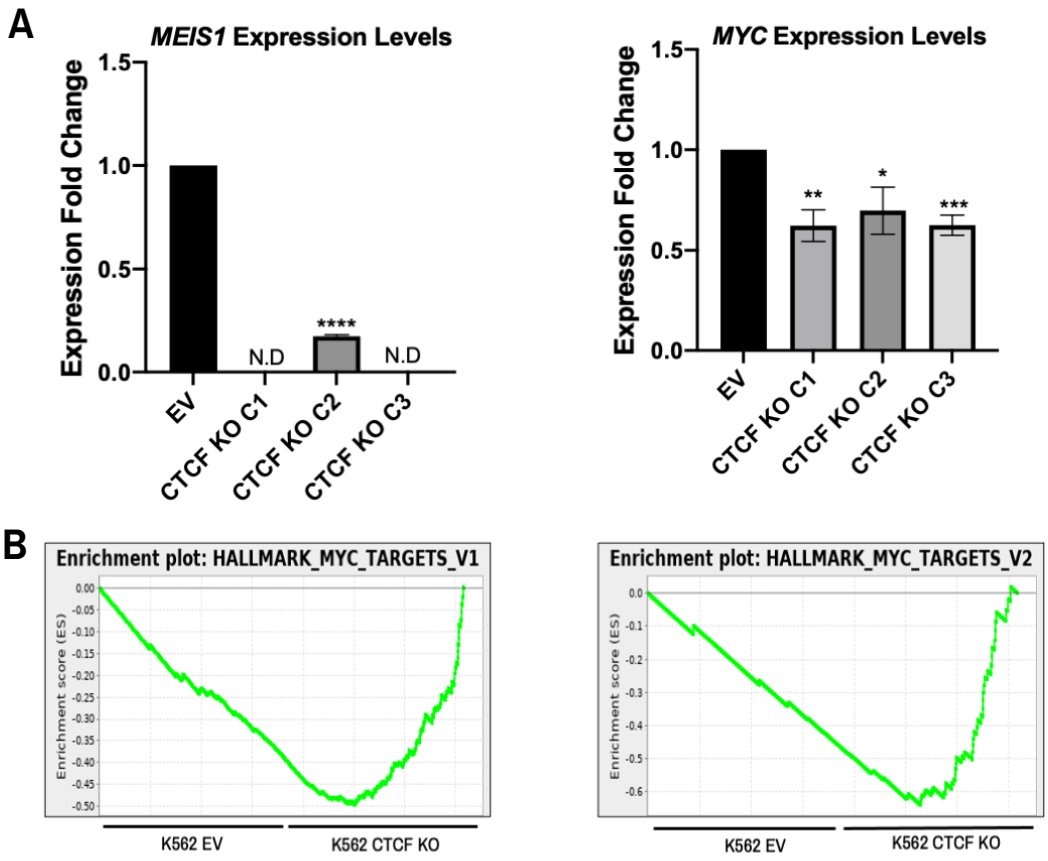

Figure 23. CTCF binding site excision results in *MEIS1* expression levels reduction in K562 cells. A. RT-qPCR results in K562 EV (empty vector) and CTCF binding site deleted cells on three clones. *MEIS1* RT-qPCR (left) and *MYC* RT-qPCR in negative control cells which went through the CRISPR process with empty vector ("empty vector") and CRISPR knockout ("CRISPR KO") cells. Data shown are average +/- standard error from three biological replicates. "N.D." indicates no gene expression was detected. One-tailed t-test was performed to evaluate significance, and the asterisks indicate the data is significant as per **** - P< 0.0001, *** - P< 0.001, ** - P< 0.01, *- P < 0.05 level. B. Gene Set Enrichment Analysis (GSEA) of two biological-replicates of RNA-seq from K562 EV and CTCF site deleted cells. Y Axis represents enrichment score (ES) at each of the samples. X Axis represents the CTCF KO and EV groups. Positive score shows positive correlation with the group that is indicated at the X axis. Negative score shows negative correlation with the group that is indicated at the X axis.

4.6 CTCF binding site excision leads to slower growth and increased apoptosis of K562 myeloid leukaemia cells

I also measured the effects of the CTCF binding site excision on K562 growth rates. Based on my cell viability assays of all three clones of K562 EV and CTCF binding site deleted cells, I observed that the growth rates of CTCF site excised cells were significantly slower over 72 hours as compared to the EV cells (**Figure 24A**).

Furthermore, my GSEA analysis of the RNA-seq from K562 EV and CTCF site excised cells showed an increased association to apoptosis gene sets and P53 tumour suppressing pathways **(Figure 24B)**. These findings may be explained in part by the downregulation of the *MEIS1* and *MYC* oncogenes expression after the loss of the CTCF binding site within the *MEIS1* FIRE. *MYC* expression is associated with cell proliferation (Delgado and León, 2010) and the activation of downstream oncogenes such as *CCND1* which is associated with leukaemia pathogenesis (Eisfeld et al., 2017, Fernandes et al., 2018). For example, the knockdown of *Myc* expression in a AML mouse model led to longer survival times (Brondfield et al., 2015). Therefore, the loss of *MEIS1* expression could result in the reduction of the expression of *MYC*, its downstream target genes and subsequently contribute to slower cell growth. In addition, the increased association to P53 tumour suppressing pathways in CTCF binding site excised K562 cells also suggest that the downregulation of oncogenes such as *MEIS1* and *MYC* has an effect on increasing P53 tumour suppressive activity. However, the understanding of the exact mechanisms of how these oncogenes may regulate the P53 tumour suppressive pathways still requires more future work.

Altogether, these may have a wide downstream effect on changing multiple genes expression thereby leading to reduced K562 myeloid leukaemia cell growth, possible alterations in apoptosis-related genes such as caspases and upregulation of associated P53 pathways **(Figure 24B**).

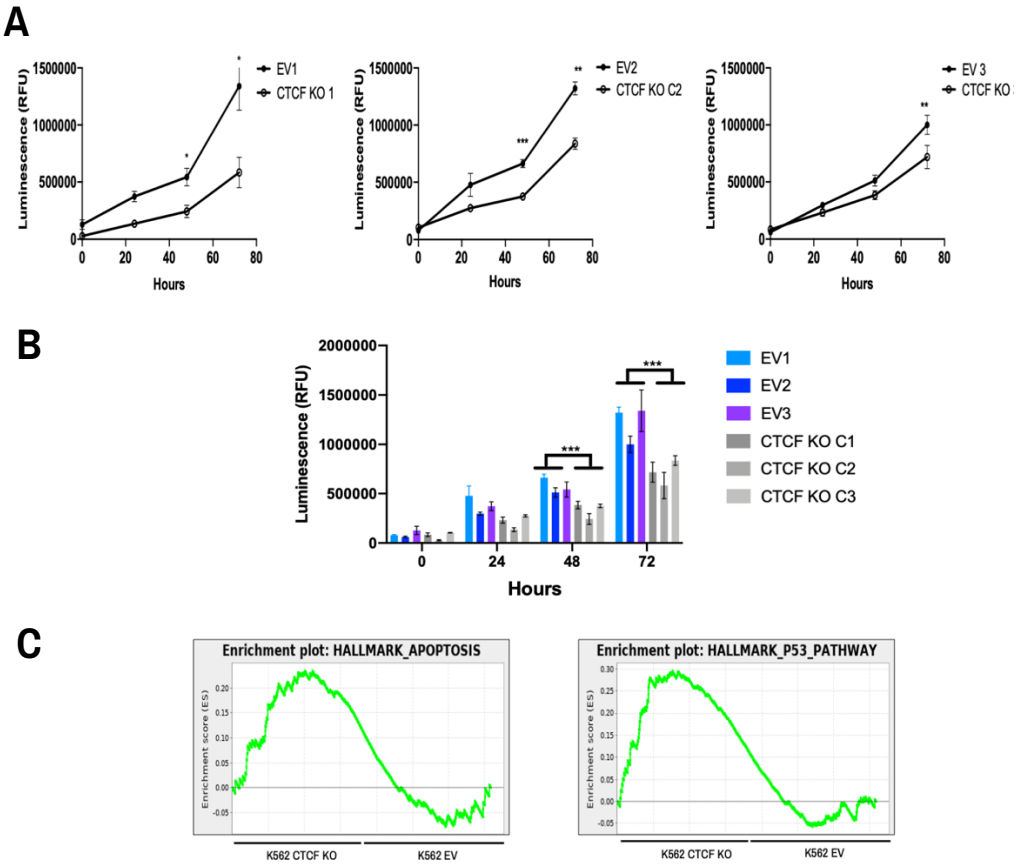

Figure 24. CTCF binding site deletion slows down K562 growth rates A. Cell viability assay was performed on three clones per treatment condition, and viable cell levels are shown on the y-axis. Y axis represents the luminescence (RFU) from the number of viable cells. X axis represents the hours of growth. Data shown are average +/- standard error from three biological replicates. One-tailed t-test was performed to evaluate significance, and the asterisks indicate the data is significant as per ** - P< 0.01, *- P < 0.05 level. B. Combined growth curve of all six K562 EV and CTCF KO clones. ***- P < 0.0001. C. Gene Set Enrichment Analysis (GSEA) of two biological-replicates of RNA-seq from K562 EV and CTCF binding site deleted cells. Y Axis represents enrichment score (ES) at each of the samples. X Axis represents the CTCF KO and EV groups. Positive score shows positive correlation with the group that is indicated at the X axis. Negative score shows negative correlation with the group that is indicated at the X axis.

4.7 Proposed model of *MEIS1* regulation via chromatin interactions in myeloid leukaemia

Collectively, I propose a model of *MEIS1* regulation in myeloid leukaemia. The intact presence of the *MEIS1* FIRE is involved in the maintenance of high *MEIS1* expression in normal CD34+ cells and a subset of myeloid leukaemias. Upon the disruption of *MEIS1* FIRE via CTCF boundary loss, I observed alterations to the

MEIS1 promoter-associated chromatin loops, and the subsequent downregulation of oncogenes such as *MEIS1*and *MYC*. The disruption of *MEIS1* FIRE also leads to slower growth and alterations in apoptosis-related genes in myeloid leukaemia **(Figure 25A)**. Here, I propose a model for the heterogeneous expression of *MEIS1* in myeloid leukaemia. The intact *MEIS1* FIRE and its associated loops are required for the high expression of *MEIS1* during the normal hematopoiesis stage of blood stem cells. Since *MEIS1* expression level was observed to be downregulated upon cell differentiation, I speculate that the *MEIS1* FIRE may be lost along with the downregulation of *MEIS1* expression upon hematopoietic cells maturation. Thus, the aberrant *MEIS1* expression may be contributed by the block in cell differentiation and conservation of the *MEIS1* FIRE in myeloid leukaemia. In addition, the heterogeneous nature of myeloid leukaemia *MEIS1* expression may be explained under two possible scenarios: Firstly, the presence of the *MEIS1* FIRE and its loops during myeloid leukaemia, thus resulting in continuous high expression of *MEIS1*. Or secondly, through the loss of *MEIS1* FIRE as observed in AML 28 and 30, resulting in lowered *MEIS1* expression (**Figure 25B**).

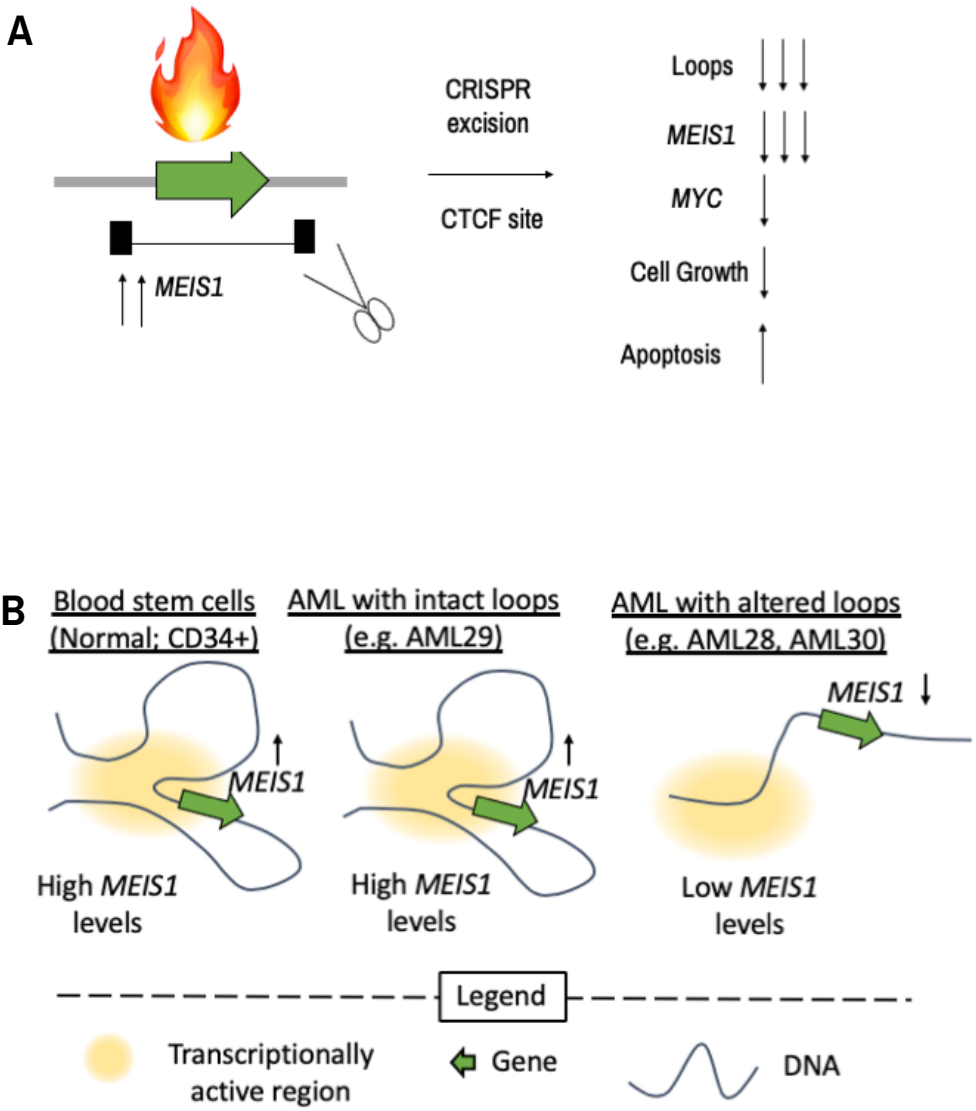

Figure 25 Schematic summary of *MEIS1* regulation in myeloid leukaemia. A. CRISPR excision of the FIRE (indicated by the "fire" icon) at *MEIS1* leads to *MEIS1* gene expression and other cellular changes in myeloid leukaemia. Up and down arrows indicate the increase and reduction in gene and pathways upregulation and downregulation respectively. B. A proposed schematic of how the loss of FIRE boundary at the *MEIS1* region results in the dysregulation of chromatin loops and downregulated *MEIS1* expression in myeloid leukaemia.

5 Super-enhancer occupancy at pre-existing chromatin interactions of AML precursor CD34+ blood stem cells at *HOXA9* is associated with high *HOXA9* expression level in myeloid leukaemia

5.1 *HOXA9* expression is co-expressed with *MEIS1* and significantly upregulated in AML

In addition to *MEIS1* regulation mechanisms, I set out to examine if chromatin interaction alterations could regulate *HOXA9* expression in myeloid leukaemia. Since MEIS1 and HOXA9 are co-factors at the protein level, I performed ddPCR on normal CD34+ and AML mononuclear cells (MNC) samples to assesses their *HOXA9* & *MEIS1* expression levels. The mononuclear cells are total populations of AML patients derived samples. My ddPCR results of normal CD34+ cells revealed that *HOXA9* and *MEIS1* are both expressed in normal CD34+ cells **(Figure 26A)**. However, *MEIS1* expression levels are significantly higher than *HOXA9* levels in the normal CD34+ group **(Figure 26B)**. These results verified previous reports that *HOXA9* and *MEIS1* are co-expressed in hematopoietic cells (Hu et al., 2009).

Similarly, I observed that *HOXA9* and *MEIS1* are expressed in AML patients, but with elevated *HOXA9* levels. Interestingly, *MEIS1* expression levels were not significantly different between normal CD34+ and AML cells **(Figure 27A)**. This supports my earlier findings that *MEIS1* levels continue to be expressed in AML in samples that possess the intact *MEIS1* FIRE. I also observed that approximately half of the 12 AML samples surveyed showed relatively proportionate expression of *HOXA9* and *MEIS1* levels (**Supplementary Figure 39A**). Furthermore, an analysis of 451 AML samples from a publicly available myeloid leukaemia database (Tyner et al., 2018) showed a positive correlation between *HOXA9* and *MEIS1* expression in this group of AML patients, with a pearson coefficient of 0.78 (**Supplementary Figure 39B**). These results strongly suggest that *HOXA9* and *MEIS1* expression levels are positively correlated in a subset of AML patients and also supports the information that *HOXA9* and *MEIS1* operate as co-factors at the protein level.

Further inspection of *HOXA9* levels in more mature CD33+ cells, showed agreement with previous literature (Ramos-Mejía et al., 2014) that *HOXA9* is downregulated upon cell differentiation from the hematopoietic stages **(Figure 27B)**. However, I observed that *HOXA9* was significantly overexpressed in AML as compared to both the normal CD33+ and CD34+ groups. My analysis of *HOXA9* and *MEIS1* expression levels in three previously Hi-C sequenced AML and three normal knee CD34+ samples showed that *HOXA9* and *MEIS1* are specifically upregulated in AML 29, but poorly expressed in AML 28 and 30 **(Figure 28A, B &C)**. Both Knee 25 and 27 CD34+ show some levels of *HOXA9* and *MEIS1* expressions as expected. Collectively, these evidences led me to hypothesize that *HOXA9* and *MEIS1* expressions are coordinated in AML as the presence of one oncogene is often seen with the other. Furthermore, I speculate that the upregulation of *HOXA9* levels in AML as compared to normal CD33+ and normal CD34+ cells may be due to the gain in chromatin loops to the *HOXA9* locus that drive oncogenic *HOXA9* overexpression in AML. Based on this hypothesis, I moved on to examine for potential differences in the TADs and chromatin loops at the *HOXA9* locus in clinical AML CD34+ cells and normal Knee CD34+ cells.

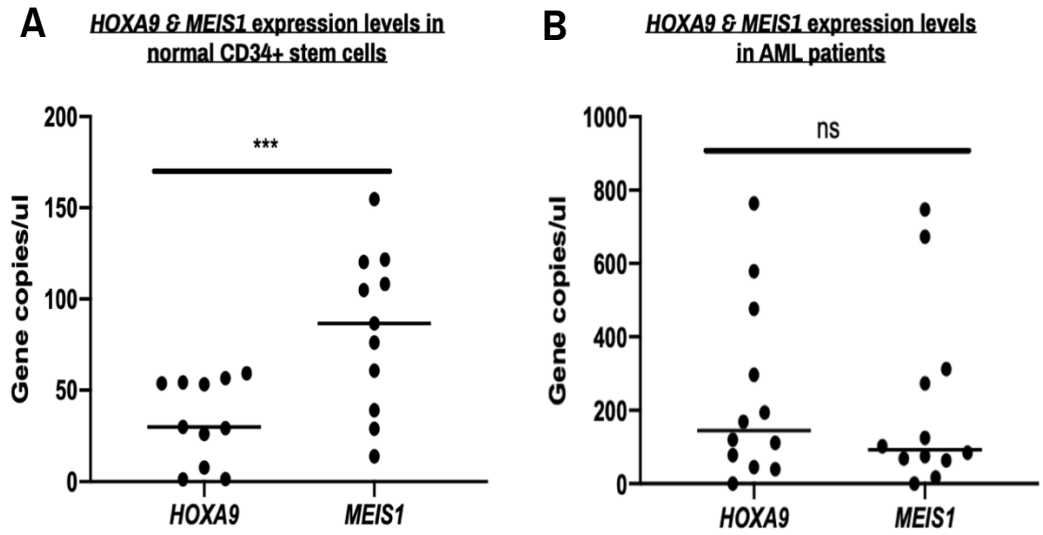

Figure 26. *HOXA9* and *MEIS1* are co-expressed in myeloid leukaemia. The ddPCR results of *HOXA9* and *MEIS1* expressions in normal CD34+ patients and AML patients. Two-tailed student t-test was performed to evaluate significance, and the asterisks indicate the data is significant as per *** -P< 0.001, ** - P< 0.01, *- P < 0.05 level.

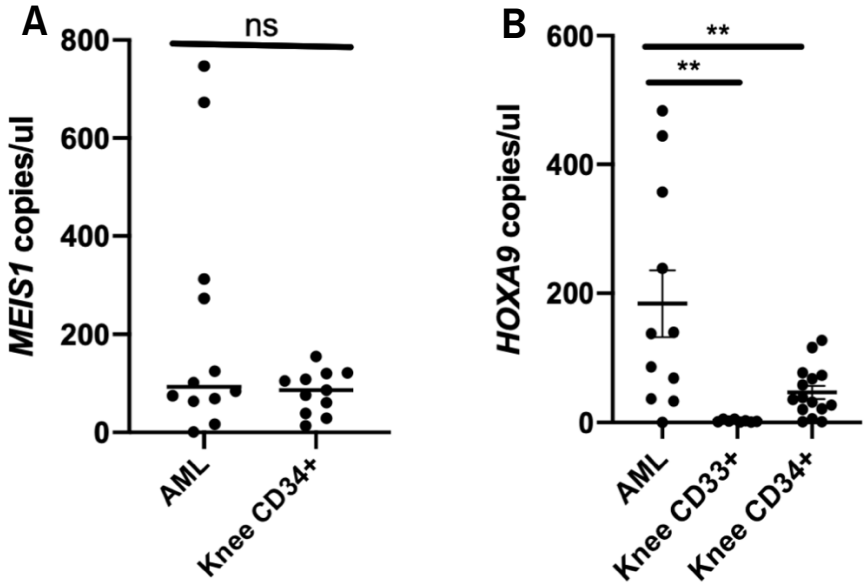

Figure 27. *HOXA9* expression levels is significantly upregulated in AML patient samples. A. The ddPCR results of *MEIS1* in AML and normal CD34+ cells. B. ddPCR results of *HOXA9* expression in AML, normal CD33+ & CD34+ cells. Two-tailed student t-test was performed to evaluate significance, and the asterisks indicate the data is significant as per *** -P< 0.001, ** - P< 0.01, *- P < 0.05 level. Error bars indicate standard error.

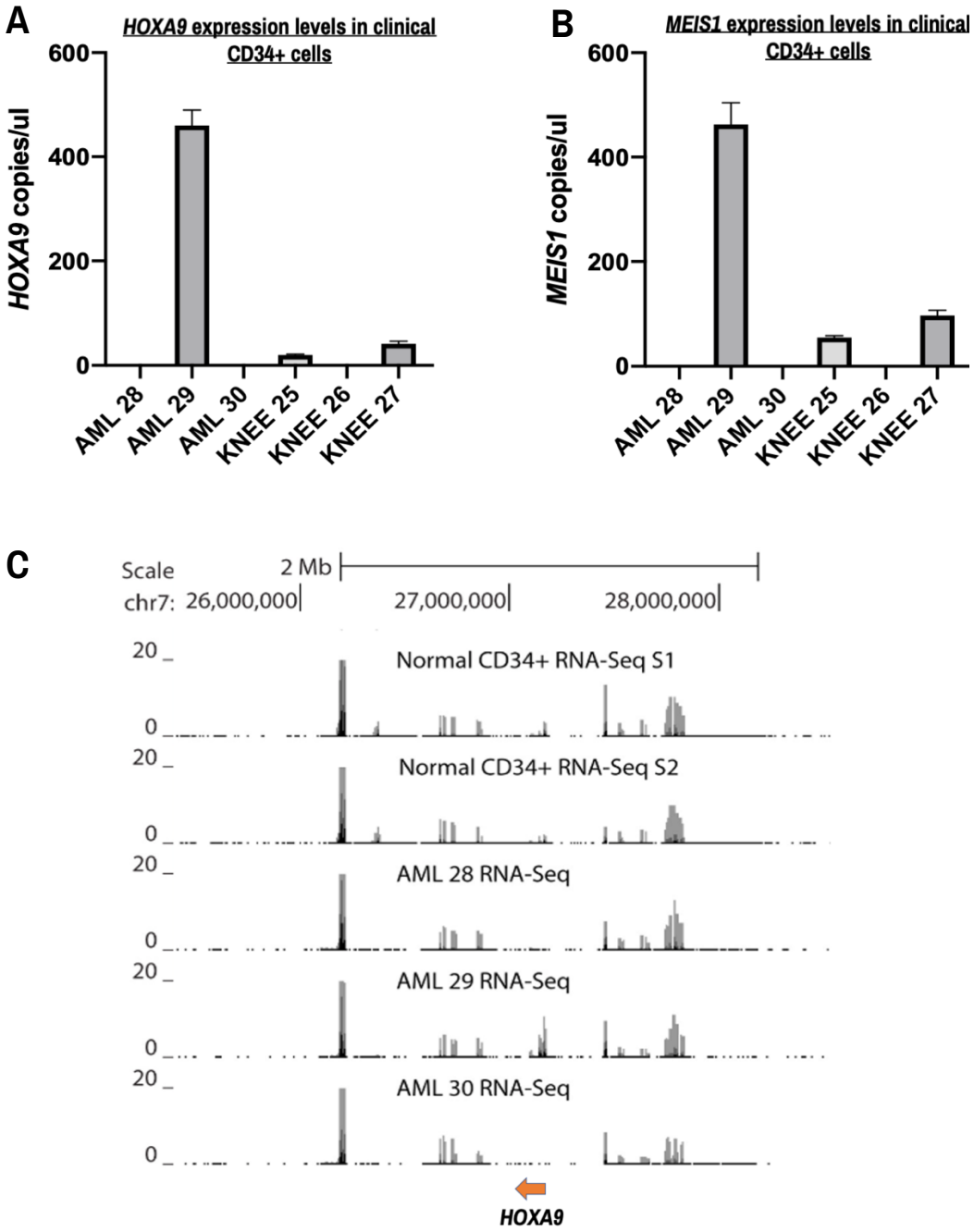

Figure 28. *HOXA9* and *MEIS1* expressions are strongly associated in AML CD34+ and normal CD34+ cells. A. The *HOXA9* expression levels in three AML and three normal CD34+ patients' samples. B. The *MEIS1* expression levels in three AML and three normal CD34+ patients' samples. Data shown indicate the average value of technical replicates performed on the same clinical sample and error bars indicate standard error. C. RNA-seq of *HOXA9* expression levels in two normal CD34+ and three AML CD34+ patients' samples. The arrow indicates the location of the *HOXA9* gene.

5.2 Topologically associating domains (TADs) remains unaltered at the *HOXA9* locus

Based on the Hi-C contact heatmaps, I examined the topologically associating domains (TADs) and loops of three AML and three normal knee CD34+ samples, for either potential alterations in chromatin loops, or gain/loss of TAD boundaries at the *HOXA9* locus. My inspection of the loops and TADs across the six CD34+ samples did not show obvious differences between their loops and TADs at the *HOXA9* locus **(Figure 29A & B)**. Since no gold standard of loop and TADs calling have been established, the minor difference between loops seen in the samples may be attributed to the algorithmic effect of the analysis package, and the presence of heterogeneity between each sample. At the TAD level, I do not observe significant differences between the Hi-C contact heatmaps of the Knee 50, AML 29 samples and THP-1 AML cell line at the *HOXA9* locus **(Figure 29B)**. Since I did not observe obvious differences between the TADs and loops of the Hi-C samples TADs, I hypothesized that they do not play a significant role in upregulating *HOXA9* expression in AML.

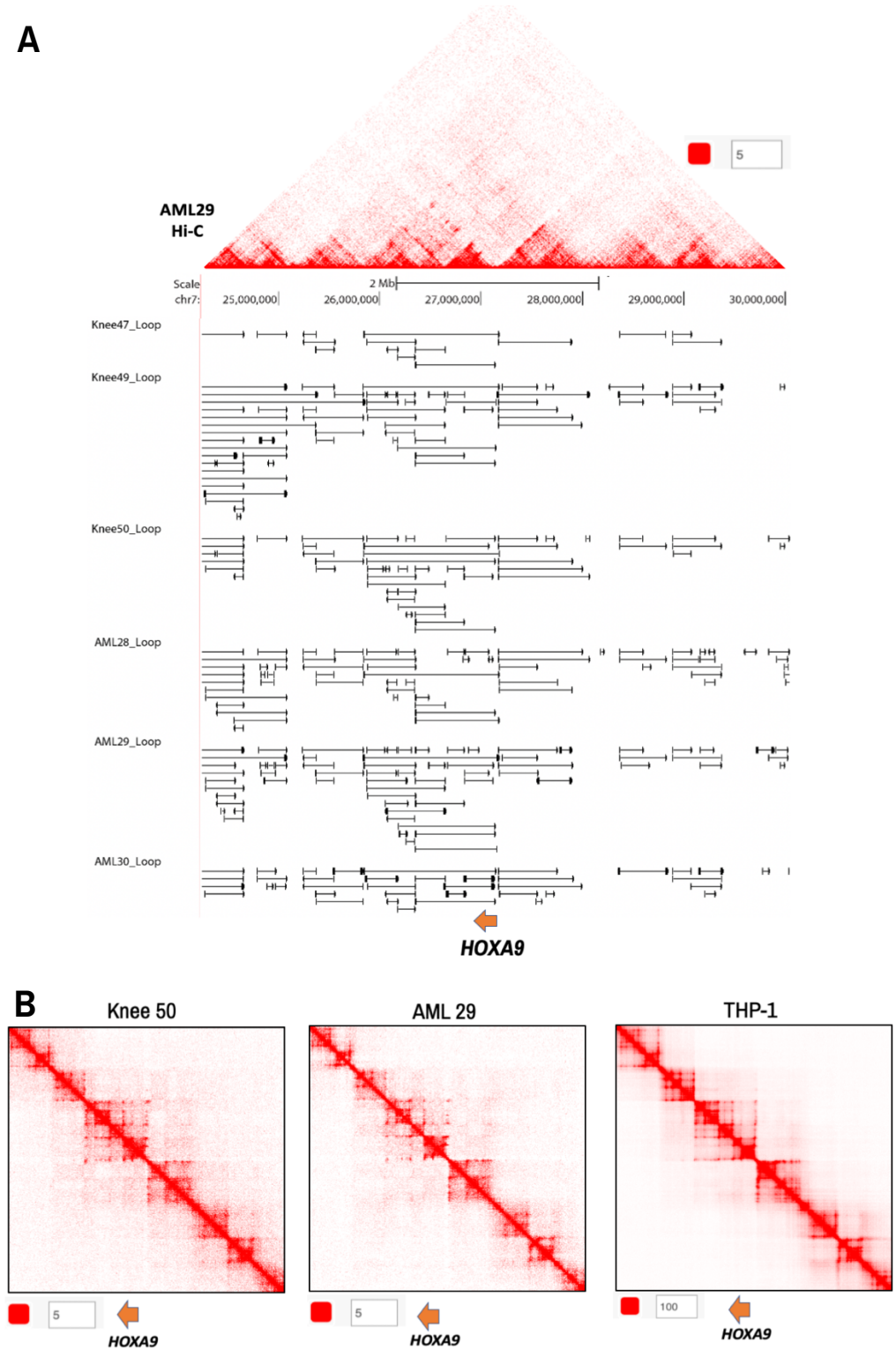

Figure 29. TADs and loops are largely unaltered at the *HOXA9* locus. A. Hi-C contact heatmap of AML29, aligned to the loops from six clinical CD34+ Knee and AML samples. Two points of a single line indicates the two points of anchors that are interacting. The location of the *HOXA9* gene is indicated below the loops. B. Hi-C contact heatmap of Knee 50, AML 29 & THP-1 at the *HOXA9* region. The arrow indicates the location of the *HOXA9* gene. Colour intensity of heatmap is indicate as a value at the bottom left-hand corner.

5.3 Chromatin interactions to *HOXA9* do not differ between THP-1 and HL-60 AML cell lines

I also designed 4C experiments in AML cell lines THP-1 and HL-60 with the *HOXA9* promoter as the bait, to ascertain that there are no significant differences between the chromatin loops at the *HOXA9* region. I then aligned the 4C results from THP-1 and HL-60 at the *HOXA9* region to the HI-C contact heatmap of THP-1 (**Figure 30A**). My 4C results from THP-1 and HL-60 showed similar chromatin interaction patterns to the *HOXA9* promoter, where I observed the spreading of interactions into the two sub-TADs that are located upstream and downstream of the *HOXA9* gene location (**Figure 30A**). I also assessed *HOXA9* expression levels of four blood cell lines- THP-1, HL-60, K562, GM12878 using ddPCR (**Figure 30B**). My results show that *HOXA9* expression levels are specifically upregulated in THP-1, but poorly expressed in HL-60, K562 and GM12878. Taken together, the Hi-C contact heatmaps, similar 4C interactions profiles of THP-1 and HL-60 and their *HOXA9* expression levels indicate that the increase in *HOXA9* expression level in THP-1 may not be a result of chromatin loop alterations at the *HOXA9* region.

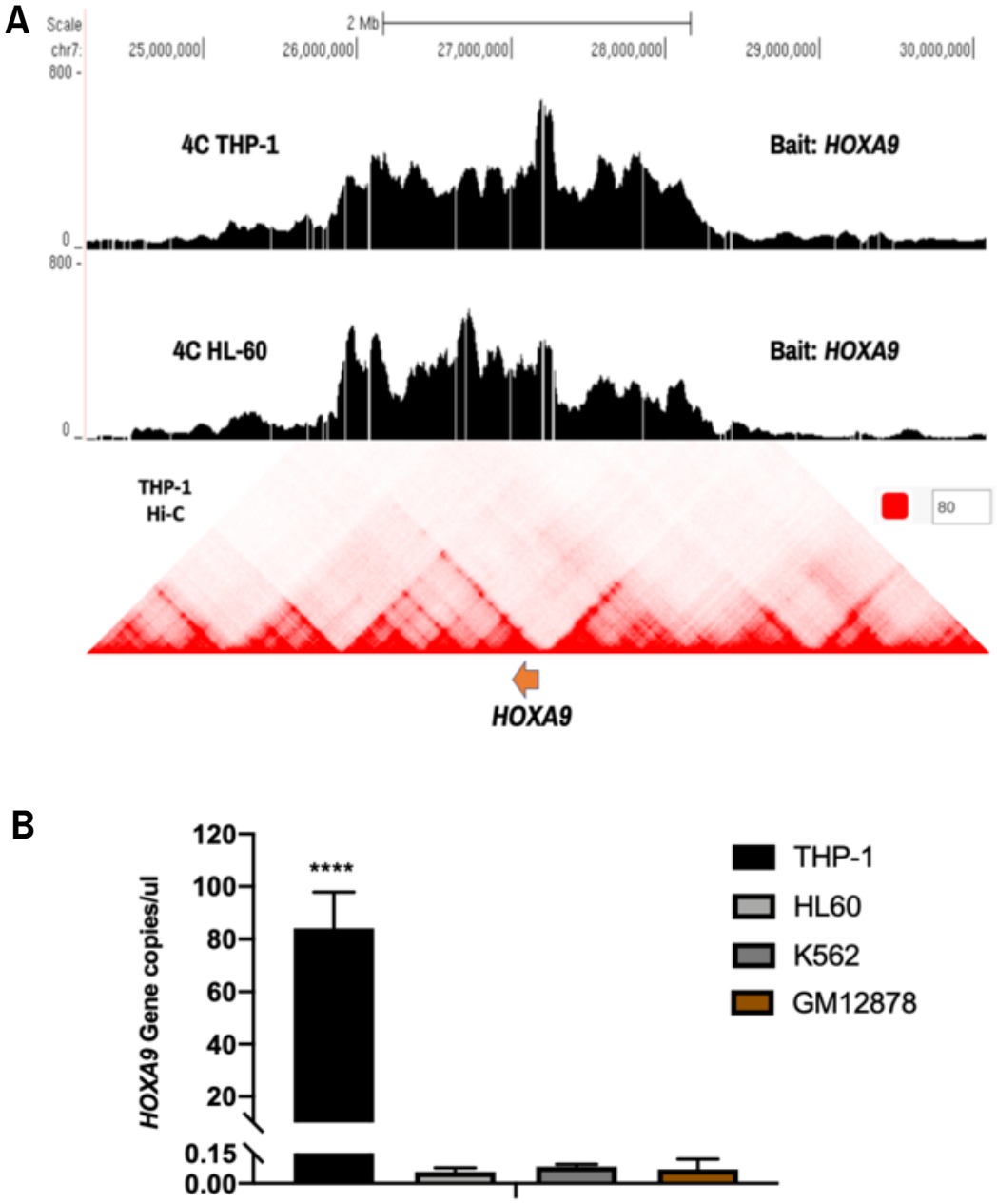

Figure 30. *HOXA9* associated chromatin interactions are similar across high & low expressing AML cell lines. A. 4C at the *HOXA9* viewpoint in HL-60 and THP-1 cells. Hi-C heatmap of TADs shown in THP-1. The arrow indicates the location of the *HOXA9* gene. Colour intensity of the Hi-C contact heatmap is shown as a value. B. ddPCR of *HOXA9* expression in blood cell lines THP-1, HL-60. K562 & GM12878. One-way ANOVA was performed to evaluate significance, and the asterisks indicate the data is significant as per **** - P< 0.0001 level. Data shown indicate the average value of three biological replicates performed on each cell line and error bars indicate standard error.

5.4 Super-enhancer occupancies are altered at the *HOXA9* region in AML

Based on my 4C and Hi-C results at the *HOXA9* locus, I concluded that the alteration of chromatin loops is unlikely to be responsible for the regulation of *HOXA9* in AML. This is because comparative analysis of the AML CD34+ cells and normal Knee CD34+ cells as well as in the AML cell lines of THP-1 and HL-60 showed no distinct alterations at the TADs and chromatin loops at the *HOXA9* locus. This result was also supported by my ddPCR results that showed that the THP-1 cell line expressed high levels of *HOXA9* while HL-60 expression of *HOXA9* is significantly lower. Taken together, these results suggest that the TADs and chromatin loops are unlikely to contribute to different level of *HOXA9* expression levels in AML and the THP-1 and HL-60 cell lines.

Instead, my new hypothesis was that epigenetic changes and acquisition of enhancers and super-enhancers at distal regions of the genome that loop to the *HOXA9* gene promoter could be the underlying means for the upregulation of *HOXA9* expression levels in AML. To investigate this hypothesis, I performed super-enhancer analysis on the publicly available H3K27ac ChIP-seq data of four blood lineage cell lines THP-1 (Mohaghegh et al., 2019), HL-60 (Barbieri et al., 2018), K562 (Burr et al., 2019) and GM12878 (Zhang et al., 2018) at the *HOXA9* region. I then aligned the Hi-C contact heatmap and loops to the super-enhancer tracks at the *HOXA9* region.

My super-enhancer analyses revealed the presence of three super-enhancer regions which interact with the *HOXA9* gene, with each super-enhancer present in a subset of AML patients (**Figure 31**). Among the three super-enhancers identified, I observed a large super-enhancer which is present in the intronic region of the *SNX10* gene (*SNX10* SE). I also discovered the existence of a smaller super-enhancer, that is approximately 300kb in size, downstream of the *HOXA9* gene. Both of these super-enhancer profiles were seen specifically only in AML samples but not in normal CD34+ samples. Comparison of these two super-enhancers also revealed that *SNX10* SE is only present in THP-1 but not the other three blood cell lines (**Figure 31**). In addition, I observed that *SNX10* SE and SE3 are present in a subset of AML patients but not in the normal CD34+

samples, which suggest that they could have been acquired during the disease pathogenesis.

Furthermore, I observed that these acquired super-enhancers are present in locations that show chromatin interactions with the *HOXA9* promoter. However, it is important to note that super-enhancer acquisition is not seen at all the locations that possess chromatin interactions to the *HOXA9* promoters. My previous inspection of the 4C results of HL-60 showed similar interaction profiles to THP-1 but without the presence of *SNX10* SE and SE3. These indicate that the *HOXA9* promoter-associated chromatin interactions are likely to be preexisting features, regardless of their super-enhancer occupancy profiles. Collectively, these findings may suggest specific super-enhancers such as *SNX10* and SE3 are acquired in AML, and subsequently hijack preexisting chromatin interactions to the *HOXA9* promoter, thus leading to upregulation of *HOXA9* gene expression, resulting in downstream cancer-related changes.

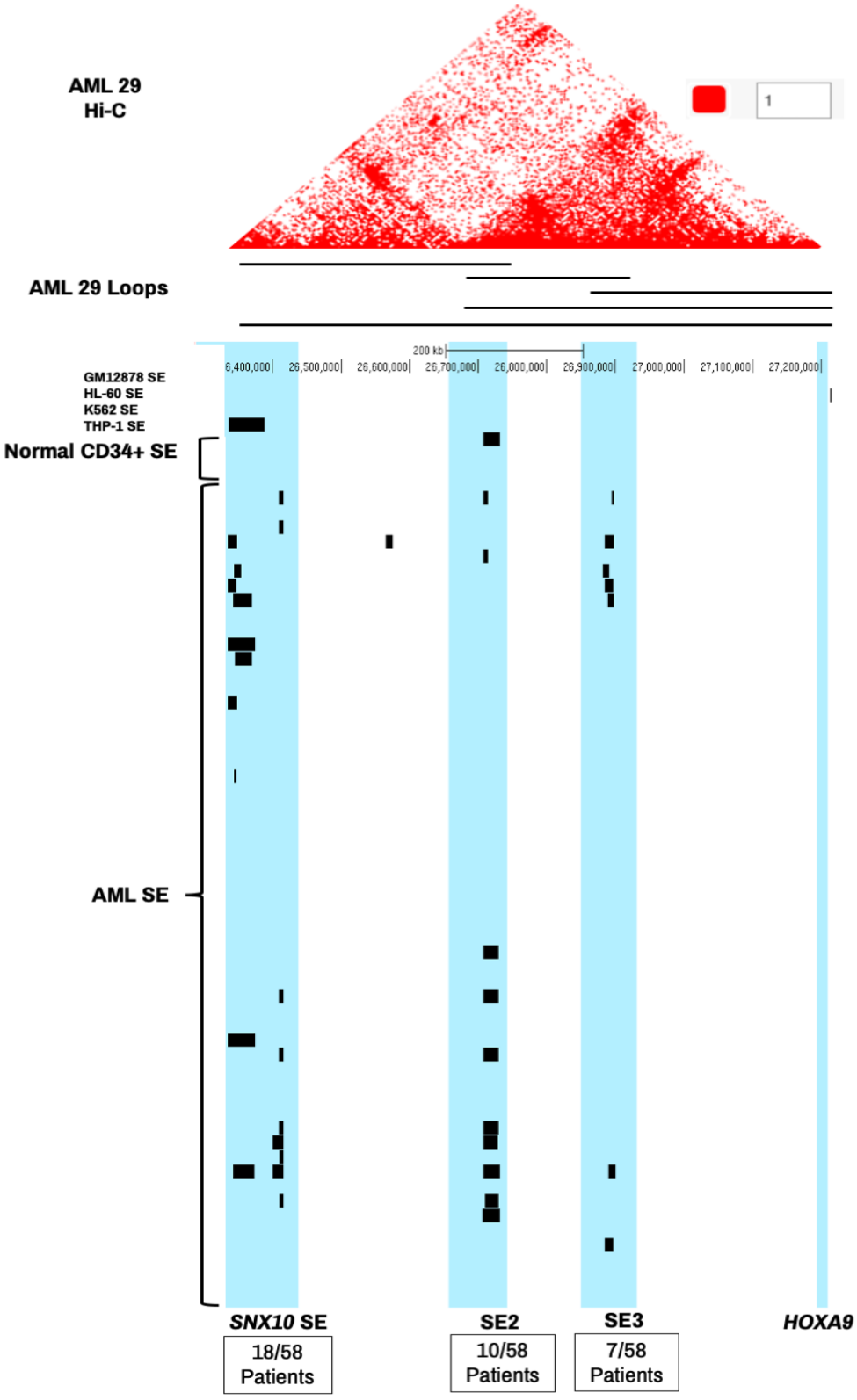

Figure 31. *HOXA9* and super-enhancers interact via chromatin loops in AML Loops between *SNX10* SE and *HOXA9* are shown together with SEs profiles from two normal CD34+ donors and 58 AML patients (each row represents one patient, and the black bars indicate SEs). Regions with many SEs found in many patients are indicated by blue stripes, as well as the *HOXA9* promoter.

5.5 Super-enhancer acquisition in AML patients at the *SNX10* genome region is associated with high *HOXA9* expression

Since *SNX10* SE is identified as a large super-enhancer with interactions to the *HOXA9* promoter, I proceed to investigate its association with *HOXA9* overexpression in AML. Using RNA-seq analysis, I first examined *HOXA9* expression levels across 62 AML MNC samples, 9 normal knee MNC samples and 13 normal knee CD34+ samples. The reason for examining these three groups of clinical samples is because AML pathogenesis occurs across various subtypes due to the block in cell differentiation at different levels of blood cells maturations, from the hematopoietic to the myelocytic stages. Therefore, the examination of *HOXA9* expression level in both AML MNC and normal knee MNC cells and normal knee CD34+ samples can provide us with a clearer outlook of the levels of *HOXA9* overexpression in AML MNC as compared to the normal MNCs as well as to the normal CD34+ cells that inherent express some housekeeping levels of *HOXA9* for cellular development processes.

My analysis shows that *HOXA9* expression levels were significantly upregulated in AML as compared to both the mixed cell population of knee MNC and CD34+ samples. This supports previous results that *HOXA9* expression is strongly upregulated in AML (Thorsteinsdottir et al., 2002) (**Figure 32A)**.

Applying the MACS2 (Zhang et al., 2008a) and ROSE (Hnisz et al., 2013a) algorithms, I further analysed the super-enhancer profiles of the 58 AML patient samples and revealed that 31% of them possess the *SNX10* SE (**Figure 32B**). Interestingly, the comparison of *HOXA9* expression levels in these group of patients with the *SNX10* SE and the normal knee MNC and normal CD34+ samples revealed that approximately 94% of these patients with the *SNX10* SE also concurrently expressed high levels of *HOXA9* that are above the level of normal CD34+ and MNC cells. Therefore, these results suggest that the *SNX10* SE is strongly associated with the upregulation of *HOXA9* in a subset of AML patients.

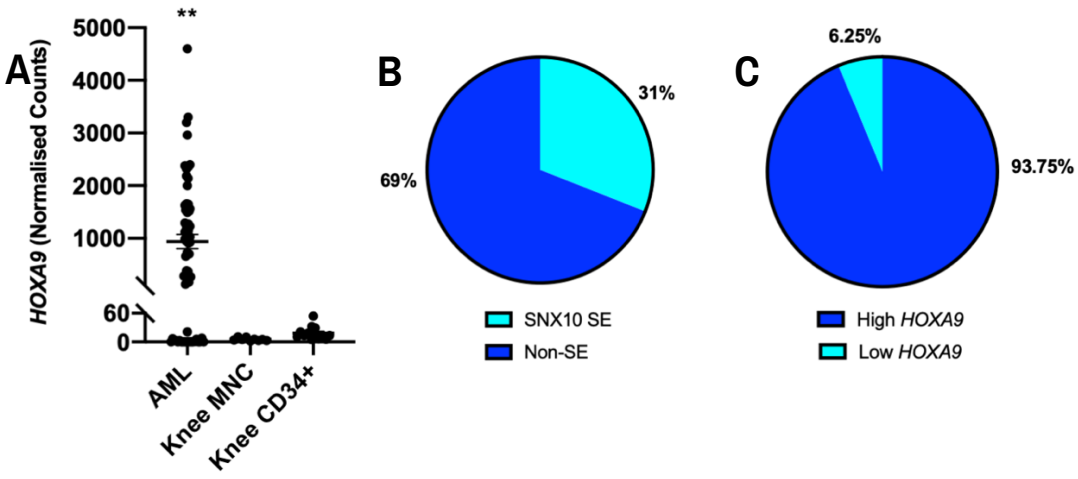

Figure 32. *HOXA9* expression is highly associated with *SNX10* super-enhancers in AML. A. ddPCR of *HOXA9* expression in AML, normal knee mononuclear cells (MNC) and CD34+ cells. Two-tailed student t-test was performed to evaluate significance, and the asterisks indicate the data is significant as per ** - P< 0.01 level. Error bars indicate standard error. B. Pie charts showing the percentage of the 58 AML patients with the *SNX10* super-enhancer. C. Pie chart showing the percentage of AML patients with *SNX10*-super-enhancer and high *HOXA9* expression.

5.6 HOXA9 & MEIS1 proteins bind to each other's promoter in THP-1 cells

Next, I sought to understand how MEIS1 and HOXA9 may mutually co-regulate each other in AML. In the first chapter, I hypothesized that the MEIS1 and HOXA9 proteins are able to bind to each other's promoters to contribute to each other's transcription. Thus, I performed chromatin immunoprecipitation qPCR (ChIP-qPCR) on the THP-1 AML cell line. First, I pulled down HOXA9 and MEIS1 proteins with their specific antibodies. I then performed qPCR on selective regions of interest such as the promoters of *HOXA9*, *MEIS1* and *MYC* oncogenes to examine for the presence of binding HOXA9 and MEIS1 proteins. My ChIP-qPCR results from the HOXA9 pulldown experiment showed that the HOXA9 protein binds to its own promoter and that of *MEIS1* and *MYC* (**Figure 33A**).

Interestingly, I also observed that the HOXA9 protein binds to the *SNX10* SE region in THP-1. These results indicate that *SNX10* SE and *HOXA9* promoter form chromatin loops with HOXA9 proteins binding at both regions within the loop. Therefore, potentially leading to upregulation of *HOXA9* gene expression in the THP-1 AML cell line. ChIP-qPCR analysis of the MEIS1 pulldown region shows that MEIS1 also binds to its own promoter as well as those of *HOXA9* and *MYC* (**Figure 33B**). The binding of both HOXA9 and MEIS1 proteins at each other's promoter suggests that they are involved in mutual regulation of their own expression in a potential regulatory feedback loop. This feedback loop may allow for increased levels of HOXA9 and MEIS1 proteins to be produced, which consequently bind to the *MYC* promoter to contribute to *MYC* upregulation in myeloid leukaemia. An example of how the MEIS1 protein can impact *MYC* expression is seen in my experiment, where the deletion of a CTCF boundary led to *MEIS1* expression downregulation and the subsequent reduction in *MYC* expression and downstream MYC associated pathways in K562 cells.

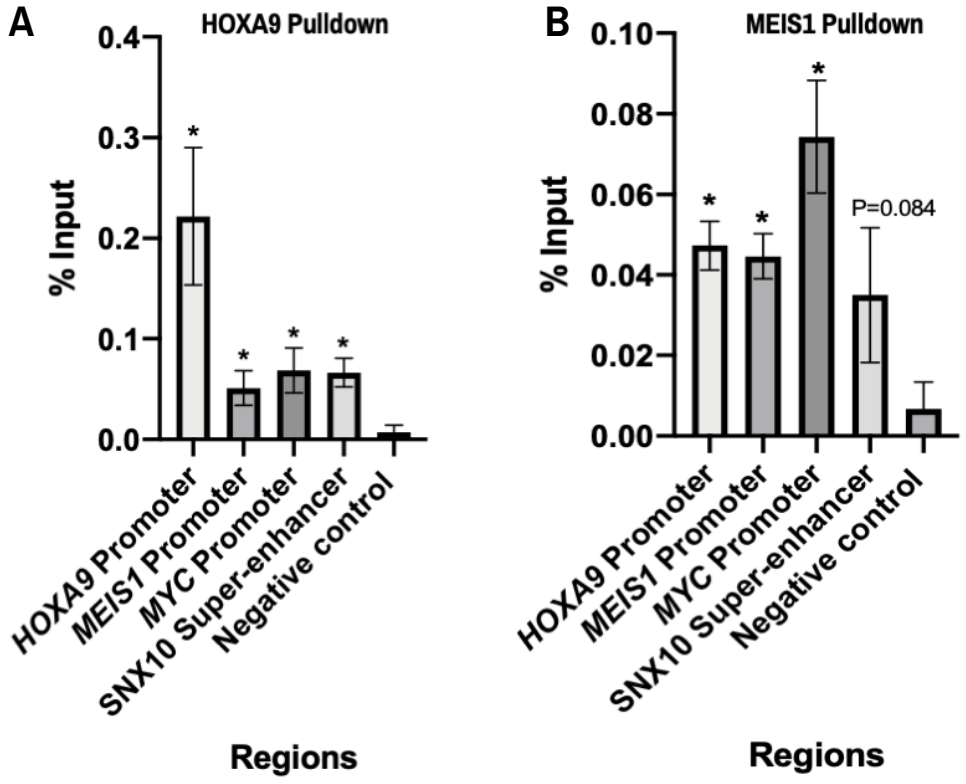

Figure 33. HOXA9 and MEIS1 proteins bind to each other's promoter to contribute to oncogene expression in THP-1 AML cells. A. ChIP-qPCR of HOXA9 antibody pulldown at the *HOXA9, MEIS1, MYC* promoters and *SNX10* super-enhancer. B. ChIP-qPCR of MEIS1 antibody pulldown at the *HOXA9*, *MEIS1*, *MYC* promoters and *SNX10* super-enhancer. One-tailed student t-test was

performed to evaluate significance, and the asterisks indicate the data is significant as per * - P< 0.05 level. Error bars indicate standard error.

5.7 Schematic of proposed mechanism: Heterogeneous super-enhancer occupancy and juxtaposition with *HOXA9* and *MEIS1* gene promoters via chromatin interaction leads to heterogeneous *HOXA9* and *MEIS1* gene expression observed in myeloid leukaemia.

The deletion of the CTCF boundary downstream of the *MEIS1* gene region resulted in the alterations and wide losses of *MEIS1* promoter-associated chromatin loops, and the consequent downregulation of *MEIS1* expression levels. This implies the importance of the *MEIS1* FIRE for the expression of *MEIS1*. In contrast, the *HOXA9* region showed abundant chromatin interactions that were similar between AML and precursor haematopoietic stem cells. On the other hand, super-enhancers that are brought in close proximity to the *HOXA9* promoter through preexisting chromatin loops were found to be acquired in both the THP-1 AML cell line and a subset of the AML patient samples. Interestingly, these super-enhancers were found absent the normal CD34+ samples.

Based on my results obtained, I propose a synergistic model that may explain the coordinated regulation of *HOXA9* and *MEIS1* expression in myeloid leukaemia. Since HOXA9 and MEIS1 are reported to be important in hematopoietic stem cells, I conceive that the presence of the *MEIS1* FIRE (as seen in the normal knee CD34+ samples) and the interaction of a few super-enhancers (SE2) with the *HOXA9* promoter, allows for the high expression of *HOXA9* and *MEIS1* levels in normal CD34+ cells (**Figure 34A**).

However, in the onset of myeloid leukaemia, two potential scenarios may occur. First being that the continuous retention of *MEIS1* FIRE due to the block in cell maturation in AML may lead to the continuous expression of *MEIS1*. Secondly, super-enhancers may be acquired at preexisting chromatin loops to the *HOXA9* promoter as in the case of the *SNX10* SE, which leads to the upregulation of *HOXA9* expression. Since *HOXA9* and *MEIS1* proteins have been found to bind each other's promoters via a positive feedback loop, the expression of these two oncogenes will be highly elevated in myeloid leukaemia (**Figure 34B**).

On the other hand, the reason for low *MEIS1* and *HOXA9* expression in certain AML patients may be attributed to their loss of the *MEIS1* FIRE, which leads to the reduction of *MEIS1* expression. This group of patients with the loss of *MEIS1* FIRE may also lack the acquired super-enhancers at the *HOXA9* locus, therefore leading to the low expression of *MEIS1* and *HOXA9* **(Figure 34C)**.

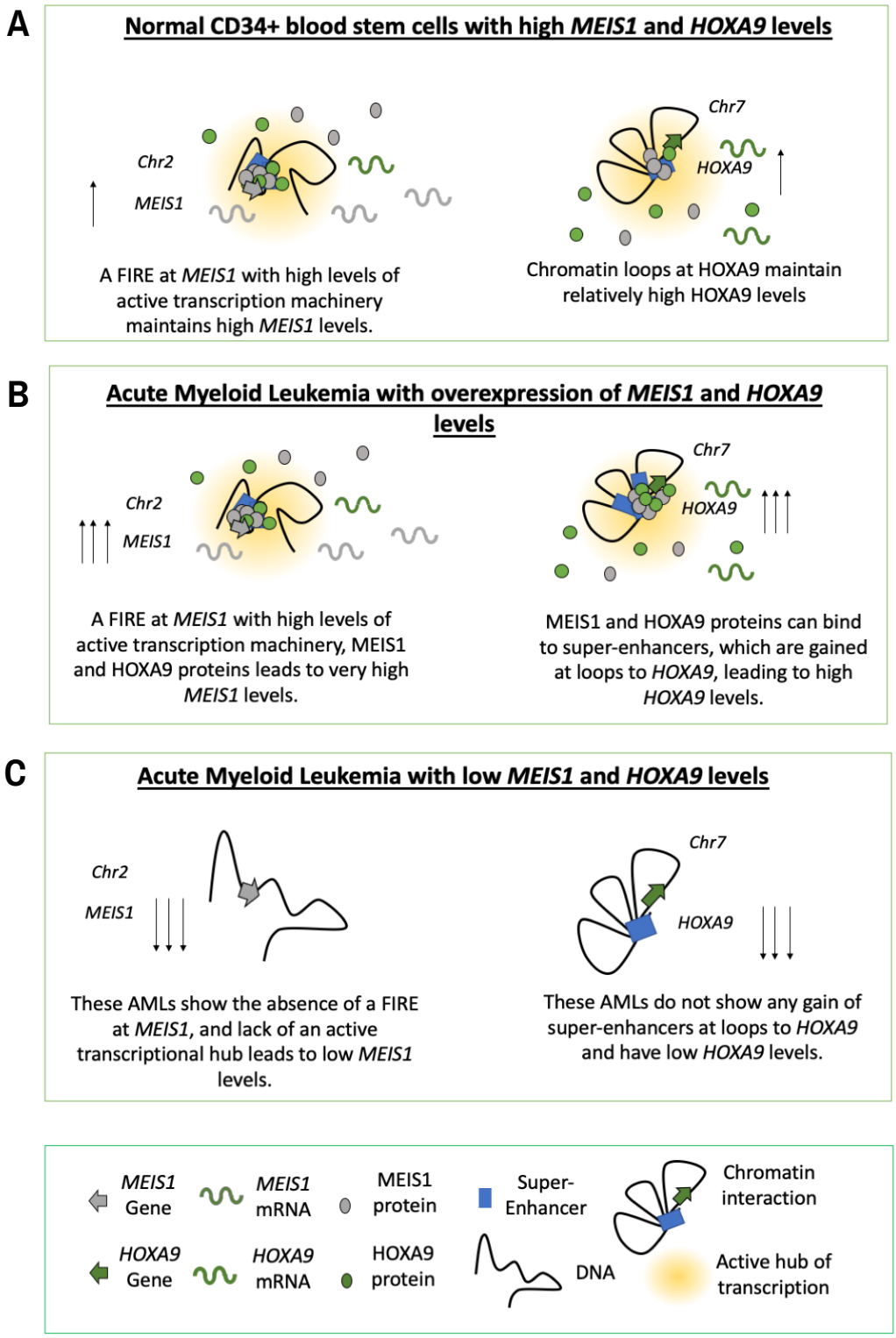

Figure 34. Proposed model of *MEIS1* & *HOXA9* regulation in myeloid leukaemia. A. Schematic of high *MEIS1* and *HOXA9* expression via *MEIS1* FIRE and super-enhancer to *HOXA9* promoter interactions during normal hematopoiesis stage. B. Schematic of increased *MEIS1* and *HOXA9* expression via intact conservation of *MEIS1* FIRE and gain of super-enhancers that hijacks existing chromatin loops to the *HOXA9* promoter in a subset of myeloid leukaemias. C. Schematic of low *MEIS1* and *HOXA9* expression via the loss of

MEIS1 FIRE with no additionally gained super-enhancer to *HOXA9* promoter interactions in a subset of myeloid leukaemias

6 Discussion & Conclusion

In the first chapter, I discussed the roles that MEIS1 and HOXA9 proteins play in the maintenance of stem cell hematopoiesis and how the loss of these important transcription factors impacts normal cellular maturation and development. Furthermore, *MEIS1* and *HOXA9* have been widely reported to be aberrantly upregulated in leukaemia, with high co-expression in myeloid leukaemias (Lawrence et al., 1999). A question that has constantly surfaced is "What regulates *HOXA9* and *MEIS1* overexpression in myeloid leukaemia?". To date, mix lineage leukaemia (*MLL*) translocation and MLL fusion proteins are widely reported to be associated with *HOXA9* and *MEIS1* upregulation in leukaemia. For example, Zeisig et al. (2004), demonstrated that the *MLL-ENL* translocation and its resulted fusion protein was essential in upregulating HOXA9 and MEIS1 proteins in hematopoietic cells. However, the understanding of mechanisms that are responsible for regulating *MEIS1* and *HOXA9* overexpression in myeloid leukaemia remains limited.

In this study, I demonstrated that the *MEIS1* FIRE is important in regulating *MEIS1* expression in K562 myeloid leukaemia cells. In addition, I also showed that the deletion of the CTCF binding site in the *MEIS1* FIRE results in large-scale chromatin loops reorganization at the *MEIS1* locus and the loss of chromatin interactions between three super-enhancers and the *MEIS1* promoter. In addressing the potential role of super-enhancers in regulating *MEIS1* and *HOXA9* expression in myeloid leukaemia, I also revealed the presence of newly acquired super-enhancers in a subset of clinical AML patient sample and the AML cell line THP-1 at the *HOXA9* locus. Finally, my ChIP-qPCR results of THP-1 cells show that HOXA9 and MEIS1 proteins bind to their own and each other's promoters to contribute to coordinated *HOXA9* and *MEIS1* regulation. I also found that HOXA9 proteins bind to a novel super-enhancer at the *SNX10* gene region in the genome and that this *SNX10* SE forms chromatin loops with the *HOXA9* promoter, thereby suggesting the role that the *SNX10* SE has in the overexpression of *HOXA9* in THP-1 cells and potentially in a subset of myeloid leukaemia with high *HOXA9* overexpression.

This improved understanding of alternative mechanisms such as the regulation of oncogenes via aberrant chromatin interactions will be beneficial towards expanding the knowledge on myeloid leukaemia pathogenesis and the development of potential treatments. Such treatments and drugs include those that are targeted to the disruption of chromatin interactions between super-enhancers and the *HOXA9* and *MEIS1* promoter for myeloid leukaemias with high *HOXA9* and *MEIS1* expression. Alternatively, the disruption of FIREs such as the *MEIS1* FIRE via CTCF deletion may also present as a potential route for gene therapy in the future.

6.1 FIREs are essential for the expression of their associated cell-type specific genes

In the first part of my study, I discovered the presence of a "Frequently Interacting Region" (FIRE) at the *MEIS1* region in the clinical AML 29 CD34+ and normal knee derived CD34+ samples as well as in the THP-1, K562 cell lines. I likewise demonstrated that the preservation of chromatin interactions at the *MEIS1* FIRE is essential in consistently promoting the high expression of *MEIS1* in the K562 myeloid leukaemia cells. The presence of the *MEIS1* FIRE with accompanied high *MEIS1* expression in normal hematopoietic CD34+ and AML 29 CD34+ samples suggest a strong association between elevated *MEIS1* expression and the presence of the *MEIS1* FIRE region.

My CRISPR-Cas9 excision of the CTCF binding site that is located downstream of the *MEIS1* gene resulted in substantial chromatin loops alterations at the *MEIS1* locus. These chromatin loops alterations include the loss of interactions between the *MEIS1* promoter and super-enhancers that are in proximity to the *MEIS1* gene. Furthermore, the loss of *MEIS1* promoter-associated chromatin loops also led to the loss and downregulation of the oncogenes *MEIS1* and *MYC* expression. Growth of CTCF boundary excised K562 cells also exhibited significantly slower growth kinetics and increased association to apoptosis and the P53 tumour-suppressing pathway. Collectively my findings suggest the importance of the conservation of *MEIS1* FIRE in ensuring continuous *MEIS1, MYC* and downstream oncogenes expression.

To date this is the first study in cancer biology that reported the presence of a FIRE at the *MEIS1* locus and depicts its role in maintaining the expression of *MEIS1* in myeloid leukaemia. As I highlighted above, FIREs are widely found to be located near cell-type specific genes such as *MEIS1*, thus suggesting their role in regulating these gene-type specific genes. Since cancer commonly features overexpressed oncogenes such as *HOXA9* and *MEIS1* that are originally expressed to serve a specific cell stage or process such as haematopoiesis, the continued retention of FIREs at such oncogenes, beyond the hematopoietic cell stages may therefore be a factor that contributes to their aberrant expression in cancers like myeloid leukaemia. Thus, FIREs are areas that should be further investigated as a potential target for cancer treatment.

6.2 CTCF is important towards the maintenance of the chromatin interactions in a FIRE

In addition, I also demonstrated that CTCF binding is important in maintaining chromatin interactions within the FIRE. The deletion of a CTCF boundary within the *MEIS1* FIRE led to large scale chromatin changes in that region.

The CTCF protein has been widely reported to play an important role in acting as an insulator in the process of TAD formation (Nora et al., 2017). In addition, Schmitt et al. (2016b) have also suggested that the CTCF protein is a feature of FIRE. However, our understanding of the role that individual CTCF binding sites in a FIRE plays remains unknown. Here I showed that the removal of the CTCF binding site leads to the abrogation and reorganization of pre-existing chromatin loops that are dependent on the insulation that CTCF binding provides. Interestingly, the removal of the downstream CTCF binding site to *MEIS1* was sufficient to induce *MEIS1* promoter linked chromatin loops changes upstream and down of the *MEIS1* gene, with loss of loops being the main type of loop alterations seen.

This observation is agreed by the work of Schmitt et al. (2016b), who analysed the FIREs from Hi-C samples of cohesin subunit SMC3 depleted and non-depleted HEK293 cells. The reason for the analysis of depleted and non-depleted cohesion subunit SMC for differences in FIREs is because cohesin proteins have

been suggested to mediate cohesin associated chromatin loops and that these include those that are formed at FIRE and TAD boundaries in tandem with CTCF proteins (Kagey et al., 2010). Their analysis was based on the evaluation of FIRE scores that are algorithm-generated, with higher scores referring to stronger regions of FIRE in the genome. The analysis revealed that FIRE bins with both CTCF and SMC binding showed a significant reduction in their FIRE scores as compared to FIREs that only had CTCF binding alone (Schmitt et al., 2016a). They further analysed HEK293 cells with CTCF knockdown (Zuin et al., 2014) and observed that the FIRE scores were most reduced at FIREs that originally possess CTCF and SMC binding (Schmitt et al., 2016b).

These published results support my findings that CTCF and cohesin proteins play a role in regulating FIREs and that the removal of a CTCF binding site at the *MEIS1* FIRE results in the downregulation of *MEIS1* and its associated downstream oncogenes such as *MYC* in the K562 myeloid leukaemia cells. This is also supported by widely reported facts that cohesin and CTCF proteins tend to function together at TAD boundaries to form insulators in the genome (Ali et al., 2016). However, it is important to note that FIRE scores were reduced but not completely abrogated upon cohesin-SMC or CTCF knockdown. This suggests that CTCF and cohesin may regulate FIREs in a slightly different manner as to TADs.

Although CTCF and cohesin are reported to act as insulators at genomic regions in the formation of TADs, the knockdown or deletion of CTCF at TAD boundaries appear to have distinctly different outcomes to CTCF disruption in FIREs (Lupiáñez et al., 2015, Guo et al., 2015). The disruption of CTCF boundaries results in the reorganization of chromatin loops and interactions of enhancer regions and genes that were originally in adjacent TADs, leading to aberrant gene expression. The reversion of the disrupted CTCF boundary leads to the recovery of TADs compartments in the genome (Nora et al., 2017).
In another piece of work, Flavahan et al. (2016) demonstrated that the hypermethylation at a CTCF boundary in mutant *IDH* (isocitrate dehydrogenase) glioma leads to the disruption of the TADs and results in an aberrant enhancer interaction to the receptor tyrosine kinase oncogene *PDGFRA* (**Figure 35**). The

treatment of this mutant *IDH* glioma with a demethylating drug consequently lead to the partial rescuing of TAD boundaries and reduction in *PDGFRA* expression (Flavahan et al., 2016). This study demonstrates the effect of CTCF loss on TAD boundary alterations and the consequent occurrence of dysregulation enhancer-gene interactions in cancer. While FIREs and TADs are both associated with CTCF and cohesin at their boundaries, the effects of CTCF and cohesin disruption are observed to have different degrees of modulation on the chromatin interactions in the region.

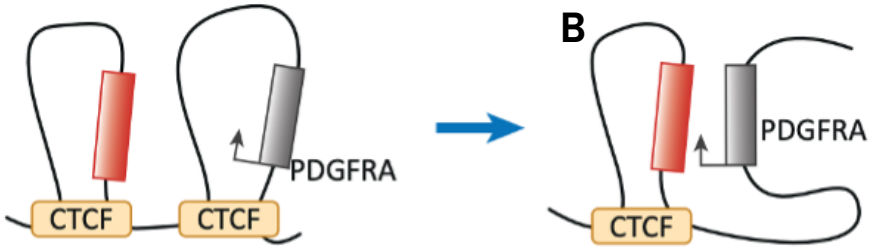

Figure 35. Schematic showing the effects CTCF gain and loss on TAD boundaries. A. Schematic of TAD boundaries that are intact with CTCF acting as insulators that prevent aberrant chromatin interactions between the enhancer and gene region. B. Schematic of disrupted TAD boundaries that loss a CTCF insulators through epigenetic modulations like methylation thus resulting in the formation of aberrant chromatin interactions between the enhancer and gene region from two adjacent TADs. The grey rectangle represents the gene location and the red rectangle represents the enhancer. The CTCF binding sites are indicated as "CTCF in a box". The schematic is adapted from my publication (See et al., 2019).

A key difference between TADs and FIREs is that TADs are often reported to be stable across different cell types while FIREs are seen to occur at regions that are near cell type-specific genes and dynamically altered according to cell type requirements. For instance, FIREs that are seen in the GM12878 B lymphoblastoid cell line are distinctly different in presence in the brain cells (Schmitt et al., 2016a). In addition, up to 90% of FIREs are reported to be present within chromatin loops that are mediated by CTCFs and that not all insulated chromatin loops or TADs contains FIREs (Schmitt et al., 2016a). The identification of the *MEIS1* FIRE at the haemotopoietic stem cell specific *MEIS1* gene region supports the above reported feature that FIREs are located at cell-type specific gene sites. As *MEIS1* is significantly downregulated upon cell differentiation, I speculate that the *MEIS1* FIRE could be potentially disrupted to a certain extent during cell maturation, leading to the downregulation of *MEIS1*

in line with my observed results that showed *MEIS1* downregulation upon the CTCF binding site excision at the CTCF FIRE in the K562 cells.

Taken together, the chromatin loop alterations occurring upstream and downstream of the *MEIS1* FIRE upon the CTCF binding site excision suggests that CTCF binding may regulate FIREs in a differential way to TADs, in the sense that the loss of the CTCF insulation results in chromatin loop reorganisations at both ends of the *MEIS1* FIRE and not result in the formation of chromatin loops in a unidirectional manner that extends towards adjacent compartments as was reported with one-sided CTCF boundary loss in TADs. This also reiterates that the level of insulation that is provided by CTCF and possibly cohesin proteins are of different levels in FIREs and TADs. However, not much is known about the role of CTCF in FIRE and thus more experiments in CTCF manipulations in FIREs have to be conducted to verify the validity of this speculation. Since a FIRE may contain multiple CTCF binding sites, some of such additional experiments may include the deletion of CTCF sites individually and collectively in the FIRE to understand if the loss of collective CTCF sites has a larger effect on the disruption of FIREs and their associated loops. In this way, we will be able to understand if CTCF are able to function independently of each other within a FIRE as we observed with a single CTCF deletion in TADs or if CTCFs within a single FIRE have a dependent relationship on each other to ensure that overall chromatin loops within a FIRE are stable. The strength of insulation by CTCF boundaries in FIRE and TADs can also be further examined via insulation score analysis to allow for a deeper understanding of how differentially FIRE and TADs are insulated in the genome.

Beside acting as insulators for 3D genome compartmentalization, CTCF also functions as a transcription factor for gene regulation. A recent study by Thiecke et al. (2020) demonstrated that CTCF and its associated cohesin proteins are essential for the maintenance of enhancer-promoters chromatin interactions and gene expression. The depletion of CTCF and cohesin resulted in the rewiring of the promoter-enhancer loops and the consequential alteration in genes transcription, thus suggesting that CTCF plays an important role in modulating steady-state gene transcription under different cellular conditions (Thiecke et al.,

2020). These findings along with my results from the deletion of the CTCF binding site at the *MEIS1* FIRE suggest that CTCFs have a role in regulating internal chromatin loops within TADs & FIREs as well as loops that serve as insulators for genome compartmentalization. As such these results may explain for the chromatin loops reorganisation that is seen upon CTCF binding site deletion. Therefore, I speculate that the presence of CTCF proteins at specific locations of the FIRE may function as scaffolds in partially confining the chromatin interactions in a specific pattern within the FIRE.

In summary, these results support my hypothesis that the excision of the CTCF binding site within the *MEIS1* FIRE will result in the alteration of chromatin loops and consequent downregulation of *MEIS1* expression. In addition, these chromatin changes and downregulation of *MEIS1* expression in CTCF excised K562 cells, also resembles what is observed with the TAD and *MEIS1* expression profiles of AML 28 and AML30 patients who lack the *MEIS1* FIRE. Interestingly, these low expression levels of *MEIS1* are accompanied by similar levels of *HOXA9* downregulation in both patients. The downregulation of both *HOXA9* and *MEIS1* oncogenes expression levels in AML 28 and AML 30 patients are in agreement with earlier findings that suggested that HOXA9 and MEIS1 oncogenic proteins work in tandem to induce aggressive myeloid leukaemia development (Lawrence et al., 1999, Li et al., 2013b). However, since only a small number of samples were analysed, more samples of AML patients need to be examined for these observations to be established.

Although previous work has been done to study the effects of *MEIS1* abrogation in leukaemia, to my knowledge, this is the first study that has experimentally addressed how chromatin interactions and super-enhancers in a FIRE contribute to the expression of the *MEIS1* oncogene in myeloid leukaemia. Furthermore, this is also one of the first few studies that examines the regulation of a FIRE in detail through experimental means such as with CRISPR Hi-C and 4C. Therefore, my work which demonstrated that the *MEIS1* FIRE and its associated super-enhancers are important in ensuring the continuous expression of *MEIS1* in K562 myeloid leukaemia will be beneficial towards the development of treatments for myeloid leukaemia with high *MEIS1* expressions. As the research into FIREs is

a relatively new area, more work has to be done to assess if different regions of FIREs have differential degrees of reliance on CTCF binding to maintain the chromatin interactions and gene transcription in that locus.

6.3 Dysregulated DNA methylation could affect CTCF binding patterns and FIRE stability in cancers

Despite the increase in knowledge of TADs in leukaemia development, the reason for the observed heterogeneous presence of the *MEIS1* FIRE in some AML patients remains to be fully understood. Recently, a study by Mujahed et al. (2020b) suggested an association between AML development and aberrant CTCF occupancies at regions that are enriched with myeloid transcription factors binding sites. Mujahed and colleagues also discovered that CTCF knock-down resulted in more cell differentiation, thus signifying the potential downregulation of hematopoiesis related genes such as *MEIS1* upon CTCF binding loss which subsequently leads to increased cell differentiation. As, certain groups of AML patients like those with *TET* mutations potentially possess higher concentrations of CTCF expression as compared to the other groups of AML patients with other types of mutations (Mujahed et al., 2020a), these findings may explain for the heterogeneous presence of the *MEIS1* FIRE in some AML patients. Moreover as I mentioned above, *MEIS1* expression have been noted to be downregulated upon cell differentiation (Ramos-Mejía et al., 2014). Taken together with my experimental results from the CTCF excision at the *MEIS1* locus, these findings support the suggestion that CTCF binding is critical toward sustaining of chromatin loops towards the *MEIS1* oncogenes and also implies that these CTCF binding at the *MEIS1* FIRE may be lost upon haematopoietic stem cells differentiation but not in myeloid leukaemia cells.

Azacytidine is a common drug used for the treatment of AML and has also been shown to upregulate CTCF levels (Mujahed et al., 2020a), thus implying that these drugs may work in part to regulate chromatin interactions in myeloid leukaemia. Azacytidine functions as a DNMT enzyme inhibitor by binding irreversibly to the DNMT enzyme, thus preventing the deposition of methylation marks in the genome DNA (Christman, 2002). For instance, the induction of pancreatic ductal adenocarcinoma cells with Azacytidine resulted in

demethylation effects on these cells' genomes and consequently restricted pancreatic ductal adenocarcinoma tumour growth in vivo (Gailhouste et al., 2018). Altogether, the use of drugs like Azacytidine that promotes genome demethylation may also affect CTCF binding patterns at the genomic level and can be applied as a sole treatment or potentially combined with alternative drugs in the treatment of myeloid leukaemia. Based on the above examples of CTCF and DNA methylation binding patterns, DNA methylation and CTCF binding are suggested to be antagonistic towards each other and thus the loss of methylation at certain genomic region will promote the binding of CTCF proteins, which possibly could lead to new genome compartmentalization and altered gene expression in that region.

Since FIREs are recently discovered features of 3D genome organization, the mechanisms behind their regulation and occurrence remain largely unknown. Therefore, the heterogenous presence of the *MEIS1* FIRE and its role in myeloid leukaemia progression require further experimentation. However, some possible explanations for the heterogeneous presence of FIRE in myeloid leukaemia can be drawn based on our current understanding of epigenetic modulator mutations that affect DNA methylation levels in this disease.

For example, *DMNT3A*, *IDH1/2* and *TET* mutations that are mentioned above are common mutations that occur in myeloid leukaemia (Weissmann et al., 2012, Wagner et al., 2010, Gross et al., 2010, Ley et al., 2010). Furthermore, their mutations are heterogeneously observed in myeloid leukaemia patients. Since DMNT3A, IDH1/2 and TET complexes modulate the methylation status of the genome, their mutations could suggest that CTCF/cohesin binding are affected in different levels between individuals with myeloid leukaemia. As I discussed above, IDH1/2 and TET2 function in antagonistic fashion to DMNT3A by promoting the demethylation of genomic DNA regions (Tahiliani et al., 2009, He et al., 2011, Ito et al., 2010, Xu et al., 2011). Since DNA methylation in the genome has an effect on CTCF/cohesin binding and the formation of TAD/FIRE boundaries, these dysregulated methylation/demethylation patterns that are resultant of DMNTs, IDHs and TET2 mutations (Flavahan et al., 2016) may thus contribute to abnormal TADs, FIREs formation in myeloid leukaemia.

For instance, AML patients who possess mutations at the *IDH* gene presented a genomic profile of global hypermethylation, with hypermethylation of gene promoters that are accompanied with silenced expressions (Figueroa et al., 2010). Therefore, I speculate that aberrant DNA methylation/demethylation patterns that are derived from heterogeneous mutations of epigenetic modulators such as IDH and TET may explain for the selective development of FIREs such as that of *MEIS1* FIRE in individual myeloid leukaemia patients.

In addition, the Nucleophosmin 1 protein (NPM1) was reported to participate in protein-protein interactions with the CTCF protein (Wang et al., 2020). The NPM1 phosphoprotein is understood to participate in a variety of key cellular metabolic activities such as the regulation of chromatin condensation for the purpose of gene transcription as well as modulate gene transcription by acting as a transcriptional partner to proteins such as YY1 and NFkB (Grisendi et al., 2006). It also has the ability to travel between the cell nucleus and cytoplasm and thus can function as a chaperone protein to bring molecules across these two cellular locations (Yun et al., 2003, Okuwaki et al., 2001). However, *NPM1* is also widely observed to be mutated in myeloid leukaemia with its mutated form of protein known as NPM1c (Falini et al., 2007).

Interestingly, (Wang et al., 2020) reported that the NPM1c protein is able to bind to CTCF proteins and subsequently carry them away from the cell nucleus to its cytoplasm. This aberrant transportation of CTCF proteins to the cytoplasm will result in the reduction of CTCF proteins in the nucleus and may thus lead to the dysregulation of CTCF mediated genome organization such as with FIREs and TADs. The study also reported that the disruption of the NPM1c protein leads to the recovery of CTCF proteins within the nucleus (Wang et al., 2020), thereby reiterating their potential role in CTCF protein regulation in myeloid leukaemia patients with the *NPM1* gene mutation. However, the full impact that CTCF reduction in the nucleus has on FIREs and TADs formation/loss in myeloid leukaemia remains to be further understood.

6.4 TADs and chromatin interactions are largely unaltered at the *HOXA9* locus, but are hijacked by newly acquired super-enhancers to promote *HOXA9* overexpression in AML

In the second part of this project, I focused on investigating alternative modes of *HOXA9* regulation via chromatin interactions involving the *HOXA9* promoter and enhancers in cis. I also seek to understand the mechanisms behind the co-expression of *HOXA9* and *MEIS1* oncogenes during AML pathogenesis.

As I reported above, my analyses of six clinical CD34+ samples (three CD34+ AML & three CD34+ normal knee samples) at the *HOXA9* region, showed no significant alterations of TADs and chromatin loops between the samples. This observation is in stark contrast to the *MEIS1* region, which showed the loss of the *MEIS1* FIRE in the samples AML 28 and AML 30. My 4C experiments with THP-1 and HL-60 demonstrated that chromatin interactions that involve the *HOXA9* promoter, remain largely unaltered in these two AML cell lines. These suggest that chromatin interactions at the *HOXA9* locus are consistently conserved between AML subtypes and hematopoietic cell phases. A possible explanation for my above observations could be that chromatin interactions function as scaffolds for gene regulation and genome organization. For instance, TADs are found to be largely stable and non-altered between cell types (Dixon et al., 2015).

However, I speculate that these chromatin loops may function to attract transcription factors during different cell phases that require the expression of certain genes such as *HOXA9* during the haematopoietic stages which will become downregulated upon cell differentiation. These hypothesized concept supports my 4C results at the *HOXA9* promoter, as the chromatin loops of my clinical normal Knee CD34+ and AML CD34+ samples and the THP-1 and HL-60 cell lines do not seem to be altered in comparison to each other.

Furthermore, my ddPCR results for THP-1 and HL-60 cells also exhibited high levels of *HOXA9* expression that are specific to THP-1 but poorly expressed in HL-60. Collectively, my Hi-C analyses from the clinical normal Knee CD34+, AML CD34+ samples and 4C experiments from the THP-1 and HL-60 cell lines,

strongly suggest that chromatin interactions at the *HOXA9* region does not strongly vary between the CD34+ and AML samples. Despite their similar pattern of chromatin interactions to the *HOXA9* promoter, a possible reason for the difference in *HOXA9* expression levels between the THP-1 and HL-60 cell lines can be explained by their differing AML subtypes. Based on the French-American-British (FAB) classification for leukaemia, THP-1 falls under the M5 subtype while HL-60 falls under the M3 subtype. In addition, previous studies have also demonstrated genetic and phenotypical differences between AML subtypes (Haferlach et al., 2005, Rose et al., 2017). Therefore, my results suggest that alternative means of *HOXA9* gene regulation may be responsible for the upregulation of *HOXA9* in THP-1 and certain subsets of leukaemia patients. These findings also agree with my speculated explanation that chromatin loops do not change between cell types but gain specific transcription factors for gene expression at different phase of the cell cycle.

Since no significant alterations are seen at the TAD and chromatin loop at the *HOXA9* locus, I proceeded to investigate if super-enhancers are gained and can contribute to the elevated levels of *HOXA9* expression that is seen in myeloid leukaemia. My analyses of H3K27ac ChIP-seq data from 58 AML MNC samples and 2 normal CD34+ samples (McKeown et al., 2017b), show the presence of three groups of super-enhancers (*SNX10* SE, SE2 & SE3) that are located upstream of the *HOXA9* gene in the AML samples. I also noticed that among these three groups of super-enhancers, only one super-enhancer (SE2) can be seen in the normal CD34+ control samples, while *SNX10* and *SE3* are absent. The overall analyses of these three super-enhancers revealed that each of these super-enhancers is only present in a subset of the AML patients (numbers indicated in the bottom boxes of **Figure 31**). Upon further examination of the AML MNC and normal CD34+ super-enhancer profiles, I also noticed that the *SNX10* SE and SE3 are specifically present in the AML samples but absent in the normal CD34+ samples. To verify if these three groups of super-enhancers are also present in blood lineage associated cell lines, I performed the same super-enhancer analysis on published H3K27ac ChIP-seq of the cell lines GM12878, HL-60, K562 and THP-1. Interestingly, only the AML cell line THP-1 possess the presence of the *SNX10* SE, while the other three cell lines do not show the

existence of any of these three super-enhancers. The heterogeneous presence of each super-enhancer in selected AML patients and the specific presence of *SNX10* SE in only THP-1 but not HL-60, suggest that the pattern of super-enhancers acquisition may be unique to each subtype of AML. Furthermore, the alignment of my super-enhancer analyses to the Hi-C contact heatmap and chromatin loops from the clinical CD34+ samples and cell lines, show-specific chromatin interactions between the super-enhancers and the *HOXA9* promoter. These findings suggest that the *SNX10* SE and SE3 could be acquired during AML pathogenesis to drive *HOXA9* overexpression via the hijacking of preexisting chromatin loops. In addition, my study is also among the few that present 4C data of chromatin interactions between three novel super-enhancers and the HOXA9 promoter in AML cells.

Such hijacking of preexisting chromatin interactions was also report in a study involving the binding of the glucocorticoid receptor to preformed chromatin loops after the induction of glucocorticoid hormone treatment (D'Ippolito et al., 2018). D'Ippolito et al. (2018), demonstrated that the activation of glucocorticoid receptors does not lead to new chromatin loop formations, but instead glucocorticoid receptors act on preexisting chromatin loops to induce genes transcription.

Super-enhancer acquisition was also observed in triple negative breast cancer. This form of breast cancer is known to be heterogenous in its expression of molecular driver gene. A study by Raisner et al. (2020) identified at least 781 uniquely acquired super-enhancers that were otherwise absent in normal breast cells. In addition, they found that a cell growth associated oncogene *BAMBI* (BMP and Activin Membrane Bound Inhibitor) is dependent on the acquisition of enhancers and super-enhancers in its vicinity (Raisner et al., 2020). The inhibition of these enhancers that are proximal to *BAMBI* resulted in the repression of *BAMBI* expression in the breast cancer cell line HCC38 (Raisner et al., 2020).

Their results are similar to my observation of super-enhancers acquisition at regions that possess chromatin interaction to the *HOXA9* promoter. In addition to the Hi-C, 4C and super-enhancer analyses, I examined the relationship between the *SNX10* SE and *HOXA9* expression in the group of 58 AML patients. My analysis revealed a strong percentage of association between *SNX10* SE presence and elevated *HOXA9* expression levels. My analysis revealed that approximately 94% of AML patients who possess the *SNX10* SE, also express high *HOXA9* expression levels that were higher than that of normal CD34+ haematopoietic stem cells and normal CD33+ cells. Taken together, these results support my hypothesis that acquired super-enhancers may be driving the overexpression of *HOXA9* in specific subsets of AML. Moreover, as I mentioned above, these results also contribute to the small pool of information for myeloid leukaemia development, that super-enhancers acquisition is a possible reason for the overexpression of oncogenes such as *HOXA9* and *MEIS1* in myeloid leukaemia.

6.5 HOXA9 and MEIS1 proteins can bind at each other's promoters to contribute to coordinated *HOXA9* and *MEIS1* co-expression via a feedback loop

HOXA9 and MEIS1 proteins are co-expressed and shown to be biomarkers for poor prognosis in myeloid leukaemia (Gao et al., 2016). In agreement with previously reported observations, my ddPCR results show that *HOXA9* and *MEIS1* expression levels are upregulated in normal CD34+ hematopoietic stem cells and interestingly, *MEIS1* expression levels are significantly upregulated in comparison to the *HOXA9* expression levels.

However, the examination of the AML MNC samples shows a different trend of *HOXA9* and *MEIS1* expression levels, where no significant differences are observed between their levels of expression. This lack of significant differences between *HOXA9* and *MEIS1* expressions in AML patients may be explained by the overexpression of *HOXA9* in AML and the unchanged levels of *MEIS1* expression that continues to be expressed at similar levels as the haematopoietic CD34+ stem cells. These observations also highlight the validity of my earlier proposed model for the regulation of *MEIS1* expression via the *MEIS1* FIRE in

AML. Since the *MEIS1* FIRE is not loss in some AML patients (As observed in AML29), *MEIS1* expression levels will remain to be high alongside *HOXA9* expression levels in AML. These findings are supported by my ddPCR results of AML MNC, normal knee CD33+ and normal knee CD34+ cells show no significant differences in *MEIS1* expression between the AML MNC and normal knee CD34+. This observation agrees with my argument that *MEIS1* levels are consistently maintained in the CD34+ and AML cells.

Next, I investigated the expression levels of *HOXA9 in* AML MNC, normal CD33+ and Knee CD34+ samples. The reason for evaluating the expression analysis in normal knee CD33+ and normal knee CD34+ cells is because *HOXA9* have been reported to be downregulated upon differentiation from the hematopoietic stem cell stage (Ramos-Mejía et al., 2014). Therefore, the CD33+ cell marker represents as an indicator of more mature cells within the blood maturation process (Moore et al., 2016) and can serve as a good comparison to assess if AML cells are indeed upregulated for *HOXA9* expression. The ddPCR results of *HOXA9* expression levels in the AML MNC, normal CD33+ and normal Knee CD34+ cells show a different trend to that of *MEIS1* expression, with *HOXA9* expression levels being significantly higher than the normal CD33+ and normal CD34+ cells. Altogether, my ddPCR findings suggest that *HOXA9* expressions are upregulated at higher levels as compared to the CD34+ hematopoietic stem cells. This may also suggest that an additional mechanism of increasing *HOXA9* expressions may be gained during AML pathogenesis.

Based on these combined observations of *HOXA9* and *MEIS1* co-expression and upregulation in myeloid leukaemia as well as the discovery of newly acquired super-enhancers like the *SNX10* SE that interacts with the *HOXA9* promoter. I hypothesized that HOXA9 and MEIS1 proteins can bind to each other's promoter as well as the promoters of downstream oncogenes such as *MYC* to regulate aberrant oncogenes expression in myeloid leukaemia in a circuitry manner. My ChIP-qPCR experiments on THP-1 cells with HOXA9 and MEIS1 protein pulldowns revealed that the HOXA9 and MEIS1 proteins bind to both their own promoters as well as the promoter of each other.

Based on this understanding of HOXA9 and MEIS1 protein binding patterns at their respective promoters as well as the observed trend of their co-expression or co-reduction in subsets of AML patients, specific treatments that are targeted towards this group of patients may be considered. For instance, the tyrosine kinase inhibitor imatinib has been used for the treatment of chronic myeloid leukaemia patients (Quintás-Cardama and Cortes, 2009), but with limited success due to the development of drug resistance (Ren, 2005). Poor responses to imatinib were also found to be correlated to patients with high *MYC* expression (Albajar et al., 2011). However, this drug can be considered for use in the treatment of patients with the loss of the *MEIS1* FIRE and low *MEIS1* expression. Since, *MYC* levels and its downstream target pathways have been suggested to be downregulated in the instance of *MEIS1* FIRE loss, the application of drugs such as imatinib which have stronger effects on low *MEIS1* and *MYC* expressing patients can be further explored.

My findings are also supported by a previous study that reported the co-binding of HOXA9 and Meis1 in transformed myeloblastic mouse cells. Huang et al. (2012), reported the identification of genome wide HOXA9 and Meis1 co-binding at several hundreds of highly conserved sites in those myeloblastic transformed mouse cells. Furthermore, my ChIP-qPCR results and analysis on previously published data (Zhong et al., 2018) of HOXA9 ChIP-seq on an AML cell line MV4;11 revealed that HOXA9 binds to its own promoter and that of *MEIS1* and *MYC* promoters in both THP-1 and MV4:11 AML cell lines (data not shown). In addition, my ChIP-seq analysis from the MV4;11 AML cell line, also shows that HOXA9 binds at SE2 and SE3 that possess preexisting chromatin loops with the *HOXA9* promoter.

These observations reiterate that HOXA9 and MEIS1 may function via positive feedback regulatory loops during AML pathogenesis by binding to their own and each other's promoters. The transcription of either one of *HOXA9* or *MEIS1* oncogenes may thus lead to the simultaneous expression of their partner protein.

Furthermore, the binding of HOXA9 and MEIS1 proteins at each other's promoters suggest that they are involved in regulating each other's transcription in leukaemia via a possible upstream master regulator like the *MEIS1* FIRE.

Another example of co-expressed and upregulated oncogenes pairs in AML are *SKI* and *c-MYB* (Frech et al., 2016). Frech et al. (2016) performed an analysis of MYB transcription factor binding sites in AML and revealed that MYB binding sites are located in the *SKI* gene regulatory region. Furthermore, the knockdown of *c-MYB* in leukaemia cells also led to the downregulation of MYB and SKI protein levels, thus demonstrating the role that MYB has in regulation *SKI* expression in leukaemia (Frech et al., 2016). The author also found that MYB and SKI proteins and functions were positively correlation in AML patients and suggested the presence of mutual transcription regulation between *c-MYB* and *SKI* oncogenes (Frech et al., 2016).

6.6 Super-enhancer can promote oncogenes overexpression in myeloid leukaemia

Considering that, my 4C and clinical Hi-C results have shown that *SNX10* SE and the *HOXA9* promoter interact via chromatin loops, the presence of bound HOXA9 protein at regions strongly suggest that the chromatin loop functions to bring the HOXA9 proteins to both the SE and the *HOXA9* promoter region for the transcription of *HOXA9* in certain groups of AMLs. However, *SNX10* SE deletion or inhibition of its associated chromatin interactions have to be conducted for firm conclusions to be drawn.

Lastly, Sun et al. (2018b) demonstrated that *HOXA9* upregulation in leukaemia results in the reprogramming of genome-wide enhancer landscape. In this study, Sun and colleagues revealed that sites with HOXA9 proteins binding were highly enriched for enhancer related H3K4me1 and H3K27ac histone marks in both myeloid and B cell lymphoblastic leukaemia. Moreover, they also observed that the HOXA9 proteins bind to a similar set of genetic landscapes in myeloid and B cell lymphoblastic leukaemia. These suggest that HOXA9 protein binding is targeted toward enhancers that are specific to the regulation of oncogenes in leukaemia.

They also revealed that leukaemic conditions with elevated levels of *HOXA9* expression result in the acquisition of new leukaemia specific enhancers with high H3K4me1 histone marks as compared to normal hematopoietic stem cells

profiles. These new enhancer regions are also enriched for HOXA9, which was otherwise absent in the hematopoietic stem cells. Furthermore, they also found that the genes that are regulated via these new enhancers showed increase in gene expression when *HOXA9* expression was upregulated.

These suggest that the HOXA9 protein binds to the new enhancer regions to upregulate leukaemia related genes. Furthermore, these reported results strongly resemble the mode of how *SNX10* SE potentially regulate *HOXA9* overexpression in subsets of AMLs via chromatin loops to the *HOXA9* promoter. These support my hypothesis that acquired super-enhancers such as the *SNX10* SE contribute to elevated *HOXA9* expression in myeloid leukaemia. In another example of how an oncogene can self-regulate its own gene expression by controlling its associated enhancers, Wang et al. (2014), demonstrated that MEIS1 proteins are able to bind to its own enhancer to regulate *MEIS1* expression. However, the discovery of novel enhancers and super-enhancers that regulate co-expressed genes via chromatin interactions remain limited. Thus, my discovery of the *SNX10* SE which forms chromatin loops with the *HOXA9* promoter and possess the binding of HOXA9 and MEIS1 proteins between can bring improved understanding towards how these two oncogenes are regulated in myeloid leukaemia.

7 Future Work

Some questions on how *MEIS1* FIRE and its associated chromatin interactions can affect *MEIS1* expression remains to be answered. These questions include: what is the parentage of AML patients with the presence of a *MEIS1* FIRE and is there any AML subtypes that are more associated with the *MEIS1* FIRE? What is the degree of influence that individual super-enhancers such as SE1, 2 & 3 have on *MEIS1* expression? And do individual super-enhancers have a bigger influence than another on *MEIS1* overexpression? Or is the collective presence of all three super-enhancers required for *MEIS1* expression in myeloid leukaemia?

7.1 Questions on the prevalence of *MEIS1* FIRE in AML patients and AML subsets that are strongly linked to *MEIS1* FIRE presence

An aspect for further investigation is on how many percent of AML patients possess the *MEIS1* FIRE and high *MEIS1* expression levels and if these patients with such characteristics belong to a specific subtype of AML. To answer this question, more clinical AML samples have to be sequenced with Hi-C seq and tested for *MEIS1* expression using ddPCR for us to gain a better understanding of the demographic of *MEIS1* FIRE and *MEIS1* expression among these AML patients. Further clinical subtyping has to be performed along with the Hi-C sequencing and ddPCR for us to associated on the types of patients with these *MEIS1* FIRE features and if they fall under specific AML subtypes. At the same time, translocations and specific gene mutations that are present in AML patients, especially those with the presence of high *MEIS1* expression and the *MEIS1* FIRE should be observed to let us gain a better understanding of potential biomarkers that could be related to such cases of AML. In addition, the question on if *MEIS1* FIRE is present in more matured cells can also investigated by performing Hi-C seq and FIRE analyses in peripheral bone marrow cells. Due to insufficient CD33+ clinical samples availability at the time of the experiment, more clinical samples of the mature CD33+ AML and normal CD33+ knee cells can be collected in the future and be tested with ddPCR to assess for their *MEIS1* expression levels.

The effect of *MEIS1* downregulation as a cause for cell maturation or as a consequence of cell maturation can also be investigated with knockout/knockdown experiments of *MEIS1* expression in haematopoietic stem cells followed by the observation of maturation markers development in these knockdown/knockout stem cells. In this way, we can gather more information on the role of *MEIS1* in haematopoiesis and stem call maturation.

7.2 Questions on the significance of CTCF & chromatin interactions for the maintenance of FIREs and their associated gene expression

Another aspect for further investigation is on the role of different CTCF boundaries in maintain a FIRE in the genome. In this study, I investigated the effects of the removal of a CTCF binding site that is downstream of the *MEIS1* genome at the 3' end of the *MEIS1* FIRE. This excision of the downstream CTCF binding site resulted in the downregulation of *MEIS1* and more importantly significant alterations of chromatin interactions that are associated with the *MEIS1* promoter and potential abrogation of the *MEIS1* FIRE in the K562 myeloid leukaemia cells. *MYC* expression was also observed to be downregulated, with downstream *MYC* gene targets suggested to be downregulated in expression in the Gene Set Enrichment Analysis. However, as these analyses were performed with two clones of CTCF KO cells, it will be important to include additional clones of RNA-seq into the analysis to get a more conclusive outcome on the genes that are strongly downregulated due to reduced *MYC* expression in the CTCF KO cells.

Furthermore, the investigation of the roles of the newly formed MEIS1 promoter associated loops in the CTCF binding site deleted cells will also be important in evaluating the impact and effect of these new loops on oncogenes expression in myeloid leukaemia. For example, an interaction was shown to be gained between the *MEIS1* promoter and the *EML4* gene. The *EML4* gene has been reported to be associated with non-small cell lung cancer. Preliminary check of the differential expressed genes from my RNA-seq data shows that *EML4* expression levels are reduced in the CTCF KO cells, but more work such as the validation of the *EML4* gene expression in the CTCF binding site deleted cells as well as 4C experiments with the *EML4* gene as the bait have to be conducted for us to

better understand the effect of this newly gained interaction and their effect on regulating other oncogene expression in its vicinity.

These investigations can also be extended to normal and *MEIS1* low AML clinical samples to examine if these new loops that are interacting with the *MEIS1* promoter are also present in them. An example of this is that these normal and low *MEIS1* AML samples can be studied for their *MEIS1* promoter associated loops through the use the Episwitch technique that allows for the study of interactions in clinical samples with limited cell number availability.

It is also imperative for further work to focus on examining the effects of the other CTCF boundary that is located upstream of the *MEIS1* gene, at the 5' end of the *MEIS1* FIRE. This is because that will allow us to gain and understand if this specific CTCF that I have excised have more control over the maintenance of the *MEIS1* FIRE and its associated chromatin loops or if the loss of the other CTCF binding site at the upstream boundary of the *MEIS1* FIRE will have similar or different effects of disrupting the *MEIS1* FIRE and the chromatin loops within. Thus, the role of the upstream CTCF binding site within the *MEIS1* FIRE can be revealed by conducting the above CRISPR-Cas9 mediated experiments that is targeted at the upstream CTCF binding site in the *MEIS1* FIRE.

4C, Hi-C and ddPCR experiments of *MEIS1* and *MYC* expressions can be performed after the excision of the upstream CTCF boundary to examine if the loss of the upstream CTCF boundary will lead to similar results as observed with my earlier experiment or if different chromatin alterations will occur.

Since the K562 cell line does not express *HOXA9*, the evaluation of *HOXA9* expression upon CTCF binding site is thus not possible. Therefore, to investigate the effects of CTCF binding site deletion on *HOXA9* expression in AML, the CTCF excision on the 5' and 3' ends of the *MEIS1* FIRE can be done on an AML cell line such as THP-1 that express *HOXA9* and *MEIS1* to investigate if the disruption of the *MEIS1* FIRE and reduced *MEIS1* expression will have a significant effect on *HOXA9* expression in AML. The use of lentivirus at higher viral titres can also be used for the CRISPR-Cas9 excision of the CTCF binding

site at the *MEIS1* FIRE in difficult to transfect cell like THP-1 cells as high titre of lentivirus has a higher chance of success with CRISPR-Cas9 experiments in hard to transfect cell lines

FIRES are partly characterized by their occupancy near genes that are specific to cell stage and type. Therefore, the presence of FIREs at other oncogenes in myeloid leukaemia should be investigated into. For example, oncogenes such as *PBX3* and *FLT3* which were reported to be critical players and partners of *MEIS1* in myeloid leukaemia may also be regulated via FIREs.

In my preliminary analysis of other oncogenes regions in clinical myeloid leukaemia samples, I observed that some AML patients possess heterogenous presence of FIRE at the *PBX3* gene locus. Interestingly, super-enhancers were also seen to be located at the postulated *PBX3* FIRE region, which suggest that *PBX3* expression may involve chromatin loops within the FIRE. However, more AML samples have to be sequenced using Hi-C seq and RNA-seq to determine if the *PBX3* region is indeed a FIRE and if so, to what degree does this *PBX3* FIRE contribute to *PBX3* expression in myeloid leukaemia.

In order to address this question about the role of chromatin interactions with the *PBX3* FIRE, the excision of chromatin loops within the potential *PBX3* FIRE can be performed to determine if any specific chromatin loops in the potential FIRE hold more influence over *PBX3* expression than the other loops. These experiments will allow us to gain a better understanding of FIRE occupancies in cancer and if FIREs are only present in selected oncogenes in diseases.

7.3 Questions on the dysregulated methylation and aberrant CTCF binding patterns in the genome and their effect on oncogenes expression in cancers

An area for further investigation could be to determine if CTCF bindings are consistent across different FIREs at different genomic regions or if these CTCF binding patterns to FIREs are specific to certain genomic areas such as the active compartments within the 3D genome organization.

Further work through the inhibition of these FIRE associated super-enhancers inhibition via gene therapy techniques such as with double stranded DNA targeting decoys or with super-enhancer targeting drugs such JQ1 will be useful for the development of targeted regions for therapy in myeloid leukaemia. For instance, DNA targeting decoy is a technique that is still in its infant stage of development but represents a potential way in which DNA transcription can be inhibited. These DNA targeting decoys function through the concept of competitive binding thus preventing transcription factors from gaining access to DNA genome such as gene promoters and super-enhancer, thereby restricting transcription (Takakura et al., 2019). In an example, Cogoi et al. (2013) designed a protein decoy that mimic a form of G-quadruplexes and introduced it into pancreatic cancer cells that carries the *KRAS* mutation. They then observed that this decoy inhibits *KRAS* expression and resulted in the reduction of pancreatic cell growth (Cogoi et al., 2013). As these decoys are designed to mimic specific transcription factors DNA binding sites, they can be used to prevent transcription factors such as HOXA9, MEIS1 and MYC from binding to the super-enhancers and promoter anchors within a chromatin loop, consequently disrupting the effects of these super-enhancers on *HOXA9* and *MEIS1* overexpression in myeloid leukaemia.

7.4 Questions on the effects of super-enhancers, chromatin interactions and their associated gene transcription in FIREs

In order to address the question on the degree of influence that individual super-enhancer has on *MEIS1* expression, CRISPR-Cas9 experiments on each of the three super-enhancer (SE1,2 &3) regions with subsequent chromatin interaction experiments such as 4C and Hi-C as well as RNA-seq have to be conducted in myeloid leukaemia cell lines such as THP-1 and K562 to understand the effects of each super-enhancer on *MEIS1* expression. By performing the excision of individual super-enhancer regions, we will be able to assess if individual super-enhancers possess a stronger effect on *MEIS1* overexpression and if the loss of individual super-enhancers within the *MEIS1* FIRE will lead to reorganization of chromatin loops to the *MEIS1* promoter.

Furthermore, the removal of all three super-enhancers (SE1 ,2 & 3) via CRISPR-Cas9, will also reveal if these three super-enhancers presence are collectively required to maintain *MEIS1* expression levels or that *MEIS1* expression levels in myeloid leukaemia are also dependent on additional chromatin loops between upstream and downstream distal regions of the *MEIS1* locus and the *MEIS1* promoter. For instance, the removal of all three super-enhancers with significant or complete abrogation of *MEIS1* expression will suggest that *MEIS1* expression is strongly reliant on these super-enhancers that are located within the *MEIS1* FIRE. However, if *MEIS1* levels are not significant reduced upon the loss of these three super-enhancers, further work involving the distal regions that forms chromatin loops to the *MEIS1* promoter should be investigated for their effect on *MEIS1* expression in myeloid leukaemia. In addition, the CRISPR and 4C experiments on selective distal regions of chromatin interactions will also allow us to better understand if the individual loss of chromatin interactions to the *MEIS1* promoter will result in alteration to *MEIS1* expression.

7.5 Questions on the prognosis of AML patients with the loss of *MEIS1* FIRE & the potential of utilizing existing oncology drugs for such groups of patients

The observed reduction in *MEIS1* expression upon the excision of its downstream CTCF binding site matches the expression levels of *MEIS1* in AML patients with the loss of *MEIS1* FIRE in their genome. This finding suggests that the AML patients who showed a loss of *MEIS1* FIRE may also experience lower expressions of oncogenes such as *MYC* and its downstream targets as well as slower leukaemia cells growth. Another question that may be asked is "Do patients who show a loss of the *MEIS1* FIRE have a better prognosis as compared to those who possess the FIRE?" Since *MEIS1* and *MYC* are widely reported biomarkers for poor prognosis in myeloid leukaemia, the downregulation of these two genes with the disruption of the *MEIS1* FIRE may result in differing prognosis outcomes for such myeloid leukaemia patients who have loss the *MEIS1* FIRE. Therefore, further work towards the determining the prognosis of this group of myeloid leukaemia patients without the *MEIS1* FIRE is essential.

As the disruption of *MEIS1* FIRE resulted in the reductions of *MEIS1* and *MYC* oncogenes in the K562 cells, an area for deeper exploration may be to access if these myeloid leukaemia cells with disrupted FIREs and chromatin loops will be more susceptible to certain drug treatments. Drugs such as imatinib which induced resistance in patients with high *MYC* expression (Albajar et al., 2011) may then be more effective in patients with lower *MEIS1* and *MYC* expression that is accompanied with a disruption of *MEIS1* FIRE.

7.6 Questions arising to the presence of *MEIS1* FIRE in different types of cancers

Besides promoting myeloid leukaemia development, *MEIS1* was reported to be involved in a variety of cancers and organ development such as in colorectal, breast, non-small-cell lung cancer (Li et al., 2014) and the development of the mammalian heart (Mahmoud et al., 2013). An isoform of the MEIS1 protein was found to be commonly downregulated in up to 83% of colorectal cancer patients, with its function suggested to be of a tumor suppressing nature (Crist et al., 2011). In addition, the *MEIS1* was also reported to regulate cell proliferation in non-small-cell lung cancer as well as in mammalian heart cardiomyocytes. The overexpression of *Meis1* in mouse cardiomyocytes resulted in a decrease in cardiomyocytes proliferation, while *Meis1* loss led to the proliferation period of these mouse cardiomyocytes (Mahmoud et al., 2013). Therefore, *MEIS1* expression can be suggested to play a role in a variety of cell types such as in the haematopoietic stem cells to more differentiated cardiomyocytes.

The role of *MEIS1* in different cell types highlights its importance in regulating cellular development. However, the regulation of *MEIS1* across different cell types and conditions are still at present largely unknown. Therefore, another area for investigation is to examine for the presence of the *MEIS1* FIRE in different types of cancers and cell types as well as to observe if this *MEIS1* FIRE regulates *MEIS1* expression in a similar manner to that in myeloid leukaemia in other cells types. These can be achieved through Hi-C sequencing and analysis of the chromatin interactions that are present in various types of cancer cell and normal cells genomes. Furthermore, RNA-seq and RT-qPCR can also be performed to evaluate corresponding *MEIS1* expression in these cell types. After which,

analysed Hi-C loops and FIRE/TAD regions can be aligned to individual sample's *MEIS1* expression profile. A possible candidate to begin with is the AML M3 cell line HL-60 as no available Hi-C data for this subtype of AML is available at the moment and since HL-60 does not express *MEIS1*, it will be interesting to investigate if the *MEIS1* FIRE is lost in this cell line. Altogether, these can provide us with more details on whether *MEIS1* FIRE occupancy and *MEIS1* expression are similarly associated in different cell types as the haematopoietic stem cells

7.7 Questions on the influence of acquired super-enhancers that hijack pre-existing chromatin loops to oncogenes promoters

In addition, further studies can be performed to conclude the influence of the acquired *SNX10* SE on *HOXA9* expression. In my analyses of the super-enhancers' profiles between the clinical AML samples and normal CD34+ samples, *SNX10* SE was among the three acquired super-enhancers that forms a chromatin loop to the *HOXA9* promoter. Moreover, I also found that AML patients who possessed this *SNX10* SE are strongly associated with high *HOXA9* expression levels that were above the normal cells. Therefore, further experiment that inhibits the chromatin interaction between the *SNX10* SE and the *HOXA9* promoter will allow us to gain a better idea of how much influence does the *SNX10* SE have on promoting *HOXA9* overexpression in AML. Since chromatin loops and HOXA9 proteins are observed to be present between the *SNX10* SE and *HOXA9* promoter, CRISPR-dCas9 experiments targeting the *SNX10* SE can be designed to inhibit the interaction between the *SNX10* SE and the *HOXA9* promoter. Through the CRISPR-Cas9 or CRISPR-dCas9-KRAB (Kruppel Associated Box) mediated excision or inhibition of *SNX10* SE, we will be able to examine for potential downregulation of *HOXA9* expression in myeloid leukaemia (**Figure 36**). By understanding the impact of *SNX10* SE on *HOXA9* expression, we can then consider if it is a suitable candidate for development as a target for gene therapy such as with techniques like CRISPR or oligonucleotide-based decoys in myeloid leukaemia.

Further 4C experiments can be performed with the acquired *SNX10* SE as the bait to examine if other genes beside *HOXA9* form loops with it and if the deletion of this acquired *SNX10* super-enhancer will have an effect on unintended genes transcription alteration. This will be important as part of the validation before the *SNX10* SE can be used as a specific target for further gene therapy purposes in myeloid leukaemia.

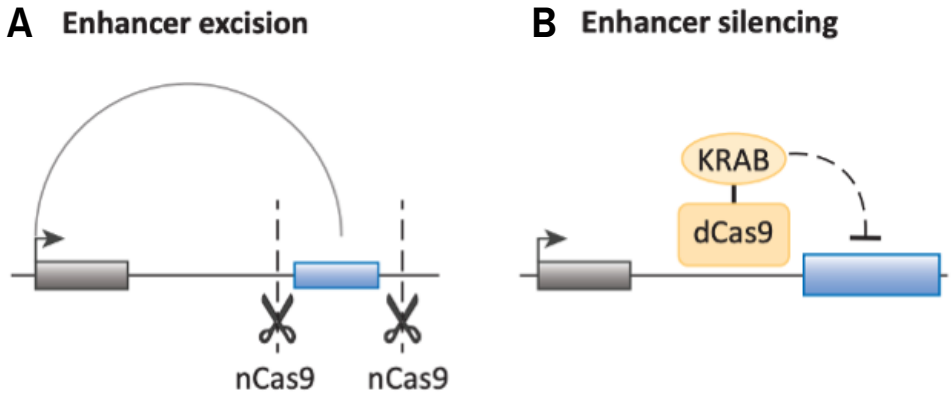

Figure 36. Schematic of CRISPR-Cas9 & dCas9-KRAB silencing of enhancers and super-enhancers in the genome. A. Illustrated image of CRISPR cut sites to the upstream and downstream of an enhancer region that is interacting with a gene via chromatin interaction. B. Illustrated image of dCas9-KRAB binding to the enhancer region that is interacting with a gene via chromatin interaction to inhibit the super-enhancer activity. The red rectangle represents the gene location and the blue rectangle represents the enhancer. The schematic is adapted from my publication (See et al., 2019).

7.8 Translational directions for the development of targeted epigenetic therapies for myeloid leukaemias

Finally, as *MEIS1* and *HOXA9* are revealed to involve chromatin interactions to super-enhancer regions that contains high amounts of transcription factors, the bromodomain inhibitor JQ1 and cyclin-dependent kinase 7 (CDK7) inhibitor THZ1 should be investigated for their effectiveness in downregulating *MEIS1* and *HOXA9* expression. JQ1 is a drug that inhibits the binding of the transcriptional co-activator BRD4 at super-enhancers, thereby downplaying the effect that super-enhancers have on oncogenes transcription (Lovén et al., 2013). The CDK7 enzyme is an essential for the process of transcription, as it functions by phosphorylating the RNA polymerase II domain (Sampathi et al., 2019). The CDK7 inhibitor THZ1 prevents CDK7 from functioning by binding to the serine residue outside the kinase domain (Cao and Shilatifard, 2014, Arner et al., 2015).

In addition, THZ1 was also reported to have an effect on inhibiting super-enhancers associated genes transcription in cancers. For instance, Chen et al. (2018) demonstrated that THZ1 induction reduced cell proliferation and invasion rates in osteosarcoma through the mechanism of inhibiting super-enhancer associated genes. Also Zhou et al. (2019) showed that THZ1 resulted in the downregulation of several super-enhancer associated genes in multiple myeloma and that THZ1 treatment resulted in the loss of chromatin interaction between a oncogene *HJURP* super-enhancer and its promoter.

Since super-enhancers are known to attract high level of transcriptional proteins such as BRD4 and CDK7 (Brown et al., 2018, Arner et al., 2015, Core et al., 2014), the application of JQ1 and THZ1 on AML subtypes with high *MEIS1* and *HOXA9* expression such as the M5 subtype, THP-1 cell line can allow us to gain a better understanding of their suitability as a treatment drug in AMLs with high *MEIS1* and *HOXA9* expression levels. For example, AML cells with high *MEIS1* and *HOXA9* expression levels can be subjected to treatment with the drug JQ1 before being analysed for potential chromatin changes between the super-enhancers and the *MEIS1* and *HOXA9* promoters with 4C or Hi-C techniques. In addition, qPCR of *HOXA9* and *MEIS1* can also be performed on the RNA that is derived from JQ1 treated AML cells to evaluate if JQ1 is effective in reducing *HOXA9* and *MEIS1* expression levels in AML. The matching of the Hi-C or 4C results with these *HOXA9* and *MEIS1* expressions from cells that are treated with JQ1 will then allow us to draw a clearer conclusion to the efficacy of JQ1 in disrupting super-enhancer to *HOXA9* and *MEIS1* promoter interactions and their subsequent effects on these two oncogenes expression.

In addition, the validation of these identified chromatin interactions in vivo should also be investigated. For example, the implantation of AML cells with CTCF binding site excision and its control cells can be done in mouse models to assess their growth kinetic in vivo and also to examine if these altered chromatin interactions at the *MEIS1* FIRE is consistent over time. The survival period of these AML cells implanted mouse models can also be observed to understand if the disruption of the MEIS1 FIRE has an effect on survival.

All in all, this pool of additional information that is obtained from my work in this project can enable us to advance towards the development of therapies such as DNA decoys and CRISPR based treatment that are targeted towards aberrant chromatin interactions in cancers such as myeloid leukaemia.

8 Conclusions

Altogether, my work in this study addresses the three aims and hypotheses that are mentioned in the following manner.

8.1 Conclusion I

In addressing my first aim and hypothesis that TADs/FIREs lead to increased transcription of *MEIS1* myeloid leukaemia, my results have demonstrated that the loss of the *MEIS1* FIRE and disruption of its associated chromatin interactions via CTCF binding site deletion led to the downregulation of *MEIS1* and its downstream *MYC* oncogene in the K562 myeloid leukamia cells. These findings support my first hypothesis and suggest that sub-TADs like FIREs are important in ensuring the organization of chromatin loops to oncogenes such as *MEIS1* and its consequent expression. However, the role of TADs and FIREs may not apply across all oncogenes in diseases such as myeloid leukaemia. The *HOXA9* locus was observed to not exhibit any alterations in TADs and chromatin loops between the clinical normal CD34+ bone marrow sample in comparison to the clinical AML CD34+ samples as well as between the AML cell line THP-1 and HL-60 which possess high and low *HOXA9* expression respectively. These suggest that altered chromatin interactions and disruption of 3D genome organisation could be specific to certain gene locations and that no ‘one rule fits all’ approach can be used to explain for the mechanisms that regulate oncogenes like *HOXA9* and *MEIS1* during disease pathogenesis.

8.2 Conclusion II

My second aim and hypothesis set out to characterize for acquired super-enhancer regions that forms chromatin loops to *HOXA9* and *MEIS1* oncogenes promoters. Based on my collective findings, I revealed that super-enhancers are

acquired in AML as compared to normal CD34+ cells and that these super-enhancers hijacked existing chromatin interaction loops that are present in the normal CD34+ cells to promote oncogenes overexpression. I presented examples of such acquired super-enhancers such as the *SNX10* SE that hijacks the unaltered chromatin loops to the *HOXA9* promoter to overexpress *HOXA9* in AML patients. This acquisition of super-enhancers in cis to oncogenes promoters may therefore be a potential way that oncogenes can be overexpressed on top of altered chromatin interactions occurrence. Furthermore, the hijacking of preexisting loops by these acquired super-enhancers could occur at both regions with altered and unaltered chromatin loops and TADs in cancers such as myeloid leukaemia.

8.3 Conclusion III

On investigating my final aim and hypothesis, my ChIP-qPCR and ChIP-seq analyses of HOXA9 and MEIS1 binding patterns at the *HOXA9*, *MEIS1*, *MYC* and *SNX10* SE regions demonstrated that HOXA9 and MEIS1 proteins are enriched at their own and each other's promoters, thus suggesting that these genes can co-regulate mutual expression in myeloid leukemia via a feedback loop. These results address my hypothesis that *HOXA9* and *MEIS1* operate in a coordinated fashion thereby leading to their high levels of co-expression in myeloid leukemia. *HOXA9* enrichment is also seen to occur at the acquired *SNX10* SE that is highly associated with *HOXA9* overexpression in AML patients.

8.4 Overview

Taken together, my results have addressed the three aims that I set out to investigate on. In addition, my findings across the three conclusions provide novel details as to how coordinated *HOXA9* and *MEIS1* overexpression occurs via the influence of the *MEIS1* FIRE, dysregulated chromatin interactions to *MEIS1* and *HOXA9* and the presence of acquired super-enhancers which are enriched in HOXA9 proteins in myeloid leukaemia pathogenesis.

8.5 Significance

Taken together, my findings uncovered potential mechanisms of *MEIS1* and *HOXA9* regulations that could contribute to aberrant oncogenes expression levels in myeloid leukaemia (**Figure 37**). I have also revealed the presence of newly acquired super-enhancers that form chromatin loops towards the *HOXA9* oncogene promoter which likely promotes *HOXA9* overexpression in AMLs. Based on these results, novel therapies such as dCas9-KRAB, CRISPR based gene therapies and drugs like the super-enhancers inhibitor JQ1 can be developed towards targeting super-enhancer-promoter chromatin interactions as well as the CTCF boundaries of FIREs that are associated with aberrant oncogenes in cancers like myeloid leukaemia. Finally, my work has uncovered new biology in the understanding of FIREs regulation via CTCF and chromatin interactions and also laid down the platform for such information to be translated into future treatment in the clinical.

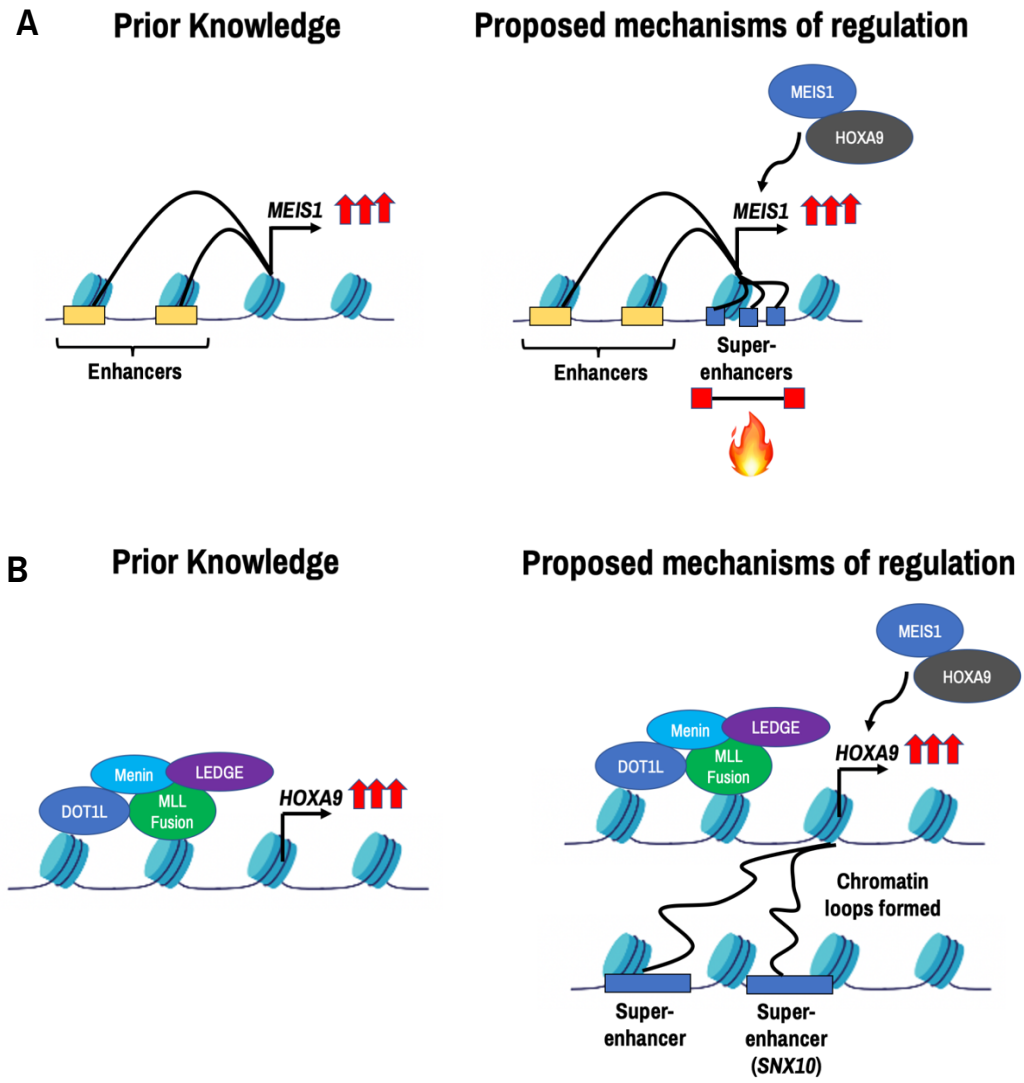

Figure 37. Schematic of prior and additional proposed mechanisms of *HOXA9 and MEIS1* regulation in myeloid leukaemia. A. Schematic of additional proposed *MEIS1* regulation via FIRE mediated chromatin interactions and the binding of HOXA9 and MEIS1 transcription factors to the *MEIS1* promoter in myeloid leukaemia. The FIRE represents the location of the *MEIS1* FIRE, with the red boxes showing the CTCF anchors at the region. B. Schematic of additional proposed mechanisms of aberrant *HOXA9* regulation via chromatin loops formation between acquired super-enhancers and the *HOXA9* promoter as well as the binding of MEIS1 and HOXA9 transcription factors to the *HOXA9* promoter in myeloid leukaemia. The red arrows indicate the upregulation of gene expression. The rectangular boxes represents the locations of enhancers and super-enhancer at the gene locus. Proteins and epigenetic modulators are represented in oval coloured shapes. (Part of illustration in this figure was generated with the website Biorender.com)

9 References

Acute Myelogenous Leukemia and Acute Promyelocytic Leukemia - in Union for International Cancer Control. 2014. *Review of Cancer Medicines on the WHO List of Essential Medicines;* [Online]. Available: http://www.who.int/selection_medicines/committees/expert/20/applications/AML_APL.pdf [Accessed May 6 2020].

Adam, R. C., Yang, H., Rockowitz, S., Larsen, S. B., Nikolova, M., Oristian, D. S., Polak, L., Kadaja, M., Asare, A. & Zheng, D. 2015. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. *Nature,* 521**,** 366-370.

Aichberger, K. J., Mayerhofer, M., Krauth, M.-T., Vales, A., Kondo, R., Derdak, S., Pickl, W. F., Selzer, E., Deininger, M. & Druker, B. J. 2005. Low-level expression of proapoptotic Bcl-2–interacting mediator in leukemic cells in patients with chronic myeloid leukemia: role of BCR/ABL, characterization of underlying signaling pathways, and reexpression by novel pharmacologic compounds. *Cancer research,* 65**,** 9436-9444.

Albajar, M., Gómez-Casares, M. T., Llorca, J., Mauleon, I., Vaqué, J. P., Acosta, J. C., Bermúdez, A., Donato, N., Delgado, M. D. & León, J. 2011. MYC in chronic myeloid leukemia: induction of aberrant DNA synthesis and association with poor response to imatinib. *Molecular Cancer Research,* 9**,** 564-576.

Ali, T., Renkawitz, R. & Bartkuhn, M. 2016. Insulators and domains of gene expression. *Current opinion in genetics & development,* 37**,** 17-26.

American Cancer Society 2020. Stem Cell Transplant for Acute Myeloid Leukemia (AML).

Anders, S., Pyl, P. T. & Huber, W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-169.

Argiropoulos, B., Yung, E. & Humphries, R. K. 2007. Unraveling the crucial roles of Meis1 in leukemogenesis and normal hematopoiesis. *Genes & Development,* 21**,** 2845-2849.

Ariki, R., Morikawa, S., Mabuchi, Y., Suzuki, S., Nakatake, M., Yoshioka, K., Hidano, S., Nakauchi, H., Matsuzaki, Y. & Nakamura, T. 2014. Homeodomain transcription factor Meis1 is a critical regulator of adult bone marrow hematopoiesis. *PloS one,* 9.

Arner, E., Daub, C. O., Vitting-Seerup, K., Andersson, R., Lilje, B., Drabløs, F., Lennartsson, A., Rönnerblad, M., Hrydziuszko, O. & Vitezic, M. 2015. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science,* 347**,** 1010-1014.

Babu, D. & Fullwood, M. J. 2015. 3D genome organization in health and disease: emerging opportunities in cancer translational medicine. *Nucleus (Austin, Tex.),* 6**,** 382-93.

Barbieri, E., Trizzino, M., Welsh, S. A., Owens, T. A., Calabretta, B., Carroll, M., Sarma, K. & Gardini, A. 2018. Targeted enhancer activation by a subunit of the Integrator complex. *Molecular cell,* 71**,** 103-116. e7.

Barutcu, A. R., Lajoie, B. R., Mccord, R. P., Tye, C. E., Hong, D., Messier, T. L., Browne, G., Van Wijnen, A. J., Lian, J. B. & Stein, J. L. 2015. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome biology,* 16**,** 214.

Bessa, J., Tavares, M. J., Santos, J., Kikuta, H., Laplante, M., Becker, T. S., Gómez-Skarmeta, J. L. & Casares, F. 2008. meis1 regulates cyclin D1 and c-myc expression, and controls the proliferation of the multipotent cells in the early developing zebrafish eye. *Development,* 135**,** 799-803.

Bickmore, W. A. & Van Steensel, B. 2013. Genome architecture: domain organization of interphase chromosomes. *Cell,* 152**,** 1270-1284.

Blinka, S., Reimer Jr, M. H., Pulakanti, K. & Rao, S. 2016. Super-enhancers at the Nanog locus differentially regulate neighboring pluripotency-associated genes. *Cell reports,* 17**,** 19-28.

Bodoor, K., Haddad, Y., Alkhateeb, A., Al-Abbadi, A., Dowairi, M., Magableh, A., Bsoul, N., Ghabkari, A. & Bodoor, K. 2014. DNA hypermethylation of cell cycle (p15 and p16) and apoptotic (p14, p53, DAPK and TMS1) genes in peripheral blood of leukemia patients. *Asian Pac J Cancer Prev,* 15**,** 75-84.

Bonnet, D. & Dick, J. E. 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine,* 3**,** 730-737.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L. & Jenner, R. G. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *cell,* 122**,** 947-956.

Brondfield, S., Umesh, S., Corella, A., Zuber, J., Rappaport, A. R., Gaillard, C., Lowe, S. W., Goga, A. & Kogan, S. C. 2015. Direct and indirect targeting of MYC to treat acute myeloid leukemia. *Cancer chemotherapy and pharmacology,* 76**,** 35-46.

Brown, J. D., Feldman, Z. B., Doherty, S. P., Reyes, J. M., Rahl, P. B., Lin, C. Y., Sheng, Q., Duan, Q., Federation, A. J. & Kung, A. L. 2018. BET bromodomain proteins regulate enhancer function during adipogenesis. *Proceedings of the National Academy of Sciences,* 115**,** 2144-2149.

Burr, M. L., Sparbier, C. E., Chan, K. L., Chan, Y.-C., Kersbergen, A., Lam, E. Y., Azidis-Yates, E., Vassiliadis, D., Bell, C. C. & Gilan, O. 2019. An evolutionarily conserved function of polycomb silences the MHC class I antigen presentation pathway and enables immune evasion in cancer. *Cancer cell,* 36**,** 385-401. e8.

Cao, F., Fang, Y., Tan, H. K., Goh, Y., Choy, J. Y. H., Koh, B. T. H., Hao Tan, J., Bertin, N., Ramadass, A., Hunter, E., Green, J., Salter, M., Akoulitchev, A., Wang, W., Chng, W. J., Tenen, D. G. & Fullwood, M. J. 2017a. Super-Enhancers and Broad H3K4me3 Domains Form Complex Gene Regulatory Circuits Involving Chromatin Interactions. *Scientific reports,* 7**,** 2186-2186.

Cao, F., Fang, Y., Tan, H. K., Goh, Y., Choy, J. Y. H., Koh, B. T. H., Hao Tan, J., Bertin, N., Ramadass, A., Hunter, E., Green, J., Salter, M., Akoulitchev, A., Wang, W., Chng, W. J., Tenen, D. G. & Fullwood, M. J. 2017b. Super-Enhancers and Broad H3K4me3 Domains Form Complex Gene Regulatory Circuits Involving Chromatin Interactions. *Sci Rep,* 7**,** 2186.

Cao, K. & Shilatifard, A. 2014. Inhibit globally, act locally: CDK7 inhibitors in cancer therapy. *Cancer cell,* 26**,** 158-159.

Casas, S., Nagy, B., Elonen, E., Aventín, A., Larramendy, M. L., Sierra, J., Ruutu, T. & Knuutila, S. 2003. Aberrant expression of HOXA9, DEK, CBL and

CSF1R in acute myeloid leukemia. *Leukemia & lymphoma,* 44**,** 1935-1941.

Castaigne, S., Pautas, C., Terré, C., Raffoux, E., Bordessoule, D., Bastie, J.-N., Legrand, O., Thomas, X., Turlure, P. & Reman, O. 2012. Effect of gemtuzumab ozogamicin on survival of adult patients with de-novo acute myeloid leukaemia (ALFA-0701): a randomised, open-label, phase 3 study. *The Lancet,* 379**,** 1508-1516.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L. & Larsson, E. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. AACR.

Chávez-González, A., Avilés-Vázquez, S., Moreno-Lorenzana, D. & Mayani, H. 2013. Hematopoietic stem cells in chronic myeloid leukemia. *Stem Cell Biology in Normal Life and Disease. Intech***,** 137-64.

Chen, D., Zhao, Z., Huang, Z., Chen, D.-C., Zhu, X.-X., Wang, Y.-Z., Yan, Y.-W., Tang, S., Madhavan, S. & Ni, W. 2018. Super enhancer inhibitors suppress MYC driven transcriptional amplification and tumor progression in osteosarcoma. *Bone research,* 6**,** 1-7.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W. & Jiang, J. 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell,* 133**,** 1106-1117.

Christman, J. K. 2002. 5-Azacytidine and 5-aza-2′-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene,* 21**,** 5483-5495.

Ciurea, S. O., Zhang, M.-J., Bacigalupo, A. A., Bashey, A., Appelbaum, F. R., Aljitawi, O. S., Armand, P., Antin, J. H., Chen, J. & Devine, S. M. 2015. Haploidentical transplant with post-transplant cyclophosphamide versus matched unrelated donor transplant for acute myeloid leukemia. *Blood***,** blood-2015-04-639831.

Cogoi, S., Zorzet, S., Rapozzi, V., Geci, I., Pedersen, E. B. & Xodo, L. E. 2013. MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic acids research,* 41**,** 4049-4064.

Collins, C., Wang, J., Miao, H., Bronstein, J., Nawer, H., Xu, T., Figueroa, M., Muntean, A. G. & Hess, J. L. 2014. C/EBPα is an essential collaborator in Hoxa9/Meis1-mediated leukemogenesis. *Proceedings of the National Academy of Sciences,* 111**,** 9899-9904.

Collins, C. T. & Hess, J. L. 2016a. Deregulation of the HOXA9/MEIS1 axis in acute leukemia. *Current opinion in hematology,* 23**,** 354.

Collins, C. T. & Hess, J. L. 2016b. Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets. *Oncogene,* 35**,** 1090-1098.

Core, L. J., Martins, A. L., Danko, C. G., Waters, C. T., Siepel, A. & Lis, J. T. 2014. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nature genetics,* 46**,** 1311.

Crist, R. C., Roth, J. J., Waldman, S. A. & Buchberg, A. M. 2011. A conserved tissue-specific homeodomain-less isoform of MEIS1 is downregulated in colorectal cancer. *PLoS One,* 6**,** e23665.

D'ippolito, A. M., Mcdowell, I. C., Barrera, A., Hong, L. K., Leichter, S. M., Bartelt, L. C., Vockley, C. M., Majoros, W. H., Safi, A. & Song, L. 2018. Pre-established chromatin interactions mediate the genomic response to glucocorticoids. *Cell systems,* 7**,** 146-160. e7.

Daramola, A. S., Perez, M., Sias-Garcia, O. S., Terrell, M. S., Wei, H. Y., Cherayil, N., Stevens, A. M., Lin, C. Y. & Yi, J. S. 2019. Rara Is a Druggable Super-Enhancer Regulated Dependency in Pediatric AML. American Society of Hematology Washington, DC.

Dawson, M. A., Prinjha, R. K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W.-I., Robson, S. C., Chung, C.-W., Hopf, C. & Savitski, M. M. 2011. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature,* 478**,** 529-533.

De Lima, M. C., Da Silva, D. B., Freund, A. P. F., Dacoregio, J. S., Costa, T. E. J. B., Costa, I., Faraco, D. & Silva, M. L. 2016. Acute Myeloid Leukemia: Analysis of epidemiological profile and survival rate. *Jornal de Pediatria,* 92**,** 283-289.

Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. 2002. Capturing chromosome conformation. *science,* 295**,** 1306-1311.

Delgado, M. D. & León, J. 2010. Myc roles in hematopoiesis and leukemia. *Genes & cancer,* 1**,** 605-616.

Deschler, B., De Witte, T., Mertelsmann, R. & Lubbert, M. 2006. Treatment decision-making for older patients with high-risk myelodysplastic syndrome or acute myeloid leukemia: problems and approaches. *Haematologica,* 91**,** 1513-1522.

Di Croce, L., Raker, V. A., Corsaro, M., Fazi, F., Fanelli, M., Faretta, M., Fuks, F., Coco, F. L., Kouzarides, T. & Nervi, C. 2002. Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science,* 295**,** 1079-1082.

Dintilhac, A., Bihan, R., Guerrier, D., Deschamps, S. & Pellerin, I. 2004. A conserved non-homeodomain Hoxa9 isoform interacting with CBP is co-expressed with the 'typical'Hoxa9 protein during embryogenesis. *Gene expression patterns,* 4**,** 215-222.

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., Ye, Z., Kim, A., Rajagopal, N. & Xie, W. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature,* 518**,** 331-336.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.

Döhner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Büchner, T., Dombret, H., Ebert, B. L., Fenaux, P. & Larson, R. A. 2016. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood***,** blood-2016-08-733196.

Dombret, H. & Gardin, C. 2016. An update of current treatments for adult acute myeloid leukemia. *Blood,* 127**,** 53-61.

Dostie, J. & Dekker, J. 2007. Mapping networks of physical interactions between genomic elements using 5C technology. *Nature protocols,* 2**,** 988-1002.

Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., Rubio, E. D., Krumm, A., Lamb, J. & Nusbaum, C. 2006. Chromosome Conformation Capture Carbon Copy (5C): a massively

parallel solution for mapping interactions between genomic elements. *Genome research,* 16**,** 1299-1309.

Druhan, L., Fasan, O. & Copelan, O. R. 2015. Acute Heart Failure in a Patient with Acute Myeloid Leukemia following Daunorubicin Treatment: a Case Report. *Journal of Leukemia***,** 1-3.

Dunwell, T., Hesson, L., Rauch, T. A., Wang, L., Clark, R. E., Dallol, A., Gentle, D., Catchpoole, D., Maher, E. R. & Pfeifer, G. P. 2010. A genome-wide screen identifies frequently methylated genes in haematological and epithelial cancers. *Molecular cancer,* 9**,** 44.

Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S. & Aiden, E. L. 2016. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell systems,* 3**,** 99-101.

Eisfeld, A.-K., Kohlschmidt, J., Schwind, S., Nicolet, D., Blachly, J. S., Orwick, S., Shah, C., Bainazar, M., Kroll, K. W. & Walker, C. J. 2017. Mutations in the CCND1 and CCND2 genes are frequent events in adult patients with t (8; 21)(q22; q22) acute myeloid leukemia. *Leukemia,* 31**,** 1278-1285.

Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R. & Coyne, M. 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature,* 473**,** 43.

Eser, U., Chandler-Brown, D., Ay, F., Straight, A. F., Duan, Z., Noble, W. S. & Skotheim, J. M. 2017. Form and function of topologically associating genomic domains in budding yeast. *Proceedings of the National Academy of Sciences,* 114**,** E3061-E3070.

Estey, E. & Döhner, H. 2006. Seminar Acute myeloid leukaemia. *Lancet,* 368**,** 1894-1907.

Faber, J., Krivtsov, A. V., Stubbs, M. C., Wright, R., Davis, T. N., Van Den Heuvel-Eibrink, M., Zwaan, C. M., Kung, A. L. & Armstrong, S. A. 2009. HOXA9 is required for survival in human MLL-rearranged acute leukemias. *Blood, The Journal of the American Society of Hematology,* 113**,** 2375-2385.

Falini, B., Nicoletti, I., Martelli, M. F. & Mecucci, C. 2007. Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features. *Blood,* 109**,** 874-885.

Fathi, A. T. & Karp, J. E. 2009. New agents in acute myeloid leukemia: beyond cytarabine and anthracyclines. *Current oncology reports,* 11**,** 346-352.

Fernandes, J. C., Alves, A. P. N. R., Coelho-Silva, J. L., Scopim-Ribeiro, R., Fenerich, B. A., Simões, B. P., Rego, E. M., Machado-Neto, J. A. & Traina, F. 2018. Increased levels of cyclin D1 negatively impacts on acute lymphoblastic leukemia overall survival. *Applied Cancer Research,* 38**,** 7.

Ferrara, F. & Schiffer, C. A. 2013. Acute myeloid leukaemia in adults. *The Lancet,* 381**,** 484-495.

Ferreira, H. J., Heyn, H., Vizoso, M., Moutinho, C., Vidal, E., Gomez, A., Martínez-Cardús, A., Simó-Riudalbas, L., Moran, S. & Jost, E. 2016. DNMT3A mutations mediate the epigenetic reactivation of the leukemogenic factor MEIS1 in acute myeloid leukemia. *Oncogene,* 35**,** 3079-3082.

Figueroa, M. E., Abdel-Wahab, O., Lu, C., Ward, P. S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A. & Fernandez, H. F. 2010. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer cell,* 18**,** 553-567.

Fiore, C., Mckeown, M. R., Lee, E., Eaton, M. L. & Fritz, C. C. 2016. SY1425 (tamibarotene) Induces Profound Transcriptional Changes in AML Tumors with High Retinoic Acid Receptor Alpha. American Society of Hematology Washington, DC.

Flavahan, W. A., Drier, Y., Johnstone, S. E., Hemming, M. L., Tarjan, D. R., Hegazi, E., Shareef, S. J., Javed, N. M., Raut, C. P. & Eschle, B. K. 2019. Altered chromosomal topology drives oncogenic programs in SDH-deficient GISTs. *Nature,* 575**,** 229-233.

Flavahan, W. A., Drier, Y., Liau, B. B., Gillespie, S. M., Venteicher, A. S., Stemmer-Rachamimov, A. O., Suvà, M. L. & Bernstein, B. E. 2016. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature,* 529**,** 110-114.

Fraser, J., Williamson, I., Bickmore, W. A. & Dostie, J. 2015. An overview of genome organization and how we got there: from FISH to Hi-C. *Microbiology and Molecular Biology Reviews,* 79**,** 347-372.

Frech, M., Teichler, S., Weber, C., Bouchard, C., Sorg, K., Bullinger, L., Bauer, U.-M. & Neubauer, A. 2016. Identification of the Oncogene SKI As a New Target Gene of the Myeloid Transcription Factor c-MYB in AML. American Society of Hematology Washington, DC.

Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., Velkov, S., Ho, A. & Mei, P. H. 2009. An oestrogen-receptor-α-bound human chromatin interactome. *Nature,* 462**,** 58-64.

Gailhouste, L., Liew, L. C., Hatada, I., Nakagama, H. & Ochiya, T. 2018. Epigenetic reprogramming using 5-azacytidine promotes an anti-cancer response in pancreatic adenocarcinoma cells. *Cell death & disease,* 9**,** 1-12.

Galmarini, C. M., Thomas, X., Calvo, F., Rousselot, P., Rabilloud, M., El Jaffari, A., Cros, E. & Dumontet, C. 2002. In vivo mechanisms of resistance to cytarabine in acute myeloid leukaemia. *British journal of haematology,* 117**,** 860-868.

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R. & Larsson, E. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling,* 6**,** pl1-pl1.

Gao, L., Sun, J., Liu, F., Zhang, H. & Ma, Y. 2016. Higher expression levels of the HOXA9 gene, closely associated with MLL-PTD and EZH2 mutations, predict inferior outcome in acute myeloid leukemia. *OncoTargets and therapy,* 9**,** 711.

Garcia-Cuellar, M. P., Steger, J., Fuller, E., Hetzner, K. & Slany, R. K. 2015. Pbx3 and Meis1 cooperate through multiple mechanisms to support Hox-induced murine leukemia. *Haematologica,* 100**,** 905-13.

Gerber, J. M., Gucwa, J. L., Esopi, D., Gurel, M., Haffner, M. C., Vala, M., Nelson, W. G., Jones, R. J. & Yegnasubramanian, S. 2013. Genome-wide comparison of the transcriptomes of highly enriched normal and chronic

myeloid leukemia stem and progenitor cell populations. *Oncotarget,* 4**,** 715.

Giralt, S., Kantarjian, H. & Talpaz, M. The natural history of chronic myelogenous leukemia in the interferon era. Seminars in hematology, 1995. 152-158.

Gorkin, D. U., Leung, D. & Ren, B. 2014. The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell,* 14**,** 771-775.

Greif, P., Eck, S., Konstandin, N., Benet-Pages, A., Ksienzyk, B., Dufour, A., Vetter, A., Popp, H., Lorenz-Depiereux, B. & Meitinger, T. 2011. Identification of recurring tumor-specific somatic mutations in acute myeloid leukemia by transcriptome sequencing. *Leukemia,* 25**,** 821-827.

Grisendi, S., Mecucci, C., Falini, B. & Pandolfi, P. P. 2006. Nucleophosmin and cancer. *Nature Reviews Cancer,* 6**,** 493-505.

Gross, S., Cairns, R. A., Minden, M. D., Driggers, E. M., Bittinger, M. A., Jang, H. G., Sasaki, M., Jin, S., Schenkein, D. P. & Su, S. M. 2010. Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. *Journal of Experimental Medicine,* 207**,** 339-344.

Guillamot, M., Cimmino, L. & Aifantis, I. 2016. The impact of DNA methylation in hematopoietic malignancies. *Trends in cancer,* 2**,** 70-83.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D. U., Jung, I., Wu, H., Zhai, Y. & Tang, Y. 2015. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell,* 162**,** 900-910.

Haferlach, T., Kohlmann, A., Schnittger, S., Dugas, M., Hiddemann, W., Kern, W. & Schoch, C. 2005. AML M3 and AML M3 variant each have a distinct gene expression signature but also share patterns different from other genetically defined AML subtypes. *Genes, Chromosomes and Cancer,* 43**,** 113-127.

Hakim, O., Sung, M.-H., Voss, T. C., Splinter, E., John, S., Sabo, P. J., Thurman, R. E., Stamatoyannopoulos, J. A., De Laat, W. & Hager, G. L. 2011. Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. *Genome research,* 21**,** 697-706.

Hansen, A. S., Cattoglio, C., Darzacq, X. & Tjian, R. 2018. Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus,* 9**,** 20-32.

Hanssen, L. L., Kassouf, M. T., Oudelaar, A. M., Biggs, D., Preece, C., Downes, D. J., Gosden, M., Sharpe, J. A., Sloane-Stanley, J. A. & Hughes, J. R. 2017a. Tissue-specific CTCF–cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nature cell biology,* 19**,** 952-961.

Hanssen, L. L. P., Kassouf, M. T., Oudelaar, A. M., Biggs, D., Preece, C., Downes, D. J., Gosden, M., Sharpe, J. A., Sloane-Stanley, J. A., Hughes, J. R., Davies, B. & Higgs, D. R. 2017b. Tissue-specific CTCF–cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nature Cell Biology*.

Hao, S. & Shao, Z. 2015. HOTAIR is upregulated in acute myeloid leukemia and that indicates a poor prognosis. *International journal of clinical and experimental pathology,* 8**,** 7223-8.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z. & Li, L. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science,* 333**,** 1303-1307.

Hershman, D. L., Mcbride, R. B., Eisenberger, A., Tsai, W. Y., Grann, V. R. & Jacobson, J. S. 2008. Doxorubicin, cardiac risk factors, and cardiac toxicity in elderly patients with diffuse B-cell non-Hodgkin's lymphoma. *Journal of Clinical Oncology,* 26**,** 3159-3165.

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-Andre, V., Sigova, A. A., Hoke, H. A. & Young, R. A. 2013a. Super-enhancers in the control of cell identity and disease. *Cell,* 155**,** 934-47.

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. A., Hoke, H. A. & Young, R. A. 2013b. Super-enhancers in the control of cell identity and disease. *Cell,* 155**,** 934-947.

Hochhaus, A., Kantarjian, H. M., Baccarani, M., Lipton, J. H., Apperley, J. F., Druker, B. J., Facon, T., Goldberg, S. L., Cervantes, F. & Niederwieser, D. 2007. Dasatinib induces notable hematologic and cytogenetic responses in chronic-phase chronic myeloid leukemia after failure of imatinib therapy. *Blood,* 109**,** 2303-2309.

Hu, Y.-L., Fong, S., Ferrell, C., Largman, C. & Shen, W.-F. 2009. HOXA9 modulates its oncogenic partner Meis1 to influence normal hematopoiesis. *Molecular and cellular biology,* 29**,** 5181-5192.

Huang, Y., Sitwala, K., Bronstein, J., Sanders, D., Dandekar, M., Collins, C., Robertson, G., Macdonald, J., Cezard, T. & Bilenky, M. 2012. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood, The Journal of the American Society of Hematology,* 119**,** 388-398.

Huret, J.-L. & Senon, S. 1997. *Atlas of genetics and cytogenetics in oncology and haematology*, CHU Poitiers.

Iijima-Yamashita, Y., Matsuo, H., Yamada, M., Deguchi, T., Kiyokawa, N., Shimada, A., Tawa, A., Takahashi, H., Tomizawa, D. & Taga, T. 2018. Multiplex fusion gene testing in pediatric acute myeloid leukemia. *Pediatrics International,* 60**,** 47-51.

Imamura, T., Morimoto, A., Takanashi, M., Hibi, S., Sugimoto, T., Ishii, E. & Imashuku, S. 2002. Frequent co-expression of HoxA9 and Meis1 genes in infant acute lymphoblastic leukaemia with MLL rearrangement. *British journal of haematology,* 119**,** 119-121.

Ito, S., D'alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C. & Zhang, Y. 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature,* 466**,** 1129-1133.

Jan, M., Snyder, T. M., Corces-Zimmerman, M. R., Weissman, I. L., Quake, S. R. & Majeti, R. 2011. Clonal evolution of pre-leukemic hematopoietic stem cells precedes human acute myeloid leukemia. American Society of Hematology.

Jelinek, J., Gharibyan, V., Estecio, M. R., Kondo, K., He, R., Chung, W., Lu, Y., Zhang, N., Liang, S. & Kantarjian, H. M. 2011. Aberrant DNA methylation is associated with disease progression, resistance to imatinib and shortened survival in chronic myelogenous leukemia. *PloS one,* 6.

Jia, Y., Chng, W.-J. & Zhou, J. 2019. Super-enhancers: critical roles and therapeutic targets in hematologic malignancies. *Journal of Hematology & Oncology,* 12**,** 77.

Jönsson, S., Olsson, B., Söderberg, J. & Wadenvik, H. 2012. Good adherence to imatinib therapy among patients with chronic myeloid leukemia—a single-center observational study. *Annals of hematology,* 91**,** 679-685.

Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., Van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B. & Levine, S. S. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature,* 467**,** 430-435.

Kantarjian, H., Ravandi, F., O'brien, S., Cortes, J., Faderl, S., Garcia-Manero, G., Jabbour, E., Wierda, W., Kadia, T. & Pierce, S. 2010. Intensive chemotherapy does not benefit most older patients (age 70 years or older) with acute myeloid leukemia. *Blood, The Journal of the American Society of Hematology,* 116**,** 4422-4429.

Kleer, C. G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S. A., Ghosh, D., Sewalt, R. G., Otte, A. P. & Hayes, D. F. 2003. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proceedings of the National Academy of Sciences,* 100**,** 11606-11611.

Kloetgen, A., Thandapani, P., Ntziachristos, P., Ghebrechristos, Y., Nomikou, S., Lazaris, C., Chen, X., Hu, H., Bakogianni, S. & Wang, J. 2020. Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nature Genetics,* 52**,** 388-400.

Koh, Y., Kim, D.-Y., Park, S.-H., Byun, H.-M., Kim, I., Yoon, S.-S., Kim, B. K., Park, E. & Park, S. 2010. Increased BCR promoter DNA methylation status strongly correlates with favorable response to imatinib in chronic myeloid leukemia patients. *Oncology letters,* 2**,** 181-187.

Kroon, E., Krosl, J., Thorsteinsdottir, U., Baban, S., Buchberg, A. M. & Sauvageau, G. 1998. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *The EMBO journal,* 17**,** 3714-3725.

Kumar, A. R., Sarver, A. L., Wu, B. & Kersey, J. H. 2010. Meis1 maintains stemness signature in MLL-AF9 leukemia. *Blood,* 115**,** 3642-3.

Kundu, M., Chen, A., Anderson, S., Kirby, M., Xu, L., Castilla, L. H., Bodine, D. & Liu, P. P. 2002. Role of Cbfb in hematopoiesis and perturbations resulting from expression of the leukemogenic fusion gene Cbfb-MYH11. *Blood, The Journal of the American Society of Hematology,* 100**,** 2449-2456.

Kuribara, R., Honda, H., Matsui, H., Shinjyo, T., Inukai, T., Sugita, K., Nakazawa, S., Hirai, H., Ozawa, K. & Inaba, T. 2004. Roles of Bim in apoptosis of normal and Bcr-Abl-expressing hematopoietic progenitors. *Molecular and cellular biology,* 24**,** 6172-6183.

Kuroda, J., Puthalakath, H., Cragg, M. S., Kelly, P. N., Bouillet, P., Huang, D. C., Kimura, S., Ottmann, O. G., Druker, B. J. & Villunger, A. 2006. Bim and Bad mediate imatinib-induced killing of Bcr/Abl+ leukemic cells, and resistance due to their loss is overcome by a BH3 mimetic. *Proceedings of the National Academy of Sciences,* 103**,** 14907-14912.

Kusio-Kobialka, M., Wolanin, K., Podszywalow-Bartnicka, P., Sikora, E., Skowronek, K., Mckenna, S. L., Ghizzoni, M., Dekker, F. J. & Piwocka, K. 2012. Increased acetylation of lysine 317/320 of p53 caused by BCR-ABL protects from cytoplasmic translocation of p53 and mitochondria-dependent apoptosis in response to DNA damage. *Apoptosis,* 17**,** 950-963.

Langmead, B. & Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods,* 9**,** 357-359.

Lawrence, H., Rozenfeld, S., Cruz, C., Matsukuma, K., Kwong, A., Kömüves, L., Buchberg, A. & Largman, C. 1999. Frequent co-expression of the HOXA9 and MEIS1 homeobox genes in human myeloid leukemias. *Leukemia,* 13**,** 1993-1999.

Lebert-Ghali, C.-É., Fournier, M., Kettyle, L., Thompson, A., Sauvageau, G. & Bijl, J. J. 2016. Hoxa cluster genes determine the proliferative activity of adult mouse hematopoietic stem and progenitor cells. *Blood,* 127**,** 87-90.

Lehnertz, B., Pabst, C., Su, L., Miller, M., Liu, F., Yi, L., Zhang, R., Krosl, J., Yung, E. & Kirschner, J. 2014. The methyltransferase G9a regulates HoxA9-dependent transcription in AML. *Genes & development,* 28**,** 317-327.

Ley, T. J., Ding, L., Walter, M. J., Mclellan, M. D., Lamprecht, T., Larson, D. E., Kandoth, C., Payton, J. E., Baty, J. & Welch, J. 2010. DNMT3A mutations in acute myeloid leukemia. *New England Journal of Medicine,* 363**,** 2424-2433.

Li, D.-P., Li, Z.-Y., Sang, W., Cheng, H., Pan, X.-Y. & Xu, K.-L. 2013a. HOXA9 Gene Expression in Acute Myeloid Leukemia. *Cell Biochemistry and Biophysics,* 67**,** 935-938.

Li, G., Gao, Y., Li, K., Lin, A. & Jiang, Z. 2020. Genomic analysis of biomarkers related to the prognosis of acute myeloid leukemia. *Oncology Letters*.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li, S., Mason, C. E. & Melnick, A. 2016. Genetic and epigenetic heterogeneity in acute myeloid leukemia. *Current opinion in genetics & development,* 36**,** 100-106.

Li, W., Huang, K., Guo, H. & Cui, G. 2014. Meis1 regulates proliferation of non-small-cell lung cancer cells. *Journal of thoracic disease,* 6**,** 850.

Li, Y., He, Y., Liang, Z., Wang, Y., Chen, F., Djekidel, M. N., Li, G., Zhang, X., Xiang, S. & Wang, Z. 2018. Alterations of specific chromatin conformation affect ATRA-induced leukemia cell differentiation. *Cell death & disease,* 9**,** 1-15.

Li, Z., Zhang, Z., Li, Y., Arnovitz, S., Chen, P., Huang, H., Jiang, X., Hong, G.-M., Kunjamma, R. B. & Ren, H. 2013b. PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood, The Journal of the American Society of Hematology,* 121**,** 1422-1431.

Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J. & Dorschner, M. O. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *science,* 326**,** 289-293.

Love, M. I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology,* 15**,** 550.

Lovén, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., Bradner, J. E., Lee, T. I. & Young, R. A. 2013. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell,* 153**,** 320-334.

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M. & Laxova, R. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell,* 161**,** 1012-1025.

Magnusson, M., Brun, A. C., Lawrence, H. J. & Karlsson, S. 2007. Hoxa9/hoxb3/hoxb4 compound null mice display severe hematopoietic defects. *Experimental hematology,* 35**,** 1421. e1-1421. e9.

Mahmoud, A. I., Kocabas, F., Muralidhar, S. A., Kimura, W., Koura, A. S., Thet, S., Porrello, E. R. & Sadek, H. A. 2013. Meis1 regulates postnatal cardiomyocyte cell cycle arrest. *Nature,* 497**,** 249-253.

Malhotra, V. & Perry, M. C. 2003. Classical chemotherapy: mechanisms, toxicities and the therapeutc window. *Cancer biology & therapy,* 2**,** 1-3.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., Dynlacht, B. D. & Reinberg, D. 2008. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Molecular cell,* 32**,** 503-518.

Martin, P., Mcgovern, A., Orozco, G., Duffus, K., Yarwood, A., Schoenfelder, S., Cooper, N. J., Barton, A., Wallace, C. & Fraser, P. 2015. Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nature communications,* 6**,** 10069.

Martindale, W. 1993. The extra pharmacopoeia 30th ed. London.

Mayer, R. J., Davis, R. B., Schiffer, C. A., Berg, D. T., Powell, B. L., Schulman, P., Omura, G. A., Moore, J. O., Mcintyre, O. R. & Frei, E. 1994. Intensive postremission chemotherapy in adults with acute myeloid leukemia. *New England Journal of Medicine,* 331**,** 896-903.

Mckeown, M. R., Corces, M. R., Eaton, M. L., Fiore, C., Lee, E., Lopez, J. T., Chen, M. W., Smith, D., Chan, S. M. & Koenig, J. L. 2017a. Superenhancer analysis defines novel epigenomic subtypes of non-APL AML, including an RARα dependency targetable by SY-1425, a potent and selective RARα agonist. *Cancer discovery,* 7**,** 1136-1153.

Mckeown, M. R., Corces, M. R., Eaton, M. L., Fiore, C., Lee, E., Lopez, J. T., Chen, M. W., Smith, D., Chan, S. M., Koenig, J. L., Austgen, K., Guenther, M. G., Orlando, D. A., Loven, J., Fritz, C. C. & Majeti, R. 2017b. Superenhancer Analysis Defines Novel Epigenomic Subtypes of Non-APL AML, Including an RARalpha Dependency Targetable by SY-1425, a Potent and Selective RARalpha Agonist. *Cancer Discov,* 7**,** 1136-1153.

Meddens, C. A., Harakalova, M., Van Den Dungen, N. A., Asl, H. F., Hijma, H. J., Cuppen, E. P., Björkegren, J. L., Asselbergs, F. W., Nieuwenhuis, E. E. & Mokry, M. 2016. Systematic analysis of chromatin interactions at disease associated loci links novel candidate genes to inflammatory bowel disease. *Genome biology,* 17**,** 247.

Meyer, C., Burmeister, T., Gröger, D., Tsaur, G., Fechina, L., Renneville, A., Sutton, R., Venn, N., Emerenciano, M. & Pombo-De-Oliveira, M. 2018. The MLL recombinome of acute leukemias in 2017. *Leukemia,* 32**,** 273-284.

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D. & Hess, J. L. 2002. MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular cell,* 10**,** 1107-1117.

Mohaghegh, N., Bray, D., Keenan, J., Penvose, A., Andrilenas, K. K., Ramlall, V. & Siggers, T. 2019. NextPBM: a platform to study cell-specific transcription factor binding and cooperativity. *Nucleic acids research,* 47**,** e31-e31.

Momparler, R. L. 2013. Optimization of cytarabine (ARA-C) therapy for acute myeloid leukemia. *Experimental hematology & oncology,* 2**,** 20.

Moore, G., Knight, G. & Blann, A. D. 2016. *Haematology*, Oxford University Press.
Morgado, E., Albouhair, S. P. & Lavau, C. 2007. Flt3 is dispensable to the Hoxa9/Meis1 leukemogenic cooperation. *Blood,* 109**,** 4020-4022.
Moskow, J. J., Bullrich, F., Huebner, K., Daar, I. O. & Buchberg, A. M. 1995. Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. *Molecular and Cellular Biology,* 15**,** 5434-5443.
Mujahed, H., Miliara, S., Neddermeyer, A. H., Bengtzen, S., Nilsson, C., Deneberg, S., Cordeddu, L., Ekwall, K., Lennartsson, A. & Lehmann, S. 2020a. AML Displays Increased CTCF Occupancy Associated to Aberrant Gene Expression and Transcription Factor Binding. *Blood*.
Mujahed, H., Miliara, S., Neddermeyer, A. H., Bengtzen, S., Nilsson, C., Deneberg, S., Cordeddu, L., Ekwall, K., Lennartsson, A. & Lehmann, S. 2020b. AML Displays Increased CTCF Occupancy Associated to Aberrant Gene Expression and Transcription Factor Binding. *Blood Journal***,** blood. 2019002326.
Mukherjee, K., Twyman, R. M. & Vilcinskas, A. 2015. Insects as models to study the epigenetic basis of disease. *Progress in biophysics and molecular biology,* 118**,** 69-78.
Mumbach, M. R., Rubin, A. J., Flynn, R. A., Dai, C., Khavari, P. A., Greenleaf, W. J. & Chang, H. Y. 2016. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nature methods,* 13**,** 919-922.
Nair, S. S. & Kumar, R. 2012. Chromatin remodeling in Cancer: A Gateway to regulate gene Transcription. *Molecular Oncology,* 6**,** 611-619.
Nakamura, T., Largaespada, D. A., Lee, M. P., Johnson, L. A., Ohyashiki, K., Toyama, K., Chen, S. J., Willman, C. L., Chen, I.-M. & Feinberg, A. P. 1996. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t (7; 11)(p15; p15) in human myeloid leukaemia. *Nature genetics,* 12**,** 154-158.
National Cancer Institute. *SEER Stat Fact Sheets: Acute Myeloid Leukemia (AML)* [Online]. Available: http://seer.cancer.gov/statfacts/html/amyl.html [Accessed May 6 2020].
Network, C. G. a. R. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl j Med,* 2013**,** 2059-2074.
Nora, E. P., Goloborodko, A., Valton, A.-L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A. & Bruneau, B. G. 2017. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell,* 169**,** 930-944. e22.
Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. 2018. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proceedings of the National Academy of Sciences,* 115**,** E6697-E6706.
Okuwaki, M., Matsumoto, K., Tsujimoto, M. & Nagata, K. 2001. Function of nucleophosmin/B23, a nucleolar acidic protein, as a histone chaperone. *FEBS letters,* 506**,** 272-276.
Ottmann, O., Saglio, G., Apperley, J. F., Arthur, C., Bullorsky, E., Charbonnier, A., Dipersio, J. F., Kantarjian, H., Khoury, H. J. & Kim, D.-W. 2018. Long-term efficacy and safety of dasatinib in patients with chronic myeloid leukemia in accelerated phase who are resistant to or intolerant of imatinib. *Blood cancer journal,* 8**,** 1-4.

Peng, Y., Xiong, D., Zhao, L., Ouyang, W., Wang, S., Sun, J., Zhang, Q., Guan, P., Xie, L. & Li, W. 2019. Chromatin interaction maps reveal genetic regulation for quantitative traits in maize. *Nature communications,* 10**,** 1-11.

Phillips-Cremins, J. E., Sauria, M. E., Sanyal, A., Gerasimova, T. I., Lajoie, B. R., Bell, J. S., Ong, C.-T., Hookway, T. A., Guo, C. & Sun, Y. 2013. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell,* 153**,** 1281-1295.

Pineault, N., Helgason, C. D., Lawrence, H. J. & Humphries, R. K. 2002. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Experimental hematology,* 30**,** 49-57.

Pott, S. & Lieb, J. D. 2015. What are super-enhancers? *Nature genetics,* 47**,** 8-12.

Prange, K. H., Mandoli, A., Kuznetsova, T., Wang, S.-Y., Sotoca, A. M., Marneth, A. E., Van Der Reijden, B. A., Stunnenberg, H. G. & Martens, J. H. 2017. MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. *Oncogene,* 36**,** 3346-3356.

Quintás-Cardama, A. & Cortes, J. 2009. Molecular biology of bcr-abl1–positive chronic myeloid leukemia. *Blood, The Journal of the American Society of Hematology,* 113**,** 1619-1630.

Raisner, R., Bainer, R., Haverty, P. M., Benedetti, K. L. & Gascoigne, K. E. 2020. Super-enhancer acquisition drives oncogene expression in triple negative breast cancer. *Plos one,* 15**,** e0235343.

Ramos-Mejía, V., Navarro-Montero, O., Ayllón, V., Bueno, C., Romero, T., Real, P. J. & Menendez, P. 2014. HOXA9 promotes hematopoietic commitment of human embryonic stem cells. *Blood, The Journal of the American Society of Hematology,* 124**,** 3065-3075.

Rao, S. S., Huang, S.-C., St Hilaire, B. G., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K.-R., Sanborn, A. L., Johnstone, S. E., Bascom, G. D. & Bochkov, I. D. 2017. Cohesin loss eliminates all loop domains. *Cell,* 171**,** 305-320. e24.

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D. & Lander, E. S. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell,* 159**,** 1665-1680.

Ren, R. 2005. Mechanisms of BCR–ABL in the pathogenesis of chronic myelogenous leukaemia. *Nature Reviews Cancer,* 5**,** 172-183.

Ribeiro, A. F. T., Pratcorona, M., Erpelinck-Verschueren, C., Rockova, V., Sanders, M., Abbas, S., Figueroa, M. E., Zeilemaker, A., Melnick, A. & Löwenberg, B. 2012. Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. *Blood, The Journal of the American Society of Hematology,* 119**,** 5824-5831.

Rio-Machin, A., Gómez-López, G., Muñoz, J., Garcia-Martinez, F., Maiques-Diaz, A., Alvarez, S., Salgado, R., Shrestha, M., Torres-Ruiz, R. & Haferlach, C. 2017. The molecular pathogenesis of the NUP98-HOXA9 fusion protein in acute myeloid leukemia. *Leukemia,* 31**,** 2000-2005.

Roche-Lestienne, C., Soenen-Cornu, V., Grardel-Duflos, N., Laı, J.-L., Philippe, N., Facon, T., Fenaux, P. & Preudhomme, C. 2002. Several types of

mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood,* 100**,** 1014-1018.

Rose, D., Haferlach, T., Schnittger, S., Perglerova, K., Kern, W. & Haferlach, C. 2017. Subtype-specific patterns of molecular mutations in acute myeloid leukemia. *Leukemia,* 31**,** 11-17.

Roumiantsev, S., Shah, N. P., Gorre, M. E., Nicoll, J., Brasher, B. B., Sawyers, C. L. & Van Etten, R. A. 2002. Clinical resistance to the kinase inhibitor STI-571 in chronic myeloid leukemia by mutation of Tyr-253 in the Abl kinase domain P-loop. *Proceedings of the National Academy of Sciences,* 99**,** 10700-10705.

Rousseau, M., Ferraiuolo, M. A., Crutchley, J. L., Wang, X. Q. D., Miura, H., Blanchette, M. & Dostie, J. 2014. Classifying leukemia types with chromatin conformation data. *Genome biology,* 15**,** R60.

Rowe, J. M., Li, X., Cassileth, P. A., Appelbaum, F. R., Schiffer, C. A., Wiernik, P. H., Litzow, M. R., Cripe, L. D., Lazarus, H. M. & Paietta, E. 2005. Very Poor Survival of Patients with AML Who Relapse after Achieving a First Complete Remission: The Eastern Cooperative Oncology Group Experience. American Society of Hematology.

Rowley, J. D. 1973. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature,* 243**,** 290-293.

Ryu, J., Kim, H., Yang, D., Lee, A. J. & Jung, I. 2019. A new class of constitutively active super-enhancers is associated with fast recovery of 3D chromatin loops. *BMC bioinformatics,* 20**,** 127.

Sabari, B. R., Dall'agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., Abraham, B. J., Hannett, N. M., Zamudio, A. V. & Manteiga, J. C. 2018. Coactivator condensation at super-enhancers links phase separation and gene control. *Science,* 361**,** eaar3958.

Sabath, D. 2013. Philadelphia Chromosome.

Sakuma, T., Nishikawa, A., Kume, S., Chayama, K. & Yamamoto, T. 2014. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Scientific reports,* 4**,** 1-6.

Salvatori, B., Iosue, I., Djodji Damas, N., Mangiavacchi, A., Chiaretti, S., Messina, M., Padula, F., Guarini, A., Bozzoni, I. & Fazi, F. 2011. Critical role of c-Myc in acute myeloid leukemia involving direct regulation of miR-26a and histone methyltransferase EZH2. *Genes & cancer,* 2**,** 585-592.

Sampathi, S., Acharya, P., Zhao, Y., Wang, J., Stengel, K. R., Liu, Q., Savona, M. R. & Hiebert, S. W. 2019. The CDK7 inhibitor THZ1 alters RNA polymerase dynamics at the 5′ and 3′ ends of genes. *Nucleic acids research,* 47**,** 3921-3936.

San José-Eneriz, E., Agirre, X., Jiménez-Velasco, A., Cordeu, L., Martín, V., Arqueros, V., Gárate, L., Fresquet, V., Cervantes, F. & Martínez-Climent, J. A. 2009. Epigenetic down-regulation of BIM expression is associated with reduced optimal responses to imatinib treatment in chronic myeloid leukaemia. *European journal of cancer,* 45**,** 1877-1889.

Sanborn, A. L., Rao, S. S., Huang, S.-C., Durand, N. C., Huntley, M. H., Jewett, A. I., Bochkov, I. D., Chinnappan, D., Cutkosky, A. & Li, J. 2015. Chromatin extrusion explains key features of loop and domain formation

in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences,* 112**,** E6456-E6465.

Schmidt, M., Rinke, J., Schäfer, V., Schnittger, S., Kohlmann, A., Obstfelder, E., Kunert, C., Ziermann, J., Winkelmann, N. & Eigendorff, E. 2014. Molecular-defined clonal evolution in patients with chronic myeloid leukemia independent of the BCR-ABL status. *Leukemia,* 28**,** 2292-2299.

Schmitt, A. D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C. L., Li, Y., Lin, S., Lin, Y. & Barr, C. L. 2016a. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell reports,* 17**,** 2042-2059.

Schmitt, A. D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C. L., Li, Y., Lin, S., Lin, Y., Barr, C. L. & Ren, B. 2016b. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep,* 17**,** 2042-2059.

Scott, M. T., Korfi, K., Saffrey, P., Hopcroft, L. E., Kinstrie, R., Pellicano, F., Guenther, C., Gallipoli, P., Cruz, M. & Dunn, K. 2016. Epigenetic reprogramming sensitizes CML stem cells to combined EZH2 and tyrosine kinase inhibition. *Cancer discovery,* 6**,** 1248-1257.

See, Y. X., Wang, B. Z. & Fullwood, M. J. 2019. Chromatin interactions and regulatory elements in cancer: from bench to bedside. *Trends in Genetics,* 35**,** 145-158.

Shima, Y., Yumoto, M., Katsumoto, T. & Kitabayashi, I. 2017. MLL is essential for NUP98-HOXA9-induced leukemia. *Leukemia,* 31**,** 2200-2210.

Shlush, L. I., Zandi, S., Mitchell, A., Chen, W. C., Brandwein, J. M., Gupta, V., Kennedy, J. A., Schimmer, A. D., Schuh, A. C. & Yee, K. W. 2014. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature,* 506**,** 328-333.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., De Wit, E., Van Steensel, B. & De Laat, W. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nature genetics,* 38**,** 1348-1354.

Splinter, E., De Wit, E., Van De Werken, H. J., Klous, P. & De Laat, W. 2012. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods,* 58**,** 221-230.

Subramanyam, D., Belair, C. D., Barry-Holson, K. Q., Lin, H., Kogan, S. C., Passegué, E. & Blelloch, R. 2010. PML-RARα and Dnmt3a1 cooperate in vivo to promote acute promyelocytic leukemia. *Cancer research,* 70**,** 8792-8801.

Sun, Y., Chen, B.-R. & Deshpande, A. 2018a. Epigenetic regulators in the development, maintenance, and therapeutic targeting of acute myeloid leukemia. *Frontiers in oncology,* 8**,** 41.

Sun, Y., Zhou, B., Mao, F., Xu, J., Miao, H., Zou, Z., Jang, Y., Cai, S., Witkin, M. & Koche, R. 2018b. HOXA9 reprograms the enhancer landscape to promote leukemogenesis. *Cancer cell,* 34**,** 643-658. e5.

Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L. & Spitz, F. 2014. Functional and topological characteristics of mammalian regulatory domains. *Genome research,* 24**,** 390-400.

Szabo, Q., Bantignies, F. & Cavalli, G. 2019. Principles of genome folding into topologically associating domains. *Science advances,* 5**,** eaaw1668.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R. & Aravind, L. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**,** 930-935.

Takakura, K., Kawamura, A., Torisu, Y., Koido, S., Yahagi, N. & Saruta, M. 2019. The clinical potential of oligonucleotide therapeutics against pancreatic cancer. *International journal of molecular sciences,* 20**,** 3331.

Tanaka, S., Miyagi, S., Sashida, G., Chiba, T., Yuan, J., Mochizuki-Kashio, M., Suzuki, Y., Sugano, S., Nakaseko, C. & Yokote, K. 2012. Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. *Blood,* 120**,** 1107-1117.

Thiecke, M. J., Wutz, G., Muhar, M., Tang, W., Bevan, S., Malysheva, V., Stocsits, R., Neumann, T., Zuber, J. & Fraser, P. 2020. Cohesin-Dependent and-Independent Mechanisms Mediate Chromosomal Contacts between Promoters and Enhancers. *Cell Reports,* 32**,** 107929.

Thiel, A. T., Feng, Z., Pant, D. K., Chodosh, L. A. & Hua, X. 2013. The trithorax protein partner menin acts in tandem with EZH2 to suppress C/EBPα and differentiation in MLL-AF9 leukemia. *Haematologica,* 98**,** 918-927.

Tholouli, E., Macdermott, S., Hoyland, J., Yin, J. L. & Byers, R. 2012. Quantitative multiplex quantum dot in-situ hybridisation based gene expression profiling in tissue microarrays identifies prognostic genes in acute myeloid leukaemia. *Biochemical and biophysical research communications,* 425**,** 333-339.

Thorsteinsdottir, U., Mamo, A., Kroon, E., Jerome, L., Bijl, J., Lawrence, H. J., Humphries, K. & Sauvageau, G. 2002. Overexpression of the myeloid leukemia–associated Hoxa9 gene in bone marrow cells induces stem cell expansion. *Blood, The Journal of the American Society of Hematology,* 99**,** 121-129.

Tijchon, E., Yi, G., Mandoli, A., Smits, J. G., Ferrari, F., Heuts, B. M., Wijnen, F., Kim, B., Janssen-Megens, E. M. & Schuringa, J. J. 2019. The acute myeloid leukemia associated AML1-ETO fusion protein alters the transcriptome and cellular progression in a single-oncogene expressing in vitro induced pluripotent stem cell based granulocyte differentiation model. *PloS one,* 14.

Tyner, J. W., Tognon, C. E., Bottomly, D., Wilmot, B., Kurtz, S. E., Savage, S. L., Long, N., Schultz, A. R., Traer, E. & Abel, M. 2018. Functional genomic landscape of acute myeloid leukaemia. *Nature,* 562**,** 526-531.

Volkova, M. & Russell, R. 2011. Anthracycline cardiotoxicity: prevalence, pathogenesis and treatment. *Current cardiology reviews,* 7**,** 214-220.

Wagner, K., Damm, F., Göhring, G., Görlich, K., Heuser, M., Schäfer, I., Ottmann, O., Lübbert, M., Heit, W. & Kanz, L. 2010. Impact of IDH1 R132 mutations and an IDH1 single nucleotide polymorphism in cytogenetically normal acute myeloid leukemia: SNP rs11554137 is an adverse prognostic factor. *Journal of clinical oncology,* 28**,** 2356-2364.

Walter, R. & Appelbaum, F. 2012. Acute myeloid leukemia stem cells and CD33-targeted immunotherapy. *Blood,* 119**,** 6198-6208.

Wang, A. J., Han, Y., Jia, N., Chen, P. & Minden, M. D. 2020. NPM1c impedes CTCF functions through cytoplasmic mislocalization in acute myeloid leukemia. *Leukemia,* 34**,** 1278-1290.

Wang, G. G., Pasillas, M. P. & Kamps, M. P. 2005. Meis1 programs transcription of FLT3 and cancer stem cell character, using a mechanism that requires interaction with Pbx and a novel function of the Meis1 C-terminus. *Blood,* 106**,** 254-264.

Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B. R., Protacio, A., Flynn, R. A., Gupta, R. A., Wysocka, J., Lei, M., Dekker, J., Helms, J. A. & Chang, H. Y. 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature,* 472**,** 120-124.

Wang, Q., Li, Y., Dong, J., Li, B., Kaberlein, J. J., Zhang, L., Arimura, F. E., Luo, R. T., Ni, J. & He, F. 2014. Regulation of MEIS1 by distal enhancer elements in acute leukemia. *Leukemia,* 28**,** 138-146.

Wang, S., Su, J.-H., Beliveau, B. J., Bintu, B., Moffitt, J. R., Wu, C.-T. & Zhuang, X. 2016. Spatial organization of chromatin domains and compartments in single chromosomes. *Science,* 353**,** 598-602.

Wang, X., Cairns, M. J. & Yan, J. 2019. Super-enhancers in transcriptional regulation and genome organization. *Nucleic acids research,* 47**,** 11481-11496.

Wang, Y., Wu, N., Liu, D. & Jin, Y. 2017. Recurrent fusion genes in leukemia: an attractive target for diagnosis and treatment. *Current genomics,* 18**,** 378-384.

Wei, C.-L., Nicolis, S. K., Zhu, Y. & Pagin, M. 2019. Sox2-dependent 3D chromatin Interactomes in transcription, neural stem cell proliferation and neurodevelopmental diseases. *Journal of experimental neuroscience,* 13**,** 1179069519868224.

Weissmann, S., Alpermann, T., Grossmann, V., Kowarsch, A., Nadarajah, N., Eder, C., Dicker, F., Fasan, A., Haferlach, C. & Haferlach, T. 2012. Landscape of TET2 mutations in acute myeloid leukemia. *Leukemia,* 26**,** 934-942.

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I. & Young, R. A. 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell,* 153**,** 307-319.

Wigle, T. J., Knutson, S. K., Jin, L., Kuntz, K. W., Pollock, R. M., Richon, V. M., Copeland, R. A. & Scott, M. P. 2011. The Y641C mutation of EZH2 alters substrate specificity for histone H3 lysine 27 methylation states. *FEBS letters,* 585**,** 3011-3014.

Wit, E. D. & Laat, W. D. 2012. A decade of 3C technologies-insights into nuclear organization. *Genes & development***,** 11-24.

Wong, P., Iwasaki, M., Somervaille, T. C., So, C. W. E. & Cleary, M. L. 2007. Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. *Genes & development,* 21**,** 2762-2774.

Wouters, B. J. & Delwel, R. 2016. Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. *Blood,* 127**,** 42-52.

Xiang, J.-F., Yin, Q.-F., Chen, T., Zhang, Y., Zhang, X.-O., Wu, Z., Zhang, S., Wang, H.-B., Ge, J. & Lu, X. 2014. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell research,* 24**,** 513-531.

Xie, H., Peng, C., Huang, J., Li, B. E., Kim, W., Smith, E. C., Fujiwara, Y., Qi, J., Cheloni, G. & Das, P. P. 2016. Chronic myelogenous leukemia–

initiating cells require Polycomb group protein EZH2. *Cancer discovery,* 6**,** 1237-1247.

Xu, K., Wu, Z. J., Groner, A. C., He, H. H., Cai, C., Lis, R. T., Wu, X., Stack, E. C., Loda, M. & Liu, T. 2012. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. *Science,* 338**,** 1465-1469.

Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S.-H., Ito, S., Yang, C., Wang, P. & Xiao, M.-T. 2011. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer cell,* 19**,** 17-30.

Yang, M., Vesterlund, M., Siavelis, I., Moura-Castro, L. H., Castor, A., Fioretos, T., Jafari, R., Lilljebjörn, H., Odom, D. T. & Olsson, L. 2019. Proteogenomics and Hi-C reveal transcriptional dysregulation in high hyperdiploid childhood acute lymphoblastic leukemia. *Nature communications,* 10**,** 1-15.

Yun, J. P., Chew, E. C., Liew, C. T., Chan, J. Y., Jin, M. L., Ding, M. X., Fai, Y. H., Li, H. R., Liang, X. M. & Wu, Q. L. 2003. Nucleophosmin/B23 is a proliferate shuttle protein associated with nuclear matrix. *Journal of cellular biochemistry,* 90**,** 1140-1148.

Zeisig, B. B., Milne, T., García-Cuéllar, M.-P., Schreiner, S., Martin, M.-E., Fuchs, U., Borkhardt, A., Chanda, S. K., Walker, J. & Soden, R. 2004. Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Molecular and cellular biology,* 24**,** 617-628.

Zhan, Y., Mariani, L., Barozzi, I., Schulz, E. G., Blüthgen, N., Stadler, M., Tiana, G. & Giorgetti, L. 2017. Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome research,* 27**,** 479-490.

Zhang, X., Zhang, Y., Zhu, X., Purmann, C., Haney, M. S., Ward, T., Khechaduri, A., Yao, J., Weissman, S. M. & Urban, A. E. 2018. Local and global chromatin interactions are altered by large genomic deletions associated with human brain development. *Nature communications,* 9**,** 1-15.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M. & Li, W. 2008a. Model-based analysis of ChIP-Seq (MACS). *Genome biology,* 9**,** R137.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W. & Liu, X. S. 2008b. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology,* 9**,** R137.

Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K. S. & Singh, U. 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra-and interchromosomal interactions. *Nature genetics,* 38**,** 1341-1347.

Zhong, X., Prinz, A., Steger, J., Garcia-Cuellar, M.-P., Radsak, M., Bentaher, A. & Slany, R. K. 2018. HoxA9 transforms murine myeloid cells by a feedback loop driving expression of key oncogenes and cell cycle control genes. *Blood advances,* 2**,** 3137-3148.

Zhou, J., Jia, Y., Tan, T. K., Chung, T.-H., Sanda, T. & Chng, W. J. 2019. Super-Enhancer Profiling Identifies Novel Oncogenes and Therapeutic Targets in Multiple Myeloma. American Society of Hematology Washington, DC.

Zhou, L., Fu, L., Lv, N., Liu, J., Li, Y., Chen, X., Xu, Q., Chen, G., Pang, B. & Wang, L. 2018. Methylation-associated silencing of BASP1 contributes to leukemogenesis in t (8; 21) acute myeloid leukemia. *Experimental & molecular medicine,* 50**,** 1-8.

Zuin, J., Dixon, J. R., Van Der Reijden, M. I., Ye, Z., Kolovos, P., Brouwer, R. W., Van De Corput, M. P., Van De Werken, H. J., Knoch, T. A. & Van Ijcken, W. F. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences,* 111**,** 996-1001.

10 Supplementary Figures

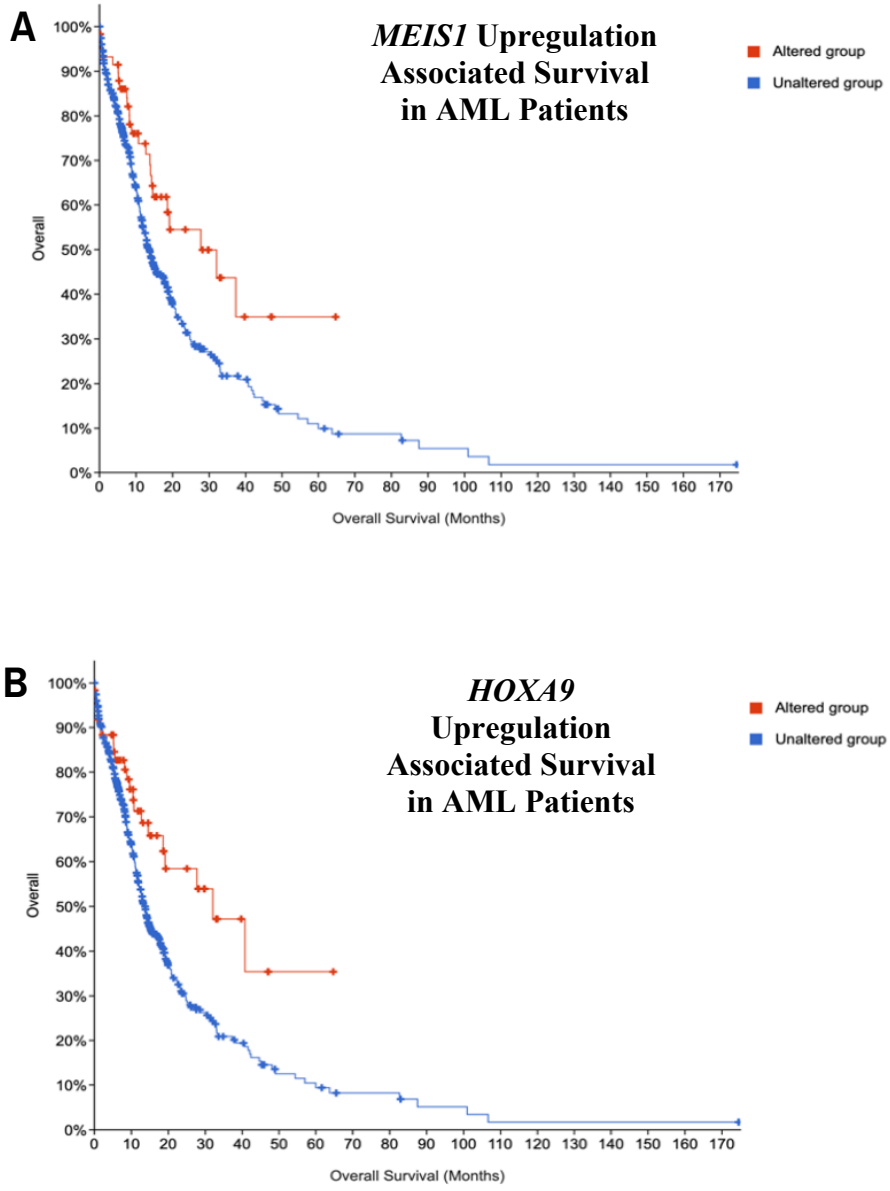

Figure 38. Survival curve of acute myeloid leukaemia patients with *MEIS1* and *HOXA9* alterations. A. Survival curve comparison the overall survival of *MEIS1* altered patients and non-altered acute myeloid leukaemia patients. B. Survival curve comparison the overall survival of *HOXA9* altered patients and non-altered acute myeloid leukaemia patients.

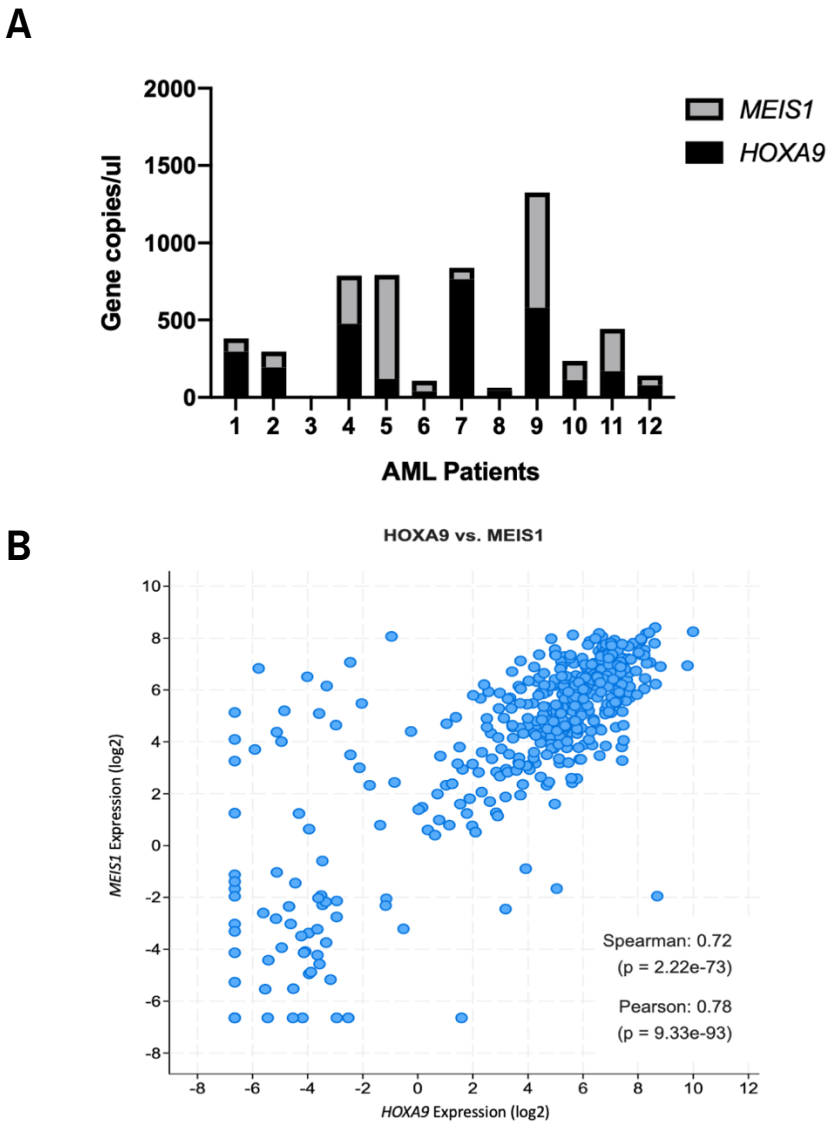

Figure 39. Expression levels of *MEIS1* and *HOXA9* in individual clinical acute myeloid leukaemia samples. A. Proportion of *MEIS1* and *HOXA9* expression levels in 12 clinical AML samples. B. Scatterplot of *HOXA9* and *MEIS1* expression levels in AML patients and their correlation analysis. X and Y axis represent the log2 function of *MEIS1* and *HOXA9* expression levels in the AML clinical samples (Tyner et al., 2018). Scatterplot conducted using the cbioportal (Cerami et al., 2012, Gao et al., 2013) .