MYC overexpression in cancer increases chromatin interactions at superenhancers and c-Myc binding sites

See, Yi Xiang

2021

See, Y. X. (2021). MYC overexpression in cancer increases chromatin interactions at superenhancers and c-Myc binding sites. Doctoral thesis, Nanyang Technological University, Singapore. <https://hdl.handle.net/10356/154939>

<https://hdl.handle.net/10356/154939>

<https://doi.org/10.32657/10356/154939>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 12:52:19 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

***MYC* overexpression in cancer increases chromatin
interactions at superenhancers and c-Myc binding sites**

SEE YI XIANG

SCHOOL OF BIOLOGICAL SCIENCES

2021

***MYC* overexpression in cancer increases chromatin
interactions at superenhancers and c-Myc binding sites**

SEE YI XIANG

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

2021

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.

20/08/2021

.....

Date

[Handwritten signature]

SEE YI XIANG

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.

20/08/2021

.....

Date _____

Dr. Melissa Jane Fullwood

Authorship Attribution Statement

This thesis contains material from 2 papers published or in revision in the following peer-reviewed journals in which I am listed as an author.

Chapter 1 contains material published in 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)

doi: <https://doi.org/10.1016/j.tig.2018.11.007>

The contributions of the co-authors are as follows:

- Dr. Melissa Fullwood provided guidance for the manuscript content and edited the manuscript drafts.
- Dr. Benny Wang Zheng Jie and I prepared the manuscript drafts and revisions together with equal levels of contribution.

Chapter 4 and 5 are published as See, Y.X., Chen, K., Fullwood, M.J., 2021. MYC overexpression leads to increased chromatin interactions at superenhancers and c-Myc binding sites. bioRxiv, 2021.2001.2004.425344. (Manuscript in revision at Genome Research)

doi: <https://doi.org/10.1101/2021.01.04.425344>

The contributions of the co-authors are as follows:

- Dr. Melissa Fullwood provided guidance for the project and edited the manuscript drafts.
- I prepared the manuscript drafts and revisions.
- I performed the following experiments for the manuscript: Western blot, quantitative Polymerase Chain Reaction (qPCR), RNA-Seq, ChIP-Seq, 4C-Seq and SIQHiC.
- I performed the RNA-Seq, ChIP-Seq and superenhancer bioinformatics analyses.
- I performed the SIQHiC bioinformatics analyses together with Ms. Chen Kaijing.

- The 4C-Seq analyses were performed using the Cancer Science Institute NGS Portal (<https://csibioinfo.nus.edu.sg/csingsportal/login/home.php>) designed and developed by Dr. Ömer An.

20/08/2021

.....
Date

.....
See Yi Xiang

Table of Contents

Statement of Originality	iii
Supervisor Declaration Statement	iv
Authorship Attribution Statement	v
ACKNOWLEDGEMENTS	x
LIST OF FIGURES	xi
LIST OF TABLES	xiv
ABBREVIATIONS	xv
SUMMARY	xvi
Chapter 1: Introduction	1
1.1 <i>MYC: master regulator and oncogene</i>	<i>1</i>
1.2 <i>c-Myc dysregulation in cancer</i>	<i>2</i>
1.3 <i>Targeting c-Myc in cancer</i>	<i>3</i>
1.4 <i>3D genome organization</i>	<i>5</i>
1.5 <i>Chromatin Conformation Capture Techniques</i>	<i>7</i>
1.6 <i>Enhancers and superenhancers</i>	<i>10</i>
1.7 <i>Enhancer and chromatin interaction alterations in cancer</i>	<i>13</i>
1.8 <i>Superenhancers and chromatin interactions as cancer biomarkers and therapeutic targets</i>	<i>16</i>
1.9 <i>c-Myc enhancer invasion in cancer</i>	<i>17</i>
1.10 <i>Aims and hypotheses</i>	<i>18</i>
Chapter 2: Materials and Methods	21
2.1 <i>Cell culture</i>	<i>21</i>
2.2 <i>Reverse transcription and quantitative polymerase chain reaction (qPCR)</i>	<i>21</i>
2.3 <i>RNA-Seq</i>	<i>22</i>
2.4 <i>Protein extraction and immunoblotting</i>	<i>22</i>
2.5 <i>Chromatin Immunoprecipitation followed by sequencing (ChIP-seq)</i>	<i>23</i>
2.6 <i>Circular chromosome conformation capture (4C)</i>	<i>24</i>
2.7 <i>Spike-in quantitative Hi-C (SIQHiC)</i>	<i>26</i>
2.8 <i>RNA-Seq data processing</i>	<i>26</i>
2.9 <i>ChIP-seq data processing</i>	<i>26</i>
2.10 <i>4C data processing</i>	<i>28</i>

2.11	<i>SIQHiC data processing</i>	28
Chapter 3: Characterizing c-Myc binding and chromatin interaction		
	profiles in high c-Myc expressing cancer cell lines	31
3.1	<i>c-Myc occupies superenhancers in high c-Myc cancer cell lines</i>	31
3.2	<i>c-Myc bound superenhancers interact with cancer-associated genes and nearby c-Myc binding sites</i>	32
3.3	<i>MYC shRNA knockdown reduced c-Myc occupancy at enhancers</i>	35
3.4	<i>MYC shRNA knockdown decreased constituent H3K27ac peaks at superenhancers and typical enhancers</i>	38
3.5	<i>Summary</i>	40
Chapter 4: MYC overexpression leads to superenhancer invasion		
41		
4.1	<i>Doxycycline induction of MYC expression in U2OS cells regulated genes involved in cancer progression</i>	41
4.2	<i>MYC overexpression increased constituent H3K27ac peaks within existing superenhancers</i>	46
4.3	<i>Superenhancers are preferentially enriched for c-Myc binding after MYC overexpression</i>	48
4.4	<i>c-Myc binding requires H3K27ac, together with canonical motif and/or G/C rich sequences</i>	52
4.5	<i>Summary</i>	55
Chapter 5: Spike-in quantitative Hi-C (SIQHiC) reveals increased chromatin interactions after MYC overexpression		
56		
5.1	<i>Spike-in Quantitative Hi-C (SIQHiC) uses mouse cell spike-in to reveal global changes in chromatin contact frequency</i>	56
5.2	<i>MYC overexpression increases global chromatin contact frequency</i>	58
5.3	<i>MYC overexpression increases chromatin contact frequency specifically at a set of gained loops</i>	61
5.4	<i>Chromatin interactions are particularly enriched at superenhancers and tend to connect c-Myc binding sites</i>	71
5.5	<i>MYC overexpression increased chromatin interaction frequency regardless of transcription activity</i>	76
5.6	<i>Summary</i>	79
Chapter 6: Conclusions and future directions		
81		
6.1	<i>c-Myc preferentially invades into active superenhancers</i>	81
6.2	<i>Spike-in Quantitative Hi-C normalizes Hi-C samples by cell count to detect global changes in chromatin contact frequency</i>	82
6.3	<i>SIQHiC revealed increased chromatin contact frequency between c-Myc binding sites at enhancers and promoters</i>	83

6.4	<i>Significance</i>	85
6.5	<i>Future directions</i>	85
References		89

ACKNOWLEDGEMENTS

I am deeply indebted to my PhD supervisor Assistant Professor Melissa Fullwood, who provided me with the opportunity to pursue my PhD, and gave me much guidance and support throughout the arduous journey. I would like to express my sincere gratitude to my Thesis Advisory Committee – Associate Professor Valerie Lin, Assistant Professor Amartya Sanyal and Assistant Professor Zhang Li-Feng – for their professional advice and encouragement.

I would like to thank Nanyang Technological University (NTU) for the financial assistance from the NTU Research Scholarship, especially for granting me an extension of candidature and waiver of school fees for one semester amidst the Covid-19 pandemic. In addition, I would like to thank the Cancer Science Institute of Singapore, National University of Singapore for providing the resources and optimal environment for research. I would also like to thank the funding agencies for this project: the National Research Foundation (NRF) Singapore and the Ministry of Education, Singapore.

I am grateful to have had a bunch of wonderful colleagues with whom I could always share my joys and frustrations. In particular, I would like to thank Ms. Chen Kaijing for performing the bioinformatics analyses for my manuscript, as well as Dr. Cai Yichao, Dr. Cao Fan, Dr. Luay Aswad and Ms. Kong Lingshi for assisting me whenever I needed help with bioinformatics. Special thanks to Dr. Tan Hong Kee, Dr. Tan Shi Hao, Dr. Benny Wang, Dr. Zhang Ying and Ms. Lim Mei Chee for the many insightful conversations that helped advance this project.

Finally, the completion of my PhD dissertation would not have been possible without my beloved family. I would like to thank my parents, See Ling Soon and Ng Siew Nui, for their unconditional love and support towards all my endeavours, especially when I left Singapore for studies and work. Most importantly, I thank my wife, Lai Si Kei, for supporting my decision to pursue doctoral studies despite having to temporarily stay apart, and for marrying me and moving to Singapore amidst the many uncertainties.

LIST OF FIGURES

Figure 1 Chromatin looping forms higher-order structures to package DNA in cells	7
Figure 2 Chromatin Conformation Capture techniques.....	9
Figure 3 Workflow for identifying superenhancers.....	12
Figure 4 Frequency of mutations at different genomic elements	13
Figure 5 Enhancer and chromatin interaction alterations contribute to oncogenesis.....	16
Figure 6 Diagrams illustrating the hypotheses of this thesis.....	20
Figure 7 c-Myc occupies superenhancers particularly in high c-Myc expressing cancer cell lines.....	32
Figure 8 c-Myc bound superenhancers interact with cancer associated genes and other c-Myc binding sites in high c-Myc expressing cancers	35
Figure 9 <i>MYC</i> shRNA knockdown reduced c-Myc occupancy at typical enhancers and superenhancers.....	38
Figure 10 <i>MYC</i> shRNA knockdown reduced the number of H3K27ac peaks within superenhancers	40
Figure 11 Superenhancers are bound by c-Myc in high c-Myc expressing cancers and lose c-Myc binding after <i>MYC</i> shRNA knockdown.	40
Figure 12 Doxycycline treatment induces <i>MYC</i> overexpression in U2OS cells	42
Figure 13 <i>MYC</i> overexpression increased c-Myc binding at differentially expressed transcripts	44
Figure 14 <i>MYC</i> overexpression modulated H3K27 acetylation and increased constituent H3K27ac peaks at superenhancers	47
Figure 15 Superenhancers are preferentially enriched in c-Myc binding sites after <i>MYC</i> overexpression.....	51
Figure 16 c-Myc binds next to H3K27ac peaks at CREs.....	52
Figure 17 Promoter c-Myc binding sites are enriched for other transcription factor motifs but not enhancer c-Myc binding sites	55
Figure 18 Superenhancers gain c-Myc binding and constituent H3K27ac peaks after <i>MYC</i> overexpression	55

Figure 19 Spike-in Quantitative Hi-C (SIQHiC) can detect global changes in chromatin contact frequency	58
Figure 20 SIQHiC normalization reveals increased chromatin contact frequency per cell after <i>MYC</i> overexpression.	60
Figure 21 <i>MYC</i> overexpression does not significantly alter chromatin compartments and topological associating domains (TADs).	62
Figure 22 <i>MYC</i> overexpression appears to decrease chromatin interaction frequency without normalization.	63
Figure 23 Normalization of APA heatmaps by SIQHiC and by local background signal reveal increased chromatin interaction frequency after <i>MYC</i> overexpression, particularly at gained loops.	64
Figure 24 Gained chromatin loops are enriched for genes differentially regulated by c-Myc.	66
Figure 25 Gained loops are associated with more significantly regulated transcripts.	67
Figure 26 <i>MYC</i> overexpression increases chromatin contact frequency at the <i>MYC</i> gene locus	69
Figure 27 <i>MYC</i> overexpression increases chromatin contact frequency and chromatin loops at superenhancers	70
Figure 28 Gained chromatin loops are enriched for c-Myc and tend to connect c-Myc binding sites together	72
Figure 29 Overexpressed c-Myc formed discrete punctate spots overlapping RNA polymerase II within the nucleus.....	73
Figure 30 c-Myc bound cis-regulatory elements preferentially gain chromatin loops.....	74
Figure 31 <i>MYC</i> overexpression increases chromatin interactions between c-Myc bound enhancers and promoters	75
Figure 32 <i>MYC</i> overexpression connected cancer-associated genes to neighbouring c-Myc bound enhancers.....	76
Figure 33 c-Myc bound transcripts gain chromatin loops regardless of differential transcript expression	78
Figure 34 c-Myc binding increases chromatin interaction frequency regardless of transcript expression levels	79

Figure 35 | c-Myc concentrates at localized regions of high transcriptional activity80

LIST OF TABLES

Table 1 qPCR Primer sequences	22
Table 2 Protein immunoblotting antibodies	23
Table 3 ChIP antibodies	24
Table 4 4C viewpoints and primer sequences	25
Table 5 ChIP-seq datasets downloaded from NCBI GEO	27
Table 6 Upregulated gene sets after <i>MYC</i> overexpression	45
Table 7 Downregulated gene sets after <i>MYC</i> overexpression	45
Table 8 Human to mouse chromatin contact ratio (HMR)	59

ABBREVIATIONS

3C	Chromatin conformation capture
4C	Circularised chromosome conformation capture
5C	Carbon copy chromosome conformation capture
AML	Acute myeloid leukaemia
BET	Bromodomain and extraterminal domain
bHLH-ZIP	Basic helix-loop-helix leucine zipper
CDK	Cyclin dependent kinases
CDS	coding sequence
ChIA-PET	Chromatin interaction analysis by paired-end tag sequencing
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation followed by sequencing
CRE	Cis-regulatory element
eRNA	Enhancer RNA
GEO	Gene expression omnibus
H3K27ac	Acetylation of histone H3 at lysine 27
H3K4me1	Monomethylation of histone H3 at lysine 4
LLPS	Liquid-liquid phase separation
PCR	Polymerase chain reaction
RT-qPCR	Real-time quantitative PCR
siRNA	Short interfering RNA
shRNA	Short hairpin RNA
SIQHiC	Spike-In quantitative Hi-C
TAD	Topologically associating domain
TCGA	The Cancer Gene Atlas

SUMMARY

The *MYC* oncogene encodes for the c-Myc protein and is frequently dysregulated across multiple cancer cell types, making it an attractive target for cancer therapy. There have been many difficulties in targeting c-Myc, due to its complex network of regulators and the unstructured nature of its protein. Thus, we are interested in looking at the downstream cancer-specific functions of c-Myc. Overexpression of *MYC* leads to c-Myc binding at active enhancers, resulting in a global transcriptional amplification of active genes. Since superenhancers are frequently dysregulated in cancer, we hypothesize that c-Myc preferentially invades into superenhancers and alters the cancer genome organization. To that end, we performed ChIP-seq, RNA-seq, 4C-seq and SIQHiC (Spike-in Quantitative Hi-C) on the U2OS osteosarcoma cell line with tetracycline-inducible *MYC*. *MYC* overexpression in U2OS cells modulated histone acetylation and increased c-Myc binding at superenhancers. SIQHiC analysis revealed increased global chromatin contact frequency, particularly at chromatin interactions connecting c-Myc binding sites. Our results suggest that c-Myc molecules are recruited to and accumulates within zones of high transcription activity, binding first at stable promoter binding sites at low expression levels, then at superenhancer binding sites when overexpressed. At the same time, the recruitment of c-Myc and other transcription factors may stabilize chromatin interactions to increase chromatin contact frequency. The accumulation of c-Myc at cancer-type specific superenhancers may then drive the expression of interacting oncogenes that each cancer is highly reliant on. By elucidating the chromatin landscape of c-Myc driven cancers, we can potentially target these chromatin interactions for cancer therapy, without affecting physiological c-Myc signalling.

Chapter 1: Introduction

1.1 *MYC*: master regulator and oncogene

The *MYC* oncogene encodes for the c-Myc transcription factor, a basic helix-loop-helix leucine zipper (bHLH-ZIP) transcription factor that is essential in regulating pathways important to cell growth, metabolism, cell cycle progression and apoptosis (Eilers and Eisenman, 2008).

c-Myc heterodimerizes with the Max protein to bind DNA, canonically at “CACGTG” E-box sequences (Blackwell et al., 1990; Blackwood and Eisenman, 1991). However, subsequent papers have shown that the c-Myc-Max complex also binds to other “CANNTG” E-box sequences (Blackwell et al., 1993) as well as non E-box sequences (Allevato et al., 2017; Guo et al., 2014; Kasibhatla et al., 2000), suggesting that c-Myc is recruited to its binding sites by other transcription factors. For example, c-Myc has been shown to interact with RNA polymerase II (Koch et al., 2007), the P-TEFb complex and TFII-F (Kalkat et al., 2018). Additionally, c-Myc co-binds to the genome with transcription factors such as YY1 and GABP (Wang et al., 2012).

c-Myc binds directly to and activates the expression of downstream target genes through two main mechanisms. c-Myc opens up the chromatin structure to facilitate transcription by recruiting nucleosome remodellers such as the SWI/SNF complex (Cheng et al., 1999), or recruiting histone acetyltransferase complexes such as the STAGA complex (Liu et al., 2003). c-Myc also regulates RNA polymerase II transcription initiation by recruiting TFII-H (Cowling and Cole, 2007), as well as transcription pause release by recruiting the P-TEFb complex (Rahl et al., 2010).

In quiescent cells, mitogenic signalling activates *MYC* expression to drive cell cycle progression from G1 to S-phase by inducing the expression of important cell cycle regulators, such as cyclin D, CDK4, CDK6 and the E2F family of genes, while inhibiting cell cycle checkpoint inhibitors, such as p15/INK4B, p21/CIP1 and p27/KIP1 (Bretones et al., 2015). c-Myc activates metabolic

transcriptional programs, such as the glycolysis and glutaminolysis pathways, to meet the increased cellular demands for energy and biosynthesis substrates (Dang, 2013). c-Myc also induces the expression of genes involved in nucleotide, protein and fatty acid synthesis, as well as ribosome and mitochondrial biogenesis (Dang, 2013). The activation of these pathways promotes cell growth and proliferation.

At the same time, c-Myc curtails the uncontrolled proliferation of normal cells by sensitizing these cells towards apoptosis and senescence. c-Myc upregulates p53 expression through upregulating p14/ARF (Hermeking and Eick, 1994), inducing apoptosis and cellular senescence. c-Myc can also induce apoptosis independent of p53 by upregulating BIM (Campone et al., 2011; Muthalagu et al., 2014).

1.2 c-Myc dysregulation in cancer

Dysregulation of c-Myc is common across most cancer types in The Cancer Gene Atlas (TCGA) dataset (Schaub et al., 2018). However, c-Myc dysregulation can be caused by a variety of mutations. First, focal amplification of the *MYC* gene occurs most frequently across different cancer cell types, especially in solid tumours such as oesophageal, gastric, ovarian and breast cancers (Beroukhi et al., 2010), and this amplification is correlated with overexpression of both mRNA and protein (Blancato et al., 2004). Second, anomalies in the network of oncogenic signalling pathways regulating *MYC*, such as the Notch1, Wnt/ β -catenin and K-Ras/MAPK pathways, often lead to increased transcription, translation and protein stability (Kress et al., 2015). This happens frequently in pancreatic ductal adenocarcinomas, where K-Ras is mutated in more than 90% of cases and induces overexpression of c-Myc. Third, *MYC* can also acquire new enhancer regulatory elements, most notably through the translocation of immunoglobulin enhancers to its gene locus via class switch recombination (Küppers and Dalla-Favera, 2001) and the formation of a superenhancer as a result of overexpression of other transcription factors (Hnisz et al., 2013). *MYC* locus rearrangements are detected in around half of multiple myelomas, frequently moving *MYC* in close proximity to enhancers and superenhancers

(Affer et al., 2014). Fourth, mutations in the coding region of *MYC* are common, especially in lymphomas, and can affect protein activation, stability and degradation (Bahram et al., 2000; Bhatia et al., 1995).

Besides the over-activation of the cell growth and proliferation pathways, c-Myc dysregulation also contributes to tumorigenesis by circumventing the usual limitations to cell proliferation. For example, c-Myc activates hTERT, which increases the tumour capacity for cell replication by extending telomere ends and maintaining genome stability (Khattar and Tergaonkar, 2017; Wu et al., 1999). c-Myc also works together with HIF proteins to promote glycolysis and angiogenesis, thereby ensuring sufficient energy production in hypoxic conditions (Li et al., 2020b).

c-Myc dysregulation drives cell growth and proliferation, but is usually not sufficient by itself to cause tumorigenesis. *MYC* overexpression in normal cells generally result in apoptosis and senescence (Evan et al., 1992; Hoffman and Liebermann, 2008). Other genetic or epigenetic mutations in the apoptotic and senescence pathways are usually required to escape these fail-safe mechanisms. For example, c-Myc overexpression is frequently accompanied by overexpression of BCL2, an anti-apoptotic protein, in diffuse large B-cell lymphomas (Aukema et al., 2011; Li et al., 2013). Overexpression of c-Myc also applies selective pressure for the loss of p53 and p14/ARF function in carcinogenesis (Beer et al., 2004; Clegg et al., 2011). Additionally, c-Myc and RAS have been shown to suppress and circumvent the apoptosis and senescence functions of each other (Hydbring et al., 2010; Kauffmann-Zeh et al., 1997).

1.3 Targeting c-Myc in cancer

Since c-Myc dysregulation is a common theme in different cancer cell types, a lot of research has been conducted on targeting c-Myc for pan-cancer therapy. However, research on direct inhibition of c-Myc has been fraught with many difficulties. The c-Myc protein itself is notoriously termed as “undruggable”, because of the unstructured nature of the protein and the lack of enzymatic binding sites or prominent pockets for small molecule inhibitor binding

(McKeown and Bradner, 2014). Inhibitors such as Omomyc (Soucek et al., 2002), 10058-F4 and 10074-G5 (Yin et al., 2003) can target the c-Myc – Max protein interaction, but potency remains low *in vivo* (Guo et al., 2009). Antisense oligonucleotides and small interfering RNAs have been used to target *MYC* mRNA transcripts, but these approaches suffer from poor delivery efficiency and rapid degradation (Whitfield et al., 2017). Additionally, the essential role of c-Myc in normal cell physiology means that toxicity is a concern in the clinical setting.

Synthetic lethality approaches have also been studied against c-Myc overexpressing cancers, targeting genes that these cancers are highly reliant on, but are not as essential in normal cells (Thng et al., 2021). These include genes involved in cell cycle and DNA repair. For example, synthetic lethality has been observed with inhibition of CDK1 (Goga et al., 2007; Horiuchi et al., 2012; Kang et al., 2014), which is essential for cell cycle and participates in DNA repair (Liao et al., 2017; Santamaría et al., 2007). Inhibition of the DNA damage response protein CHK1 has also been explored as a potential synthetic lethal target for high c-Myc expressing cancers (Cole et al., 2011; Höglund et al., 2011; Russell et al., 2013). However, these candidates may not exhibit synthetic lethality across all c-Myc overexpressed cancers. Although 2 synthetic lethal screens had identified common functional pathways that c-Myc driven cancers are vulnerable to, different candidate synthetic lethal genes were identified (Kessler Jessica et al., 2012; Toyoshima et al., 2012), partly due to the different tissue types that were used.

Indirect inhibition of c-Myc is challenging as well, due to the complex network of transcriptional and translational regulators and enhancer elements controlling its expression (Koh et al., 2016). Many studies have targeted individual upstream regulators of c-Myc such as the KRAS, Wnt/ β -catenin, MAPK and PI3K pathways, but the effects are not c-Myc specific and do not affect c-Myc regulation by the other regulatory pathways (McKeown and Bradner, 2014).

For these reasons, instead of exploring strategies targeting c-Myc expression directly or indirectly, this thesis investigates the downstream consequences of c-

Myc overexpression common across different cancer cell types, specifically the role of superenhancers and long-range chromatin interactions in c-Myc activation of oncogenes.

In sections 1.4 and 1.5, I will introduce chromatin interactions and the current techniques to study them. I will review how chromatin interactions connect enhancer and promoter regulatory elements to regulate gene expression in section 1.6. I will discuss how alterations in the enhancer-ome and chromatin interactions contribute to oncogenesis in section 1.7, and how these alterations can be used as biomarkers or therapeutic targets in section 1.8. Finally, in section 1.9, I will introduce c-Myc enhancer invasion as a mechanism for oncogene activation.

1.4 3D genome organization

DNA needs to be tightly packed into the confines of a nucleus, yet still allow selected loci to be accessed by transcriptional factors and other DNA binding proteins. To accomplish that, DNA first winds around histones to form nucleosomes, then folds and loops in a regulated fashion to form duplex and complex interactions between linearly distant genomic loci (Babu and Fullwood, 2015) (**Figure 1A**). This chromatin looping brings cis-regulatory elements (CREs), such as promoters, enhancer and insulators, together into close spatial proximity, allowing them to act on one another to regulate transcription.

Further folding at megabase scales segregate chromatin into topologically associating domains (TADs), such that chromatin interactions preferentially occur within these domains rather than across domains (Dixon et al., 2012) (**Figure 1B**, left). TAD formation appears to be dependent on CTCF and cohesin binding (Nora et al., 2017; Nuebler et al., 2018; Rao et al., 2017; Sanborn et al., 2015; Schwarzer et al., 2017). In the loop extrusion model, chromatin passes through the ring-like structure of the cohesin complex, until the cohesin complex meets two CTCF molecules at the TAD boundaries (Rao et al., 2014). Many papers have shown that TADs are predominantly stable across cell types (Dixon et al., 2012; Dixon et al., 2015; Schmitt et al., 2016; Smith et al., 2016), but there is also evidence of cell-type specific TAD boundaries (Narendra et al., 2016;

Sauerwald et al., 2019).

The genome is also organized into transcriptionally active and open euchromatin regions, and transcriptionally silent and compact heterochromatin regions, known as A/B compartments respectively. Genomic loci from the same compartment tend to localize in the same space and preferentially interact with one another (Lieberman-Aiden et al., 2009). A/B compartments appear to be made up of units of TADs that preferentially associate with each other, forming a checkerboard pattern (**Figure 1B**, right) (Dixon et al., 2015). However, depletion of CTCF and cohesin complex proteins eradicated TADs but compartments are still observed (Nuebler et al., 2018; Rao et al., 2017; Schwarzer et al., 2017). Recent papers have suggested that active and inactive genomic loci separately and aggregate at localized concentrations of associated chromatin factors to form phase-separated compartments (Larson et al., 2017; Sabari et al., 2018; Strom et al., 2017), in contrast to the physical segregation of TADs by architectural proteins.

Figure 1 | Chromatin looping forms higher-order structures to package DNA in cells

(A) DNA is wrapped around histones to form nucleosomes, before folding and looping to form interactions. Figure reproduced with permission from Babu and Fullwood (2015). (B) Chromatin interaction are enriched within topologically associating domains (TADs) but depleted between TADs (left). Open, transcriptionally active domains and closed, transcriptionally silent domains form separate compartments (termed as A/B compartments respectively), whereby chromatin interactions preferentially form between like compartments (right). Figure reproduced with permission from Szabo et al. (2019).

1.5 Chromatin Conformation Capture Techniques

Chromosome conformation capture (3C) and related techniques have been developed in the past decade to study these chromatin interactions. In 3C (Dekker et al., 2002), chromatin regions in close proximity are cross-linked using formaldehyde, digested using restriction enzymes, and ligated under dilute conditions. This yields a 3C library containing circular DNA formed from interacting loci, and interaction frequency between known pairs of genomic loci can be quantified through real-time quantitative polymerase chain reaction (qPCR) methods using primers amplifying the ligation junction (one-to-one) (**Figure 2A**).

To assay all interactions involving a particular genomic locus of interest, termed as the viewpoint or bait, circular chromosome conformation capture (4C) (Simonis et al., 2006; van de Werken et al., 2012; Zhao et al., 2006) adds a second restriction enzyme digest and ligation step after 3C library generation to reduce the size of ligated circular DNA. This is followed by inverse PCR using primers at the viewpoint, thus amplifying its ligated interacting partners (one-to-many). These PCR products can then be assayed using microarrays or next generation sequencing to identify the interacting partners (**Figure 2B**).

To increase throughput, 3C-carbon copy (5C) (Dostie et al., 2006) was developed, where primers are designed at all digestion sites of a chosen restriction enzyme in a large genomic area, to amplify and assay all possible interactions within this area (many-to-many) (**Figure 2C**).

As sequencing costs decrease with the advent of next-generation sequencing, genome-wide chromatin interactions can be identified using Hi-C (all-to-all) (Lieberman-Aiden et al., 2009). After formaldehyde cross-linking and restriction enzyme digestion, 5' sticky ends are filled with biotinylated nucleotides, followed by blunt-end ligation of interacting fragments. This allows for enrichment of ligated fragments for sequencing via streptavidin pull-down.

Genome-wide chromatin interactions associated with specific proteins or genomic sequences can be specifically enriched for. In chromatin interaction analysis by paired-end tag sequencing (ChIA-PET), genome-wide interactions mediated by a protein of interest are identified, by combining 3C and chromatin immunoprecipitation (ChIP) techniques (Fullwood et al., 2009). Instead of restriction enzyme digest and biotin-labelling as in Hi-C, cross-linked chromatin is sonicated and immunoprecipitated with antibodies targeting the protein of interest, followed by proximity ligation of interacting fragments. On the other hand, Capture Hi-C (Dryden et al., 2014; Hughes et al., 2014) uses a custom panel of DNA probes to assay chromatin interactions associated with known genomic regions of interest, such as promoters (Mifsud et al., 2015) and enhancers (Madsen et al., 2020).

Even though techniques like Hi-C and ChIA-PET offer global profiles of chromatin interactions, the high sequencing depth required, coupled with high sequencing costs, means that these techniques suffer from lower resolution compared to 3C and 4C techniques (Lajoie et al., 2015). As such, these global approaches are optimally used for analysis of higher order chromatin structures, such as TADs and chromosome compartmentalization.

Figure 2 | Chromatin Conformation Capture techniques

(A) 3C quantifies an individual interaction between 2 genomic loci. (B) 4C identifies all interactions to a particular genomic loci, termed as the viewpoint or bait (shaded in yellow). (C) 5C assays all interactions within a large genomic region. (D) Hi-C identifies genome-wide chromatin interactions, including inter- and intra-chromosomal interactions. Figure reproduced with permission from See et al. (2019).

1.6 Enhancers and superenhancers

Enhancers are CREs that regulate the expression of gene(s) linearly distant from itself. They are characterised by open chromatin regions marked by monomethylation of histone H3 at lysine 4 (H3K4me1) and acetylation of histone H3 at lysine 27 (H3K27ac) (Creyghton et al., 2010). Enhancers are orientation independent and can regulate distant genes both upstream and downstream of itself (Banerji et al., 1983; Gillies et al., 1983), by forming chromatin interactions and recruiting transcription factors such as chromatin remodellers and transcriptional machinery to these gene promoters.

Enhancer-promoter chromatin interactions are stabilized by proteins such as the Mediator and cohesin complexes (Kagey et al., 2010). Recent papers have also identified YY1 as an architectural protein that binds to both enhancers and promoters to mediate chromatin interactions through homodimerization. (Beagan et al., 2017; Weintraub et al., 2017). Different subsets of enhancers are activated in different cell types, and their interactions with promoters are consequently cell-type specific as well, thereby playing an important role in cell-type specific control of gene expression (Heintzman et al., 2009; Javierre et al., 2016; Nord et al., 2013; Pellacani et al., 2016).

Active enhancers can be transcribed by RNA polymerase II, frequently bidirectionally, yielding short-living noncoding enhancer RNA (eRNA) (Andersson et al., 2014; Kim et al., 2010; Mikhaylichenko et al., 2018). The functions of eRNAs remain unclear, but several models have been proposed. eRNA expression may merely be an unintended consequence of RNA polymerase II loading on the enhancer instead of the promoter (Struhl, 2007). The process of transcribing eRNAs may remodel chromatin at the promoter to increase accessibility and activate transcription (Natoli and Andrau, 2012). eRNAs may also interact with proteins to recruit transcription activators (Wang et al., 2011) or act as decoys to sponge up transcription repressors (Klattenhoff et al., 2013). Lastly, eRNAs may mediate chromatin looping of the enhancer to activate gene promoters (Hsieh et al., 2014; Lai et al., 2013; Yang et al., 2016).

Clusters of enhancers, known as superenhancers, have been shown to work in tandem to accumulate transcriptional activators at genes that are crucial to cell identity (Hnisz et al., 2013; Loven et al., 2013; Whyte et al., 2013). Our lab had recently shown that superenhancers connect to more genomic loci through chromatin interactions compared to typical enhancers, regulating multiple genes that are linearly distant (Cao et al., 2017). These multivalent chromatin and protein interactions are proposed to create phase-separated condensates at superenhancers – membrane-less compartments of high protein density within the nucleoplasm (Hnisz et al., 2017). These superenhancer condensates sequester and concentrate transcription factors, such as BRD4, RNA polymerase II and the Mediator complex (Boehning et al., 2018; Cho et al., 2018; Lu et al., 2018; Sabari et al., 2018), which drives the transcription of interacting genes.

The earliest papers had first defined superenhancers by the enrichment of master transcription factors in embryonic stem cells (Whyte et al., 2013) and Mediator protein (Loven et al., 2013). However, Hnisz et al. showed that superenhancers across different cell types can be well defined by H3K27ac enrichment, instead of using cell type specific master transcription factors (Hnisz et al., 2013), and most journal papers have followed with this definition since then. Superenhancers are identified by first stitching together enhancers that are less than 12.5 kilobases (kb) apart, ranking these stitched enhancers by their H3K27ac ChIP-seq signal, and finally selecting stitched enhancers to the right of the inflection point on the curve, thereby selecting for enhancers with substantially higher H3K27ac (**Figure 3**).

Step 1. Identification of enhancer locations

Step 2. Clustering of enhancers

Step 3. Identify Super-enhancers

Figure 3 | Workflow for identifying superenhancers

H3K27ac ChIP-Seq peaks are first identified, and peaks within 2.5kb of transcription start sites are filtered out as promoter H3K27ac peaks while the remaining H3K27ac peaks are labelled as enhancer H3K27ac peaks. Enhancer H3K27ac peaks within 12.5kb of each other are stitched together and ranked based on their H3K27ac ChIP-Seq signal. The point of inflection is used to stratify stitched enhancers into superenhancers and typical enhancers. Adapted with permission from Pott and Lieb (2015).

1.7 Enhancer and chromatin interaction alterations in cancer

Since the enhancer-ome and chromatin interact-ome play a major role in regulating gene expression, it is unsurprising that alterations in both can lead to pathological changes that contribute to oncogenesis.

Enhancer sequences are frequently mutated in cancer, at a rate similar to promoter and coding sequences (Weinhold et al., 2014) (**Figure 4**). These mutations can form cancer-specific enhancer and superenhancer signatures (Akhtar-Zaidi et al., 2012; Ooi et al., 2016), highlighting their importance in cancer. Genome wide association studies identified *MYC*-associated enhancers as risk loci for multiple cancer types (Ahmadiyeh et al., 2010), showing that enhancer mutations can contribute to oncogenesis by dysregulating the expression of key oncogenes.

Figure 4 | Frequency of mutations at different genomic elements

Promoters, coding sequences (CDS) and enhancers are mutated at roughly the same rate in cancer. Figure reproduced with permission from Weinhold et al. (2014).

The translocation of entire enhancers to the vicinity of oncogenes can lead to dysregulated oncogene expression. In adenoid cystic carcinoma, the translocation of a *NFIB*-associated superenhancer with MYB transcription factor binding sites to the *MYB* gene locus forms enhancer-promoter chromatin interactions that drive MYB protein expression through a positive feedback loop (Drier et al., 2016) (**Figure 5A**).

Cancer-associated enhancer-promoter chromatin interactions can be formed due to the aberrant expression of chromatin factors to drive oncogene expression. In T-cell acute lymphoblastic leukaemia, overexpression of the c-MAF protein promotes the interaction between the *TAL1* gene on chromosome 1 to an enhancer on chromosome 16, thereby upregulating *TAL1* expression (Patel et al., 2013) (**Figure 5B**).

Additionally, point mutations can create or remove transcription factor binding sites that are required for chromatin interactions. A T-to-C mutation upstream of the *TERT* gene allows the GABPA protein to bind and facilitate chromatin interactions that drive *TERT* expression in multiple cancers (Akıncılar et al., 2016) (**Figure 5C**).

Mutations can also disrupt TADs, resulting in aberrant chromatin interactions. For example, IDH mutations in gliomas result in reduced CTCF insulation at a TAD boundary near the *PDGFRA* gene, allowing chromatin interactions to form between *PDGFRA* and an enhancer from a neighbouring TAD (**Figure 5D**).

Since epigenetic dysregulation events can drive oncogenesis, these events can potentially be targeted for cancer diagnosis or treatment. In the next section, I give examples of how superenhancers and chromatin interactions have been studied as biomarkers or therapeutic targets.

A Enhancer Acquisition / Loss

B Overexpression of TF

C Gain / Loss of TF Binding Site

D Gain / Loss of TAD boundaries

Figure 5 | Enhancer and chromatin interaction alterations contribute to oncogenesis

(A) Translocation of an enhancer to the vicinity of the *MYB* locus aberrantly overexpresses MYB protein in adenoid cystic carcinoma. (B) In T-cell acute lymphoblastic leukaemia, overexpression of c-MAF protein mediates the formation of a new inter-chromosomal chromatin interaction between *TALI* on chromosome 1 and the TIL16 enhancer element on chromosome 16, thereby upregulating *TALI* expression. (C) A single nucleotide polymorphism in the *TERT* gene promoter allows GABPA protein to bind and mediate the formation of long-range chromatin interactions to upregulate *TERT* expression. (D) In gliomas, mutation of the IDH protein hypermethylates CTCF binding sites and prevents CTCF from binding, resulting in the loss of TAD insulation. This allows an enhancer to form long range chromatin interactions with the *PDGFA* gene and upregulate its expression. TF: transcription factor; TAD: topologically associating domain. Figure reproduced with permission from See et al. (2019).

1.8 Superenhancers and chromatin interactions as cancer biomarkers and therapeutic targets

Chromatin interactions and superenhancers can be used to guide diagnosis and treatment of cancers that are driven by epigenetic alterations instead of coding mutations.

Tumours have been shown to elicit epigenetic alterations in the tumour immune microenvironment, which can be assayed for using peripheral blood. Cancer-specific chromatin interactions were identified in peripheral blood cells of melanoma (Bastonini et al., 2014) and prostate cancer (Alshaker et al., 2021) patients that were absent in normal individuals, potentially acting as a diagnostic biomarker for these cancers.

Studies have been conducted on the potential of cancer-associated superenhancers as predictive biomarkers for cancer therapy. McKeown et al. (2017) clustered acute myeloid leukaemia (AML) patients based on their superenhancer profile and identified a novel subset of AML patients with a superenhancer associated with the *RARA* gene. The presence of this superenhancer can predict for increased sensitivity towards *RARA*-targeting compounds. A similar study by Mack et al. (2017) on ependymoma patient samples identified superenhancer-associated genes which can potentially be targeted in different subsets of patients. Two research groups have also profiled

the superenhancer landscape of nasopharyngeal carcinoma (Ke et al., 2017) and lung adenocarcinoma (Yuan et al., 2017a) cell lines to identify superenhancer associated genes with potential as biomarkers.

Superenhancers have emerged as a novel target for cancer therapy in recent years. Loven et al. (Loven et al., 2013) showed that superenhancers are exceptionally enriched for BRD4, a BET (bromodomain and extraterminal domain) protein that mediates transcription activation, and these BRD4-enriched superenhancers are associated with well-known oncogenes in multiple myeloma, small-cell lung cancer and glioblastoma multiforme. Small molecule inhibitors of BET proteins, such as JQ1, can be titrated to preferentially deplete BRD4 at superenhancers associated with oncogenes such as c-Myc (Loven et al., 2013). JQ1 and other BET inhibitor drugs have now undergone various stages of clinical trials as single agents or in combination with other inhibitors (Shorstova et al., 2021).

Superenhancers are also highly sensitive to inhibition of factors involved in the transcriptional machinery, because they require high densities of these factors to drive transcriptional activity (Bradner et al., 2017). Hence, cyclin dependent kinases (CDK) involved in transcription have emerged as potential drug targets to disrupt superenhancers as well. Inhibitors of CDK7, such as THZ1, have been shown to preferentially downregulate genes associated with superenhancers in neuroblastoma (Chipumuro et al., 2014), small cell lung cancer (Christensen et al., 2014), T-cell acute lymphoblastic leukaemia (Kwiatkowski et al., 2014), triple negative breast cancer (Wang et al., 2015), nasopharyngeal carcinoma (Yuan et al., 2017b), oesophageal squamous cell carcinoma (Jiang et al., 2017), and melanoma (Eliades et al., 2018). As with BET inhibitors, CDK inhibitors have entered clinical trials as cancer therapeutics (Sava et al., 2020).

1.9 c-Myc enhancer invasion in cancer

At physiological levels, c-Myc primarily occupies promoters of actively transcribing genes. When c-Myc is overexpressed, c-Myc binding increases at active promoters and additionally occupies active enhancers, resulting in a genome-wide increase in the transcription of active genes (Lin et al., 2012).

Binding of c-Myc to inactive genes is insignificant, and these genes remain silent, suggesting that c-Myc binds only to open, active chromatin (Lin et al., 2012; Nie et al., 2012; Sabo et al., 2014). Even so, these active genes are not upregulated equally. Canonical binding sites become saturated and c-Myc binds to lower affinity binding sites with non-canonical motifs, increasing the expression of genes associated with malignant transformation (Walz et al., 2014). N-Myc, which has high sequence and functional homology with c-Myc (Kohl et al., 1986; Malynn et al., 2000), also invades into enhancers when overexpressed (Zeid et al., 2018).

There is much debate about the transcriptional amplification effects of c-Myc enhancer invasion: whether it is global or selective. Proponents of the global amplification model argue that c-Myc binds non-selectively to all active regulatory elements to amplify all active genes in a general manner, and that cell type specific patterns of gene regulation arise from cell-type specific sets of active genes (Lin et al., 2012; Nie et al., 2012). Advocates of the selective amplification model contend that c-Myc targets a specific set of genes, and that global transcriptional amplification is a secondary response (Kress et al., 2015; Muhar et al., 2018; Sabo et al., 2014). Nevertheless, both models agree that c-Myc overexpression leads to transcriptional amplification by aberrantly invading into enhancers.

1.10 Aims and hypotheses

Overexpression of c-Myc has been shown to amplify active gene expression globally, but it is still unclear why oncogenes are specifically amplified in high *MYC* expressing cancers. Since superenhancers are frequently dysregulated in cancer and are enriched in chromatin interactions, we speculate that c-Myc may preferentially invade into superenhancers and alter the chromatin interaction landscape to aberrantly regulate oncogene expression in cancer.

In this thesis, I investigate the chromatin interaction landscape in c-Myc enhancer invasion, particularly at superenhancers. Specifically, the aims and hypotheses of this study are:

Aims

- 1) Characterize the c-Myc binding and chromatin interaction profiles at enhancers and superenhancers in high c-Myc expressing cancers (Chapter 3)
- 2) Investigate whether overexpressed c-Myc preferentially invades into superenhancers (Chapter 4)
- 3) Investigate whether *MYC* overexpression leads to changes in the global enhancer-ome (Chapter 4)
- 4) Examine whether *MYC* overexpression alters the chromatin interaction landscape (Chapter 5)

Hypothesis

- 1) c-Myc preferentially invades into superenhancers after *MYC* overexpression (**Figure 6A**).
- 2) Enhancer-promoter chromatin interactions facilitate the migration of c-Myc from canonical binding sites at promoters to lower affinity binding sites at superenhancers (**Figure 6B**).
- 3) c-Myc enhancer invasion leads to increased chromatin interaction frequency between c-Myc bound loci (**Figure 6C**).
- 4) Superenhancers gain c-Myc binding and aberrantly regulate gene expression (**Figure 6D**).

Significance

This thesis helps to improve our understanding of c-Myc overexpression in cancer, from the perspective of chromatin interactions and superenhancer invasion. If our hypotheses are true, high c-Myc expressing cancers may be reliant on chromatin interactions to facilitate oncogene expression, which makes these cancers exquisitely susceptible to drugs targeting chromatin interaction factors such as CTCF and cohesin. The disruption of chromatin interactions may then uniquely target the oncogenic consequence of c-Myc overexpression without affecting the important physiological functions of c-Myc.

Figure 6 | Diagrams illustrating the hypotheses of this thesis

(A) c-Myc preferentially invades into superenhancers when overexpressed. (B) Superenhancers interact with gene promoters, allowing overexpressed c-Myc accumulating at gene promoters to spread to lower affinity binding sites at the superenhancers. (C) Chromatin interactions may preferentially form between c-Myc binding sites. (D) Together, c-Myc superenhancer invasion connects superenhancers to gene promoters, resulting in aberrant gene expression. Broken lines indicate large genomic distances. Chromatin interactions are shown as loops. The black boxes represent enhancer H3K27ac peaks. The beige boxes represent c-Myc molecules. The arrows represent transcription start sites. Red wavy lines represent expressed transcripts.

Chapter 2: Materials and Methods

2.1 Cell culture

U-2OS osteosarcoma cells (HTB-96, ATCC) with a doxycycline-inducible c-Myc system were kindly provided by Elmar Wolf's lab from Universität Würzburg, Germany. K562 chronic myelogenous leukaemia cell line (CCL-243, ATCC) were kindly provided by Daniel Tenen's lab from Cancer Science Institute of Singapore (CSI), Singapore. SNU16 gastric adenocarcinoma cells (CRL-5974, ATCC) were kindly provided by Richie Soong's lab from CSI, Singapore. 3T3 mouse fibroblast cells were kindly provided by Polly Chen's lab from CSI, Singapore. Cell lines were authenticated by STR profiling and mycoplasma testing was performed using the MycoAlert™ PLUS Mycoplasma Detection Kit (Lonza).

U-2OS and 3T3 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Gibco), supplemented with 10% tetracycline-free foetal bovine serum (FBS) (Clontech) and 50 U/ml penicillin/streptomycin (Gibco). K562 and SNU16 cells were grown in Roswell Park Memorial Institute (RPMI) 1640 media (Nacalai Tesque), supplemented with 10% FBS (Hyclone) and 50 U/ml penicillin/streptomycin. *MYC* expression was induced in U-2OS cells by the addition of 1µg/ml doxycycline (Clontech) for 30h.

2.2 Reverse transcription and quantitative polymerase chain reaction (qPCR)

RNA was purified from cells using the RNeasy Mini Kit (QIAGEN) and reverse transcribed into cDNA using the qScript cDNA Supermix (Quantabio), both according to the manufacturers' protocol. qPCR was performed with primers (IDT) listed in **Table 1**, using the GoTaq qPCR Master Mix (Promega) on the Quantstudio 3 or 5 systems (Applied Biosystems). Genes were quantified relative to *TBP* or *RPLP0* using the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen, 2001).

Table 1 | qPCR Primer sequences

Primer Name	Sequence
<i>TBP</i> _qPCR_F	TAATCCCAAGCGATTTGCTG
<i>TBP</i> _qPCR_R	CAGTTGTCCGTGGCTCTCTT
<i>RPLP0</i> _qPCR_F	TTCATTGTGGGAGCAGAC
<i>RPLP0</i> _qPCR_R	CAGCAGTTTCTCCAGAGC
<i>MYC</i> _qPCR_F	CACCAGCAGCGACTCTGA
<i>MYC</i> _qPCR_R	GATCCAGACTCTGACCTTTTGC

2.3 RNA-Seq

RNA was purified from cells using the RNeasy Mini Kit (QIAGEN). Ribosomal RNA depletion and library preparation was performed using the TruSeq Stranded Total RNA LT Kit (Illumina) according to the manufacturer's protocol. Paired-end 2 x 100bp sequencing was performed on the Illumina HiSeq 2500.

2.4 Protein extraction and immunoblotting

Cells were lysed in RIPA Lysis and Extraction buffer (Thermo Scientific) supplemented with Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific) and homogenized by pipetting. Protein concentration was measured using the Pierce™ BCA Protein Assay Kit (Thermo Scientific).

Equal amounts of protein were diluted in Laemmli sample buffer (Bio-Rad) supplemented with β -mercaptoethanol (Bio-Rad) and resolved on a 12% SDS-PAGE gel at 180V. Proteins were transferred from the gel onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad). The membrane was blocked in Tris buffered saline (TBS) containing 0.05% Tween-20 and 5% non-fat milk (TBS-T) for 1 hour, followed by overnight incubation with the primary antibody and 1 hour incubation with the secondary antibody, washing thrice with TBS-T between steps. Antibody-bound proteins were visualized using the Clarity™ Western ECL Substrate (Bio-Rad) on the ImageQuant LAS 500 (GE Healthcare). Antibodies used for immunoblotting are listed in **Table 2**.

Table 2 | Protein immunoblotting antibodies

Target	Manufacturer	Cat No.	Lot No.
c-Myc	abcam	ab32072	GR295111-33
Vinculin	Sigma-Aldrich	V9131	036M4797V

2.5 Chromatin Immunoprecipitation followed by sequencing (ChIP-seq)

20 million cells were fixed with 1% formaldehyde and sonicated using the truChIP Chromatin Shearing Kit (Covaris) on the ME220 Focused-Ultrasonicator (Covaris) according to the manufacturer's protocol.

H3K27ac and c-Myc bound DNA were immunoprecipitated using antibodies listed in **Table 3**. 3.5µg of each antibody was rotated with 15µl of Dynabeads Protein G magnetic beads (Invitrogen) at 4°C overnight and washed thrice with Beads Wash Buffer (0.1% Triton X-100 in PBS) to remove unbound antibodies. Antibody-bound beads were combined with sonicated chromatin from 5 million cells and rotated at 4°C overnight. Beads were then washed thrice with Shearing Buffer D3 (Covaris), once with High Salt Washing Buffer (50mM HEPES pH 7.5, 350mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS), once with Lithium Chloride Wash Buffer (10mM Tris pH 8.0, 250mM LiCl, 1mM EDTA, 0.5% NP-40, 0.5% Sodium Deoxycholate) and once with Tris-EDTA buffer.

Chromatin was eluted from the magnetic beads in 100µl elution buffer (50mM Tris pH 8.0, 10mM EDTA, 1% SDS) supplemented with 2µl of 0.5mg/ml RNase A (QIAGEN) at 55°C for 2 hours. Chromatin was then decrosslinked with the addition of 2µl of 20mg/ml proteinase K (Ambion) at 55°C for 4 hours and 37°C overnight. For total input, chromatin from 500,000 cells was treated with RNase A and decrosslinked similarly. Immunoprecipitated chromatin was purified using the MinElute PCR Purification Kit (QIAGEN) according to manufacturer's protocol.

ChIP-seq libraries were prepared using the ThruPLEX DNA-seq Kit (Rubicon) and sequenced paired-end 2 x 100bp on the HiSeq 2500 (Illumina).

Table 3 | ChIP antibodies

Target	Manufacturer	Cat No.	Lot No.
H3K27ac	abcam	ab4729	GR150367-1
c-Myc	Santa Cruz	sc-764	H1712
IgG	Santa Cruz	sc-2027	H2615

2.6 Circular chromosome conformation capture (4C)

40 million cells were cross-linked in 30ml of phosphate buffered saline (PBS) supplemented with 1% formaldehyde at room temperature for 10 minutes, followed by quenching with 0.25M glycine for 5 minutes. Cells were washed thrice with PBS supplemented with 1mM phenylmethylsulfonyl fluoride (Sigma Aldrich). Nuclei were isolated by lysing cells in 4C Lysis Buffer (10mM Tris-HCl pH 8.0, 10mM NaCl, 5mM EDTA, 0.5% Igepal CA-630) at 4°C for 10 minutes, followed by homogenization using a dounce homogenizer (Wheaton) (50 strokes using Pestle B).

Nuclei were permeabilized at 37°C with the addition of 0.3% SDS for 1 hour, followed by 2% Triton-X 100 for 1 hour. Primary enzyme digestion with HindIII-HF (NEB) was done at 37°C for 18h, before proximity ligation in dilute conditions with T4 DNA Ligase (Thermo Scientific) at 16°C overnight. Crosslinking was reversed with the addition of 55µg/ml Proteinase K (Ambion) at 65°C for 4 hours. DNA was purified using phenol/chloroform extraction followed by ethanol precipitation to yield the 3C library.

Secondary enzyme digestion with DpnII (NEB) was done on the 3C library at 37°C overnight, followed by proximity ligation and de-crosslinking as described in the previous paragraph.

4C libraries were generated for each viewpoint through nested inverse-PCR using Phusion DNA Polymerase (Thermo Scientific), with primers listed in **Table 4**. DNA fragments between 200 and 1000bp were isolated from the 4C libraries by gel excision after running the 4C libraries on a 4-20% gradient TBE gel (Thermo Scientific). The gel slices were shredded and macerated in Tris EDTA buffer at 37°C overnight and DNA was precipitated using ethanol. The 4C libraries were

multiplexed in equimolar ratio prior to single-end 1 x 150bp sequencing on the Illumina MiSeq sequencer.

Table 4 | 4C viewpoints and primer sequences

Viewpoint	Genomic coordinates	Primer	Sequence
U2OS_chr2_ PROC_SE	chr2:128165324 -128165967	Outer F	TTAGCCACGAAGTGTCC TGC
		Outer R	AGGAGAACTGTTTTTCA TTGCCATG
		Nested F	GGTGGAGTTTCCTTCCCT GT
		Nested R	TGATTGGATGCAGCACT TGG
K562_chr3_ TFRC_SE	chr3:195834970 -195835703	Outer F	CACTGCTTACAAATGCG AAT
		Outer R	GCATGCTGGAGATAAGA GAA
		Nested F	CCCTCCACTATCTCCCTT C
		Nested R	ACGTGGCAATGTGTCTA AG
K562_chr6_ TNFAIP3_E	chr6:138276376 -138276971	Outer F	GGAAGAAAATGAGGAG CAGA
		Outer R	TTTCACTTGGTACTTGGT GT
		Nested F	GAAAGCGCAGGTTCTTA GA
		Nested R	CTTGGTGTGGTGATAGTT GA
K562_chr15_ MFGE8_E	chr15:89528798 -89529662	Outer F	TAAGAAACAACCCCTGA ACC
		Outer R	ATTTCCACCAAGAACCT TTG
		Nested F	TCTAACAGAAAAGGTGG CTC
		Nested R	CAGGAGTTTGAAGCTAC AGT
SNU16_chr8_ _YWHAZ_S E	chr8:102,021,412-102,023,300	Outer F	GCCATTCTTGGTGCATA GCC
		Outer R	CCTCCTGGGTTCATGCCA TT
		Nested F	ACAAGACTCCAAGGTGC CTC
		Nested R	ACTACAGGTGCCTACCA CCA

SNU16_chr3 _OPA1_SE	chr3:193600321 -193603043	Outer F	TTGTTTGTATGAGAGCC AGGGAG
		Outer R	ATTTAAGCTCCCTACTGC TGGG
		Nested F	TTTAAGAGGACAGGAAT GAGGGC
		Nested R	ACTGAATATAGAGCCCC AGACCA

2.7 Spike-in quantitative Hi-C (SIQHiC)

Cell counts for experiment cell samples and control mouse 3T3 cells were averaged from three readings using the Countess II automated cell counter (Thermo Scientific). 10 million cells were fixed with 2% formaldehyde using the Arima Hi-C Kit (Arima Genomics) following manufacturer's instructions. 1 million fixed control mouse 3T3 cells were added to 4 million fixed experiment cells before subsequent steps of restriction enzyme digest, biotin end filling and ligation using the Arima-HiC kit according to the manufacturer's protocol. Libraries were prepared using the KAPA Hyper-Prep Kit (KAPA), according to the Arima-HiC kit protocol. SIQHiC libraries were sequenced paired-end 2 x 150bp on the HiSeq 4000 (Illumina).

2.8 RNA-Seq data processing

RNA-seq libraries were mapped to the hg19 genome and read counts for UCSC RefSeq transcripts were obtained using kallisto v0.44.0 (Bray et al., 2016). Differentially expressed transcripts were identified using sleuth v0.30.0 (Pimentel et al., 2017) (fdr<0.01, |beta|>1).

2.9 ChIP-seq data processing

H3K27ac and c-Myc ChIP-seq datasets for K562, SNU16 and LNCaP cell lines were obtained from the NCBI Gene Expression Omnibus (GEO) database as listed in **Table 5**.

ChIP-seq libraries were mapped to the hg19 genome using BWA-MEM (Li, 2013) version 0.7.5a-r405 using default settings. PCR duplicates were removed

using samtools (Li et al., 2009) version 1.7 and reads falling within the ENCODE consensus blacklisted regions (Encode Project Consortium, 2012) were removed using BEDTools (Quinlan and Hall, 2010) version 2.26.0. ChIP-seq signals were visualized using deepTools (Ramírez et al., 2014) version 3.2.1 (bamCompare -normalizeUsing RPKM --operation subtract -bs 1). Bam files of biological replicates were merged before calling ChIP-seq peaks using MACS2 (Zhang et al., 2008) version 2.1.2. H3K27ac ChIP-seq peaks within 2kb of transcription start sites were labelled as active promoter peaks while the rest were labelled as putative enhancer peaks.

Superenhancers were identified by stitching together H3K27ac enhancer peaks within 12.5kb of each other and ranked based on H3K27ac enrichment, using an in-house script as described previously (Cao et al., 2017). When the diagonal line connecting the first and last value slides along the H3K27ac enrichment curve, the point where this line is tangential to the curve was chosen as a cut-off to separate superenhancers from typical enhancers (**Figure 3**).

To identify differential H3K27ac peaks, bam files from all replicates and conditions were merged and a common list of H3K27ac peaks was called using MACS2. Read counts for the common list of H3K27ac peaks were obtained using Rsubread (Liao et al., 2019). Differential H3K27ac peaks were identified using DESeq2 (Love et al., 2014) version 1.24.0 (padj<0.01 and padj<0.1 respectively).

Table 5 | ChIP-seq datasets downloaded from NCBI GEO

ChIP-seq dataset	Library Accession	Reference
K562 H3K27ac	GSE29611	Encode Project Consortium (2012)
K562 c-Myc	GSE31477	
SNU16 H3K27ac	GSE76153	Ooi et al. (2016)
LNCaP / NTC H3K27ac	GSE135942	Whitlock et al. (2020)
LNCaP / MYC ⁶²⁷ H3K27ac		
LNCaP / NTC c-Myc		
LNCaP / MYC ⁶²⁷ c-Myc		

2.10 4C data processing

4C libraries were analysed with the help of Dr. Omer An (Bioinformatics Core Facility, Cancer Science Institute of Singapore, Singapore), using the CSI NGS portal (An et al., 2020) (<https://csibioinfo.nus.edu.sg/>).

4C libraries were mapped to the hg19 genome using BWA-MEM (Li, 2013) version 0.7.5a-r405 with default settings. Significant 4C interactions were identified using the r3Cseq pipeline ($q < 0.05$) (Thongjuea et al., 2013).

2.11 SIQHiC data processing

SIQHiC libraries were analysed with the help of Ms. Kaijing Chen (Melissa Fullwood lab, Nanyang Technological University, Singapore).

SIQHiC libraries were analysed using the Juicer (v1.5) pipeline (Durand et al., 2016) with some modifications to obtain hic matrix files. Since SIQHiC libraries include human and mouse DNA, paired reads were separated and mapped as single reads using BWA-MEM (Li, 2013) to an artificial reference genome combining hg19 human and mm10 mouse genome sequences. Human and mouse chromosomes were appended with “H” and “M” respectively, to avoid chromosome name duplication. After mapping, reads were paired together again and split into 2 files. Paired reads both mapping to human chromosomes were placed together as Human-Human paired reads (H-H), while paired reads both mapping to mouse chromosomes were placed together as Mouse-Mouse paired reads (M-M). Ambiguous read pairs that separately mapped to different species were discarded. H-H and M-M reads were processed using Juicer separately, filtering out duplicates, intra-fragment reads and reads with MAPQ < 30 to generate .hic matrix files for each species. H-H reads of biological replicates were also merged and processed in the same way to generate merged .hic matrix files for downstream analyses.

The ratio between human and mouse contacts (HMR) for each sample was calculated. Since the mouse 3T3 cells in each sample come from the same

population and were spiked into the human cells at the same ratio of 1:4, we expect to obtain a similar HMR ratio across all samples. Hence, the relative difference between the HMR of different samples (SIQHiC ratio) reflects the changes in global chromatin contact frequency. The SIQHiC ratio was calculated such that samples with lower HMR are normalized against samples with the highest HMR. Hi-C contact matrices from merged biological replicates were normalized using the SIQHiC ratio by adding a custom normalization vector to the original contact matrix .hic file using the Juicer ‘addNorm’ subroutine, where the magnitude of each bin was the SIQHiC ratio raised to the power of (-0.5). In this way, the Hi-C matrices are cell-count normalized to the sample with the highest HMR to prevent scale-up extrapolation errors.

The SIQHiC normalized matrix for each chromosome was extracted from the .hic file using the Juicer ‘dump’ subroutine, combined together, and finally reassembled into a SIQHiC normalized .hic matrix file using the Juicer ‘pre’ subroutine. The sample with the highest HMR has a SIQHiC normalization vector of 1, thus its SIQHiC normalized matrix is equivalent to the non-normalized matrix. Both the original and SIQHiC normalized .hic matrices were balanced using the Knight-Ruiz algorithm for subsequent analyses.

Aggregate peak analyses (APA) was performed using the Juicer ‘apa’ subroutine using both non-normalized and SIQHiC normalized contact matrices using the parameters “-k KR -u -n 30”. APA heatmaps were generated using pheatmap v1.0.12 in R.

Compartments were identified using the Juicer ‘eigenvector’ and ‘pearsons’ subroutines at 100kb resolution. Since eigenvalues signs are arbitrary, we correlated eigenvectors with the number of gene promoters within each 100kb bin, such that gene rich regions were assigned positive eigenvalues. Topologically associating domains (TADs) were identified using the Juicer Arrowhead subroutine with default parameters at 50kb and 100kb resolutions.

Loops were called using the Juicer ‘hiccups’ subroutine with parameter “-k KR -m 1024 --ignore_sparsity”. Lost, common and gained loops were identified

using pgltools (Greenwald et al., 2017) by merging and intersecting loops from different samples using a 10kb intersecting window at each loop anchor. Hi-C loops were annotated according to their overlap with c-Myc binding sites, promoters, typical enhancers and superenhancers using pgltools.

Virtual 4C tracks were extracted from the Hi-C matrices by using the Juicer ‘dump’ subroutine, dumping a 5kb window against its entire chromosome using a 5kb bin size.

Chapter 3: Characterizing c-Myc binding and chromatin interaction profiles in high c-Myc expressing cancer cell lines

3.1 c-Myc occupies superenhancers in high c-Myc cancer cell lines

Lin et al. (2012) had shown that overexpressed c-Myc invades into enhancers in the c-Myc inducible P493-6 Burkitt lymphoma cell line, and observed c-Myc occupancy at active enhancers in three high c-Myc cancer cell lines of different tumour types. However, it is unclear whether c-Myc preferentially occupies superenhancers and associates with oncogenes via chromatin interactions in high c-Myc expressing cancers. In this chapter, I characterised the c-Myc binding and chromatin interaction profiles at superenhancers in 3 high c-Myc expressing cell lines: K562 chronic myelogenous leukaemia (CML), SNU16 gastric carcinoma and LNCaP prostate adenocarcinoma.

H3K27ac and c-Myc ChIP-seq data for SNU16 and K562 cells were obtained from the NCBI GEO database as listed in **Table 5**. I identified 14,129 typical enhancers and 254 superenhancers in SNU16 cells, of which 32.4% of typical enhancers and 98.4% of superenhancers were occupied by c-Myc (**Figure 7A**). Similarly, in K562 cells, 19.9% of the 15,822 typical enhancers and 78.5% of the 944 superenhancers identified were bound by c-Myc (**Figure 7A**). In comparison, 3.4% of typical enhancers and 22.5% of superenhancers were bound by c-Myc in the low c-Myc expressing U2OS osteosarcoma cell line, markedly lower than in SNU16 and K562 cells (**Figure 7A**). Superenhancers and typical enhancers were similarly enriched for canonical CACGTG and non-canonical CANNTG motifs in both K562 and SNU16 cells (**Figure 7B-C**), indicating that the differences in c-Myc enrichment are not due to differences in motif enrichment. These results confirm that c-Myc binding at enhancers correlates with c-Myc expression, and additionally reveal that superenhancers are particularly enriched in c-Myc binding.

Figure 7 | c-Myc occupies superenhancers particularly in high c-Myc expressing cancer cell lines

(A) Percentage of c-Myc bound (shaded) typical enhancers and superenhancers in the high *MYC* expressing SNU16 and K562, and the low *MYC* expressing U2OS cells. Total number of typical and superenhancers are noted above each bar. (B-C) Percentage of canonical CACGTG and non-canonical CANNTG motifs within typical enhancers (TE) and superenhancers (SE) in (B) K562 and (C) SNU16 cells (N.S.: not significant, chi-square test).

3.2 c-Myc bound superenhancers interact with cancer-associated genes and nearby c-Myc binding sites

Next, I performed 4C-seq at selected c-Myc bound typical enhancers and superenhancers in SNU16 and K562 to characterize their chromatin interaction profiles (**Figure 8**). c-Myc bound enhancers generally interact with the nearest gene promoters, but also skip over nearby genes to interact with more distant ones (**Figure 8D and F**), consistent with previous observations of gene skipping in K562 cells (Cao et al., 2017). These enhancers can also exhibit interaction directionality bias, favouring interactions with either upstream (**Figure 8B and E**) or downstream loci (**Figure 8C and D**). These results highlight the importance of analysing c-Myc enhancer invasion using chromatin interaction

studies, instead of association by proximity.

SNU16 and K562 c-Myc bound typical enhancers and superenhancers interact with cancer-associated genes, such as *PABPC1* (Zhu et al., 2015) (**Figure 8A**), *HES1* (Yao et al., 2017) (**Figure 8C**), *TNFAIP3* (Schmitz et al., 2009; Troppan et al., 2015) (**Figure 8E**), *TFRC* (Liu et al., 2014) and *RNF168* (Chroma et al., 2016) (**Figure 8F**). Chromatin interactions also connect c-Myc bound enhancers with other c-Myc bound loci in the vicinity (**Figure 8A-F**), bringing these loci together in close spatial proximity. From another perspective, c-Myc is concentrating at regions where enhancers and promoters congregate – areas of high transcriptional activity.

Figure 8 | c-Myc bound superenhancers interact with cancer associated genes and other c-Myc binding sites in high c-Myc expressing cancers

Significant 4C chromatin interactions ($q < 0.05$) at selected c-Myc bound superenhancers and typical enhancers in (A-C) SNU16 and (D-F) K562 cells. 4C viewpoint loci are highlighted in blue. Tracks shown from top to bottom are: (1) 4C chromatin interaction signal, (2) significant chromatin interactions ($q < 0.05$), (3) typical enhancers (4) superenhancers, (5) c-Myc ChIP-seq peaks and (6) UCSC genes (Hsu et al., 2006). Tracks were visualized using the UCSC Genome Browser (<http://genome.ucsc.edu>) (Kent et al., 2002).

3.3 MYC shRNA knockdown reduced c-Myc occupancy at enhancers

Although I have shown that c-Myc is highly enriched at superenhancers in high c-Myc expressing cancers, it is not clear whether these superenhancers were already occupied by c-Myc prior to increased *MYC* expression. To that end, I reanalysed publicly available H3K27ac and c-Myc ChIP-seq data of LNCaP cells expressing short hairpin RNA (shRNA) targeting *MYC* (*MYC*⁶²⁷) or non-targeting control (NTC) (Whitlock et al., 2020) (**Table 5**). shRNA knockdown of *MYC* reduced *MYC* gene expression by 30% compared to NTC (**Figure 9A**).

As with SNU16 and K562 cells, one third of typical enhancers and almost all superenhancers were occupied by c-Myc in LNCaP cells (**Figure 9B**). In total, 17,972 c-Myc binding sites were lost and 8,986 c-Myc binding sites were gained after *MYC* knockdown (**Figure 9C**). Majority of lost binding sites were at stitched enhancers, whereas promoter c-Myc binding sites remained mostly stable (**Figure 9C**), indicating that c-Myc binding sites were preferentially lost at enhancers when c-Myc concentration decreased. *MYC* knockdown significantly reduced the proportion of c-Myc bound typical enhancers and superenhancers (5.4% and 4.1% respectively) (**Figure 9B**), with 16% of stitched

enhancers losing at least 1 c-Myc binding site (**Figure 9D**). Common c-Myc binding sites were more enriched for canonical CACGTG motifs and less enriched for non-canonical CANNTG motifs compared to lost and gained c-Myc binding sites (**Figure 9E**), suggesting that c-Myc binds more stably to the canonical motif. However, the CANNTG motif occurs in around 65% of c-Myc binding sites while the CACGTG motif occurs in less than 10% (**Figure 9E**), indicating that c-Myc occupies lower-affinity binding sites at high c-Myc concentrations. Typical enhancer and superenhancers were more enriched for CANNTG motifs compared to promoters, while all three CREs were similarly enriched for CACGTG motifs (**Figure 9F-G**), indicating that c-Myc binds predominantly at lower-affinity binding sites at enhancers. These results show that c-Myc binding at enhancers is dependent on c-Myc concentration.

C Changes in c-Myc binding sites at CREs after *MYC* knockdown

D Loss/Gain in c-Myc peaks at NTC stitched enhancers after *MYC* knockdown

E LNCaP CREs after *MYC* shRNA knockdown

F LNCaP_Control CREs

G LNCaP_MYCKD627 CREs

Figure 9 | MYC shRNA knockdown reduced c-Myc occupancy at typical enhancers and superenhancers

(A) RT-qPCR of *MYC* expression relative to *GAPDH* in LNCaP cells expressing non-targeting control (NTC) or *MYC* targeting (*MYC*⁶²⁷) shRNA, using primers across exons 1 and 2 (left panel) and across exons 2 and 3 (right panel). Bars and whiskers show mean and standard deviation of 10 independent replicates (****: $P < 0.0001$, Student's *t* test). Figure adapted with permission from Whitlock et al. (2020). (B) Percentage of c-Myc bound typical enhancers and superenhancers in LNCaP/NTC and LNCaP/*MYC*⁶²⁷ cells. Total number of typical enhancers and superenhancers are noted above each bar. (****: $P < 0.0001$, chi square test) (C) Gain and loss of c-Myc binding sites at stitched enhancers and promoters after *MYC* shRNA knockdown. (D) Frequency distribution of loss or gain in c-Myc peaks within stitched enhancers after *MYC* shRNA knockdown. (E) Percentage of canonical CACGTG and non-canonical CANNTG motifs at lost, common and gained c-Myc binding sites (*: $P < 0.05$, ***: $P < 0.001$, ****: $P < 0.0001$, chi square test). (F-G) Percentage of canonical CACGTG and non-canonical CANNTG motifs at promoter (Pro), typical enhancer (TE) and superenhancer (SE) c-Myc binding sites in (F) LNCaP Control cells and (G) LNCaP MYCKD627 cells (****: $P < 0.0001$, N.S.: not significant, chi square test).

3.4 MYC shRNA knockdown decreased constituent H3K27ac peaks at superenhancers and typical enhancers

MYC overexpression has been shown to modulate histone acetylation and methylation. *MYC* overexpression in P493-6 cells hyper-acetylates various lysine residues on histones H3 and H4 at a panel of c-Myc target promoters, including H3K9, H3K14, H3K18, H4K5 and H4K12, but did not significantly modulate H3K27 acetylation (Martinato et al., 2008). *MYC* knockout in naïve B cells also did not significantly modulate global H3K27 acetylation levels (Kieffer-Kwon et al., 2017). However, it is unclear whether changes in c-Myc levels elicits changes in H3K27 acetylation patterns at enhancers and superenhancers.

MYC shRNA knockdown in LNCaP cells marginally reduced the number of stitched enhancers from 46,565 to 43,208, with a loss of 6,162 stitched enhancers and a gain of 2,774 stitched enhancers (**Figure 10A**). Although 276 superenhancers were unique to LNCaP / NTC cells (**Figure 10C**), all LNCaP / NTC superenhancers overlap with 1 or more LNCaP / *MYC*⁶²⁷ stitched enhancer (**Figure 10D**), showing that these superenhancers remained acetylated but did not meet the H3K27ac signal cut-off after *MYC* knockdown (see **Figure 3**). *MYC* shRNA knockdown resulted in a general decrease of constituent H3K27ac peaks in superenhancers and typical enhancers, with 45% of superenhancers losing

between 1 and 24 constituent H3K27ac peaks and 32% of typical enhancers losing between 1 and 6 constituent H3K27ac peaks (**Figure 10E-F**). Taken together, these results suggest that c-Myc may modulate H3K27 acetylation at superenhancers and typical enhancers.

Figure 10 | MYC shRNA knockdown reduced the number of H3K27ac peaks within superenhancers

(A-C) Overlap of (A) all stitched enhancers, (B) typical enhancers and (C) superenhancers between LNCaP / NTC and LNCaP / MYC⁶²⁷ cells. (D) Number of stitched enhancers that overlap at each LNCaP / NTC superenhancers after MYC knockdown. (E-F) Gain and loss of H3K27ac peaks at LNCaP / NTC (E) superenhancers and (F) typical enhancers after *MYC* shRNA knockdown.

3.5 Summary

In support of my first hypothesis, these results show that c-Myc binds at typical enhancers and particularly occupies majority of superenhancers in 3 high c-Myc expressing cancer cell lines, and these typical enhancers and superenhancers connect to multiple genes and other c-Myc bound loci via chromatin interactions (Figure 11). Superenhancers remained acetylated after *MYC* shRNA knockdown, but c-Myc binding sites and H3K27ac constituent peaks within superenhancers were lost, suggesting that these superenhancers were present but not occupied by c-Myc at physiological conditions (Figure 11). However, it is unclear whether these observations are immediate effects of *MYC* overexpression or acquired after long term adaptation to high *MYC* conditions. Therefore, in the next chapter, I explore the immediate effects of *MYC* overexpression in a cell line with low endogenous *MYC* expression.

Figure 11 | Superenhancers are bound by c-Myc in high c-Myc expressing cancers and lose c-Myc binding after MYC shRNA knockdown.

Superenhancers are bound by c-Myc in high c-Myc expressing cancers and connect to multiple genes via long-range chromatin interactions. Superenhancers lose c-Myc binding after *MYC* shRNA knockdown. Chromatin interactions are shown as loops. The black boxes represent enhancer H3K27ac peaks. The beige boxes represent c-Myc molecules.

Chapter 4: *MYC* overexpression leads to superenhancer invasion

4.1 Doxycycline induction of *MYC* expression in U2OS cells regulated genes involved in cancer progression

In the previous chapter, I have shown that c-Myc occupies one third of typical enhancers and almost all superenhancers in high c-Myc expressing cancer cell lines. Superenhancers and typical enhancers lose c-Myc binding and H3K27ac constituent peaks when *MYC* expression is knocked down, while promoters mostly retain c-Myc binding. However, these cancers have accumulated adaptations to survive and thrive with high c-Myc levels, making it unclear whether these observations are direct effects of c-Myc overexpression.

Therefore, to investigate the immediate changes in c-Myc enhancer binding associated with *MYC* overexpression, I used the U2OS osteosarcoma cell line inserted with a tetracycline-inducible *MYC* cassette (Pajic et al., 2000) as a model. Unlike the LNCaP *MYC* shRNA knockdown model, U2OS cells are not reliant on high *MYC* expression, have relatively low endogenous expression of *MYC* protein comparable to non-transformed cells (Lorenzin et al., 2016) (**Figure 12A**), and have been used as a model in previous studies on *MYC* overexpression (Lorenzin et al., 2016; Walz et al., 2014).

Treating U2OS cells with 1µg/ml doxycycline (DOX) for 30h increased *MYC* expression by around 35 times as shown by qPCR and RNA-seq (**Figure 12C and D**), accompanied with a concomitant increase in c-Myc protein expression (**Figure 12B**). In this work, I will refer to doxycycline treated and untreated U2OS cells as High *MYC* and Low *MYC* cells respectively.

Figure 12 | Doxycycline treatment induces *MYC* overexpression in U2OS cells

(A) Western blot showing low endogenous expression of c-Myc in U2OS cells (boxed in red), at a level comparable to other non-transformed cell lines (IMEC, HMLE and MCF10A) and lower than transformed cell lines (Hela and HCT116). Figure adapted with permission from Lorenzin et al. (2016) (B) Western blot of c-Myc in U2OS cells before and after 30h of doxycycline induction. Vinculin was used as a loading control. (C) RT-qPCR of *MYC* expression in U2OS cells with and without doxycycline (DOX) induction at 0h and 30h time points. *MYC* expression was calculated relative to *TBP* and normalized against the 0h DOX-condition. Error bar shows standard error across 3 biological replicates. (D) RNA-seq normalized transcripts per million (TPM) expression of *MYC* in U2OS low *MYC* and high *MYC* cells. Error bar shows standard error across 2 biological replicates.

RNA-seq of Low *MYC* and High *MYC* cells identified 1,390 significant differentially expressed transcripts (DETs) ($|\text{beta}| > 1$, $q < 0.05$, Wald test), of which 548 transcripts were upregulated and 842 transcripts were down regulated after *MYC* overexpression (**Figure 13A**). Endogenous c-Myc bound 24.5% of upregulated DETs, further increasing to 36.9% after overexpression (**Figure 13B**). In contrast, 7.1% of downregulated DETs were bound by c-Myc, marginally increasing to 12.7% after overexpression (**Figure 13B**). Gene set enrichment analysis (GSEA) of DETs showed significant enrichment of previously published c-Myc regulated gene sets (**Figure 13C**). *MYC* overexpression activated pathways related to cancer progression, including ribosome biogenesis, translation, mitochondrial biogenesis and splicing (**Table 6**). Upregulation of genes in the p38 mitogen activated protein kinase (MAPK) pathway, which is associated with cell motility in osteoblasts (Rodríguez-Carballo et al., 2016), and downregulation of cell adhesion related gene sets suggest a shift towards a cell migration phenotype (**Table 6** and **Table 7**).

Figure 13 | MYC overexpression increased c-Myc binding at differentially expressed transcripts

(A) Significantly upregulated (red) and downregulated (blue) differentially expressed transcripts (DETs) after MYC overexpression ($|\text{beta}| > 1$, $\text{FDR} < 0.05$, Wald test). Top 10 significant DETs are labelled. (B) Upregulated and downregulated DETs bound by c-Myc in Low MYC and High MYC cells. (C) Enrichment plots of 2 previously published c-Myc regulated gene sets generated using GSEA (Subramanian et al., 2005).

Table 6 | Upregulated gene sets after *MYC* overexpression

Up-regulated gene sets	NES	p-value	q-value
General c-Myc up-regulated			
KIM MYC AMPLIFICATION TARGETS UP	1.80	0.000	0.000
SCHUHMACHER MYC TARGETS UP	1.77	0.000	0.001
Wnt signalling			
WP WNT SIGNALING PATHWAY NETPATH	1.69	0.000	0.037
Ribosome biogenesis			
REACTOME RRNA PROCESSING	1.86	0.000	0.000
REACTOME RRNA MODIFICATION IN THE NUCLEUS AND CYTOSOL	1.69	0.000	0.033
Translation			
REACTOME TRANSLATION	1.77	0.000	0.001
Mitochondrial biogenesis			
REACTOME MITOCHONDRIAL TRANSLATION	1.67	0.000	0.044
Splicing			
KEGG SPLICEOSOME	1.80	0.000	0.057
p38 MAPK pathway			
BIOCARTA P38MAPK PATHWAY	1.67	0.002	0.040
WP P38 MAPK SIGNALING PATHWAY	1.66	0.000	0.048
Embryonic stem cell related			
BENPORATH ES 1	1.77	0.000	0.001
MUELLER PLURINET	1.66	0.000	0.046

Table 7 | Downregulated gene sets after *MYC* overexpression

Down-regulated gene sets	NES	p-value	q-value
Cell Adhesion			
NABA CORE MATRISOME	-1.97	0.000	0.000
KEGG FOCAL ADHESION	-1.89	0.000	0.007
WP FOCAL ADHESION	-1.87	0.000	0.016
NABA ECM GLYCOPROTEINS	-1.84	0.000	0.028

4.2 *MYC* overexpression increased constituent H3K27ac peaks within existing superenhancers

I first identified typical enhancers and superenhancers in Low *MYC* and High *MYC* cells by performing H3K27ac ChIP-seq. *MYC* overexpression modulated acetylation at H3K27, significantly increasing acetylation at 168 promoter and 72 enhancer peaks, and decreasing acetylation at 4 promoter and 48 enhancer peaks ($|\log_2\text{foldchange}| > 1$, $\text{FDR} < 0.01$, Wald Test) (**Figure 14A**). Enhancer H3K27ac peaks within 12.5kb of each other were stitched together and ranked by H3K27ac signal, and superenhancers were identified using the geometric point of inflection as a cut-off (**Figure 14B-E**). Around 28,500 stitched enhancers were identified in Low *MYC* and High *MYC* cells, of which 75% were common (**Figure 14C**).

Since cancer cells frequently acquire oncogenic superenhancers to drive oncogene expression, I wanted to know whether *MYC* overexpression by itself is sufficient to create *de novo* oncogenic superenhancers. 75% of the superenhancers were common between Low *MYC* and High *MYC* cells, with 164 superenhancers gained after *MYC* overexpression (**Figure 14E**). However, almost all High *MYC* superenhancers overlapped with at least 1 Low *MYC* stitched enhancer, demonstrating that the gained superenhancers were pre-existing enhancers and not previously silent (**Figure 14F**). On the other hand, 62% of the High *MYC* superenhancers gained constituent H3K27ac peaks after *MYC* overexpression (**Figure 14G**). This is consistent with my previous observation that LNCaP superenhancers remained active but lose constituent H3K27ac peaks after *MYC* shRNA knockdown (**Figure 10C-F**). Taken together, these results indicate that *MYC* overexpression modulated H3K27 acetylation at pre-existing enhancer clusters, but did not activate novel superenhancers at previously silent loci.

Figure 14 | *MYC* overexpression modulated H3K27 acetylation and increased constituent H3K27ac peaks at superenhancers

(A) Promoter and enhancer H3K27ac peaks with significant changes in acetylation signal ($|\log_2\text{foldchange}| > 1$, $\text{FDR} < 0.01$, Wald Test). (B) Rank ordering of stitched enhancers to identify superenhancers in Low *MYC* (top panel) and High *MYC* cells (bottom panel). (C-E) Overlap of all stitched enhancers (C), typical enhancers (D) and superenhancers (E) between Low *MYC* and High *MYC* cells. (F) Number of Low *MYC* stitched enhancers that overlap at High *MYC* superenhancers after *MYC* overexpression. (G) Gain and loss of H3K27ac peaks at High *MYC* superenhancers after *MYC* overexpression.

4.3 Superenhancers are preferentially enriched for c-Myc binding after *MYC* overexpression

Next, I overlapped c-Myc ChIP-seq peaks with CREs to examine c-Myc enhancer invasion. *MYC* overexpression increased c-Myc binding sites from 7,458 to 11,911, with a loss of 1,797 and gain of 6,349 binding sites (**Figure 15A-B**). The percentage of c-Myc binding sites at typical enhancers increased from 14.0% to 17.2%, and superenhancer binding sites increased from 3.6% to 4.2%, whereas promoter binding sites remained similar (61.5% to 61.8%) (**Figure 15A**), recapitulating previous observations of enhancer invasion. Endogenous c-Myc promoter binding sites were mostly stable, losing 6.3% of binding sites and gaining 70.0% more binding sites after *MYC* overexpression (**Figure 15B**). On the other hand, c-Myc binding sites at stitched enhancers were less stable, losing 41.8% of endogenous c-Myc binding sites and gaining 108.5% more binding sites after *MYC* overexpression (**Figure 15B**).

Since c-Myc protein concentration increased after *MYC* overexpression, I normalized c-Myc ChIP-seq signal at promoters, typical enhancers and superenhancers by their “MYC share” – the proportion of c-Myc binding out of the genome wide c-Myc binding signal (de Pretis et al., 2017). Consistent with de Pretis et al. (2017), promoters with low c-Myc enrichment in Low *MYC* cells showed the highest gains in MYC share after *MYC* overexpression, while promoters with the highest enrichment for c-Myc maintained their MYC share (**Figure 15C**). I extended the MYC share normalization to typical enhancers and superenhancers and observed the same trend of preferential gain in MYC share at typical enhancers and superenhancers with low pre-bound c-Myc (**Figure 15D-E**). Accordingly, *MYC* overexpression increased the proportion of c-Myc bound superenhancers from 21.9% to 46.2%, comparable to the increase in c-Myc bound promoters (34.5% to 53.1%) and in contrast to the low c-Myc occupancy at typical enhancers (3.2% to 6.5%) (**Figure 15F**). Furthermore, 26% of superenhancers gained c-Myc binding sites after *MYC* overexpression, compared to 4% of typical enhancers (**Figure 15G-H**). Since superenhancers typically span larger genomic regions than typical enhancers, I wanted to test whether the enrichment of c-Myc at superenhancers was due to size. Genomic

intervals of CREs were randomly shuffled within the same chromosome and overlapped with c-Myc binding sites to determine the random chance of c-Myc binding at CREs. Actual superenhancers overlapped with c-Myc binding sites significantly more than 1000 iterations of randomly shuffled genomic loci (**Figure 15F**), indicating that c-Myc binds preferentially at superenhancers regardless of size. Additionally, c-Myc binds next to H3K27ac peaks and not randomly within the non-acetylated regions between H3K27ac peaks in typical enhancers and superenhancers (**Figure 16A**).

Together, these results show that c-Myc binds preferentially adjacent to constituent H3K27ac peaks within superenhancers.

Figure 15 | Superenhancers are preferentially enriched in c-Myc binding sites after *MYC* overexpression

(A) Percentage of c-Myc binding sites falling within typical enhancers, superenhancers and promoters in Low *MYC* and High *MYC* cells. (***: $P < 0.001$, *: $P < 0.05$, N.S.: not significant, chi-square test) (B) Gain and loss of c-Myc binding sites at stitched enhancers and promoters after *MYC* overexpression. Percentages shown are relative to the number of c-Myc binding sites in Low *MYC* cells. (C-E) Density scatter plots showing *MYC* share of (C) promoters, (D) typical enhancers and (E) superenhancers in Low *MYC* cells against the fold change in *MYC* share after *MYC* overexpression. Trend lines show local polynomial regression fitting. Triangles mark outlier data points (top and bottom 5%). (F) Percentage of typical enhancers, superenhancers and promoters bound by c-Myc compared to random chance. Percentage of c-Myc bound CREs are shown as orange dots. Box plots show 1000 iterations of c-Myc occupancy at randomly shuffled genomic loci of the same size and on the same chromosome as the actual CREs. (G-H) Gain and loss of c-Myc peaks at (G) typical enhancers and (H) superenhancers after *MYC* overexpression.

Figure 16 | c-Myc binds next to H3K27ac peaks at CREs

(A) Average c-Myc and H3K27ac ChIP-seq signal at promoters (left), typical enhancers (middle) and superenhancers (right), centred at c-Myc ChIP-seq peaks. (B-D) Average c-Myc and H3K27ac ChIP-seq signal at lost (left), stable (middle) and gained (right) c-Myc binding sites at (B) typical enhancers, (C) superenhancers and (D) promoters.

4.4 c-Myc binding requires H3K27ac, together with canonical motif and/or G/C rich sequences

To explore the differences between c-Myc binding sites at promoters and enhancers, I compared the H3K27ac ChIP-seq signal and motif occurrence at lost, stable and gained c-Myc binding sites within typical enhancers, superenhancers and promoters. Stable c-Myc binding sites had the highest c-Myc and H3K27ac ChIP-seq signals, while gained c-Myc binding sites had lower H3K27ac signal (**Figure 16B-D**). However, lost c-Myc binding sites at promoters also had relatively high H3K27ac signal (**Figure 16B**), showing that H3K27ac is required but insufficient for c-Myc binding.

Promoters, typical enhancers and superenhancers had similar enrichment of the canonical CACGTG motif (**Figure 17A-B**). However, typical enhancers and superenhancers were markedly more enriched for the non-canonical CANNTG motif compared to promoters (**Figure 17A-B**). As expected, common c-Myc binding sites were most enriched for high-affinity CACGTG motifs and least enriched for lower-affinity CANNTG motifs compared to gained and lost c-Myc binding sites (**Figure 17C**), demonstrating preferential binding to the canonical motif. Intriguingly, the percentage of common binding sites with CANNTG motifs was higher than CACGTG motifs, yet 20% of gained c-Myc binding sites contained high-affinity CACGTG motifs (**Figure 17C**), suggesting that c-Myc binds stably to its canonical motif, but other factors are involved in recruiting c-Myc to specific binding sites.

I then wondered whether there may be different chromatin factors recruiting c-Myc to promoters and enhancers. I identified HOCOMOCO transcription factor binding motifs (Kulakovskiy et al., 2018) that co-occurred within 200bp of c-Myc ChIP-seq peak summits at promoters and enhancers. Promoter c-Myc

binding sites were highly enriched in transcription factor binding motifs, with 35 motifs occurring in more than 50% of promoters (**Figure 17D**). The most enriched motifs are for known promoter binding factors, including the KLF/SP family of transcription factors, containing guanine/cytosine (G/C) rich sequences (**Figure 17E**). In contrast, no single motif occurred in more than 50% of typical enhancers and superenhancers (**Figure 17D**). Besides the MYC and MAX family motifs, enhancers were similarly enriched for motifs containing G/C-rich sequences, such as MAZ, ZBTB17, KLF15 and EGR2, albeit at a lower enrichment (**Figure 17F**). Previous studies have shown that c-Myc binds to G/C rich probes *in vitro* (Allevato et al., 2017) and to demethylated CpG islands (CGI) *in vivo* (Guo et al., 2014). Although CGIs are usually associated with gene promoters, CGIs are also enriched at enhancers and these CGI-containing enhancers are linked to stronger regulatory activity (Bell and Vertino, 2017; Pachano et al., 2021). Here, motif analysis suggests that overexpressed c-Myc may bind at enhancers through non sequence specific binding with G/C rich sequences.

Figure 17 | Promoter c-Myc binding sites are enriched for other transcription factor motifs but not enhancer c-Myc binding sites

(A-B) Percentage of canonical CACGTG and non-canonical CANNTG motifs occurring at promoter (Pro), typical enhancer (TE) and superenhancer (SE) c-Myc binding sites in (A) Low *MYC* and (B) High *MYC* cells. (C) Percentage of canonical CACGTG and non-canonical CANNTG motifs occurring at lost, stable and gained c-Myc binding sites. (D) Percentage of HOCOMOCO transcription factor motifs co-occurring at superenhancer c-Myc binding sites versus promoter c-Myc binding sites. Each point represents 1 motif. The grey line shows equal percentage co-occurrence at promoters and superenhancer c-Myc binding sites. (E-F) Top enriched transcription factor motifs co-occurring at (E) promoter and (F) enhancer c-Myc binding sites.

4.5 Summary

In summary, overexpressed c-Myc preferentially binds near superenhancer constituent H3K27ac peaks in U2OS cells, and modulated H3K27 acetylation levels at pre-existing enhancers (**Figure 18**), consistent with observations in high c-Myc expressing cell lines as shown in Chapter 3. Although oncogenic superenhancers are frequently activated in cancers, *MYC* overexpression by itself was unable to activate novel superenhancers. Given that c-Myc does not bind to closed chromatin (Guccione et al., 2006; Lin et al., 2012; Soufi et al., 2012), other pioneer transcription factors are likely required for the activation of novel oncogenic superenhancers. Although open chromatin, demarcated by H3K27ac, is required for c-Myc binding, it is insufficient for c-Myc recruitment and other factors influence c-Myc binding, such as recognition of the c-Myc canonical motif and G/C rich sequences. Since superenhancers are highly enriched for chromatin interactions with promoters, chromatin interactions may also participate in c-Myc superenhancer invasion. In the next chapter, I explore whether chromatin interactions contribute to the preferential binding of c-Myc at superenhancers.

Figure 18 | Superenhancers gain c-Myc binding and constituent H3K27ac peaks after MYC overexpression

The black boxes represent enhancer H3K27ac peaks. The beige boxes represent c-Myc molecules. The arrows represent transcription start sites.

Chapter 5: Spike-in quantitative Hi-C (SIQHiC) reveals increased chromatin interactions after *MYC* overexpression

5.1 Spike-in Quantitative Hi-C (SIQHiC) uses mouse cell spike-in to reveal global changes in chromatin contact frequency

In the previous chapters, I showed that c-Myc overexpression resulted in superenhancer invasion, and these c-Myc bound superenhancers are associated with multiple chromatin interactions with nearby gene promoters and other c-Myc bound loci. This raised a question: does c-Myc overexpression alter the chromatin interaction landscape? Since c-Myc enhancer invasion increases gene transcription at promoters, it may also stabilize chromatin interactions between gene promoters and enhancers, resulting in global changes in chromatin contact frequency.

Genome-wide chromatin interaction frequency can be assessed through Hi-C sequencing. Traditionally, normalization across Hi-C samples is based on the assumption that chromatin interactions are largely stable across biological conditions (Lun and Smyth, 2015; Stansfield et al., 2018). Yet, global changes in chromatin contact frequency can result from changes in key chromatin architectural proteins such as CTCF and cohesin complex proteins (Rao et al., 2017; Zuin et al., 2014), as well as global changes in histone modifications such as H3K27ac (Gryder et al., 2019; Rosencrance et al., 2020). In such experiments, normalizing samples by the total number of chromatin interactions obscures the differences in chromatin contact frequency (**Figure 19, left**).

One strategy to account for such global changes is to perform cell-count normalization, by adding an external standard to cell samples in proportion to the cell count. This strategy has already been used in other “omics” techniques, such as spiking in synthetic RNA standards for RNA-Seq (Lovén et al., 2012) or cells from an orthogonal species for ChIP-Seq (Orlando et al., 2014). Recently, a cell-count normalized protocol for Hi-ChIP, termed as Absolute Quantification of Architecture HiChIP (AQuA-HiChIP) (Stanton et al., 2018), used a mouse cell spike-in strategy to perform the normalization.

Therefore, to determine whether there are global changes in chromatin contact frequency after *MYC* overexpression, I adapted the AQuA-HiChIP protocol for use with the commercially available Arima-HiC kit (Arima Genomics) to perform Spike-In Quantitative Hi-C (SIQHiC) on Low *MYC* and High *MYC* cells. Briefly, SIQHiC involves the addition of cross-linked mouse 3T3 cells into each sample of cross-linked Low *MYC* and High *MYC* cells at a ratio of 1:4 before subsequent steps of enzyme digestion, ligation, enrichment and library preparation. By mixing mouse and human cells at the start of the protocol, the ratio between human and mouse chromatin contacts (HMR) is preserved through the enrichment and amplification steps (**Figure 19, right**). Since the mouse 3T3 spike-in cells are untreated, while the human samples are treated with either doxycycline or vehicle (water), differences in the HMR can be attributed solely to the human samples. Hence, the relative difference between the HMRs of 2 samples (SIQHiC ratio) reflects the differences in global chromatin contact frequency between them (**Figure 19, right**).

Figure 19 | Spike-in Quantitative Hi-C (SIQHiC) can detect global changes in chromatin contact frequency

(A) (**Left panel**) Global changes in chromatin contact frequency cannot be detected when Hi-C samples are normalized by the total number of chromatin interactions. (**Right panel**) The SIQHiC protocol normalizes samples by spiking in mouse cells to human samples at a fixed cell number ratio. The ratio between human and mouse chromatin contacts (HMR) is preserved through enrichment, amplification and sequencing of chromatin interactions, and reflects the difference in global chromatin contact frequency between the 2 samples. (B) Brief overview of the SIQHiC workflow.

5.2 MYC overexpression increases global chromatin contact frequency

In this study, SIQHiC was performed on Low *MYC* and High *MYC* cells in duplicate. After sequencing, paired reads were mapped to a combined human and mouse genome and separated by species. Although human and mouse cells were mixed before restriction enzyme digest and ligation, in-nucleus ligation ensured that only around 0.3% of total read pairs mapped separately to different species. Human and mouse paired reads were then filtered to remove duplicates, low quality reads and intra-fragment reads to obtain the human and mouse Hi-C

contacts (**Table 8**). HMRs of Low *MYC* duplicates were similar (33.4 and 36.4), consistent with expectations that chromatin contact frequencies are stable at physiological *MYC* levels (**Table 8**). HMRs of High *MYC* duplicates were more variable (146.9 and 78.1) and were 2.1 to 4.4 fold higher than Low *MYC* duplicates (**Table 8**), suggesting an equivalent increase in chromatin contacts after *MYC* overexpression.

SIQHiC normalization was applied by scaling down Low *MYC* Hi-C contact matrices by the SIQHiC ratio such that mouse chromatin contacts are equal between samples (**Figure 20A**). Without SIQHiC normalization, Hi-C chromatin interaction heatmaps appear to show a decrease in chromatin interaction frequency after *MYC* overexpression (**Figure 20B**). SIQHiC normalization revealed that *MYC* overexpression increased chromatin interaction frequency instead (**Figure 20B**).

Table 8 | Human to mouse chromatin contact ratio (HMR)

	Unique reads	Human Contacts	Mouse Contacts	HMR	SIQHiC ratio Low <i>MYC</i> : High <i>MYC</i>
Low <i>MYC</i> Rep 1	310,510,965	266,111,828	7,972,364	33.38	1 : 4.40
High <i>MYC</i> Rep 1	302,156,580	260,697,775	1,774,657	146.90	
Low <i>MYC</i> Rep 2	322,597,046	274,563,436	7,534,757	36.44	1 : 2.14
High <i>MYC</i> Rep 2	328,864,586	280,568,109	3,594,411	78.06

Figure 20 | SIQHiC normalization reveals increased chromatin contact frequency per cell after *MYC* overexpression.

(A) Cartoon illustrating Hi-C contacts before (top panel) and after (bottom panel) SIQHiC normalization. Low *MYC* contacts were scaled down such that the number of mouse contacts in both conditions were the same. (B) Hi-C contact matrix heatmaps of a region on chromosome 1. Top panel: No normalization. Bottom panel: SIQHiC normalization.

5.3 *MYC* overexpression increases chromatin contact frequency specifically at a set of gained loops

Since *MYC* overexpression increased global chromatin contact frequency, I questioned whether there is a global restructuring of the chromatin interaction landscape. At the megabase scale, A/B compartment analysis showed few instances of compartment switching after *MYC* overexpression (**Figure 21A**). Similar numbers of TADs were identified in Low *MYC* and High *MYC* cells, with more than 60% overlap between them (**Figure 21B and C**). These results suggest that active and repressive domains remain stable after *MYC* overexpression.

7,266 significant chromatin loops were identified in Low *MYC* cells, slightly increasing to 7,910 in High *MYC* cells after *MYC* overexpression (**Figure 22A**). 4,478 loops (~60%) were common between Low *MYC* and High *MYC* cells, while 2,782 loops were lost and 3,429 loops were gained after *MYC* overexpression (**Figure 22A**). I performed an aggregate peak analysis (APA) on lost, common and gained loops to compare the aggregate signal of these loop sets across both conditions. APA stacks the midpoint of all loops in each loop set on top of each other, as well as the 105kb x 105kb contact matrices surrounding these loops. The resultant APA heatmaps illustrate the relative difference between the aggregate peak signal (P) and the local background (represented by the lower left corner – LL), indicated by the peak to lower left ratio (P2LL) score. Without normalization, APA heatmaps show a 30% decrease in peak signal of all loops after *MYC* overexpression, particularly at lost loops (**Figure 22B and C**). However, it is evident that the local background signals are also decreased for all loop sets between both conditions (**Figure 22C**). Normalizing for cell count using SIQHiC revealed a 2.1 fold increase in peak signal of all loops, with a 2.2 fold increase at common loops and 3.3 fold increase at gained loops (**Figure 23A and C**). When APA heatmaps are normalized by their local background signal (LL), P2LL scores of all loops increased by 1.3 fold, predominantly at gained loops (2.1 fold increase) (**Figure 23B and D**). Taking both normalizations into consideration, the global increase in chromatin contact frequency is not merely due to an increase in random associations but particularly focused at a set of gained loops.

Figure 21 | *MYC* overexpression does not significantly alter chromatin compartments and topological associating domains (TADs).

(A) A/B compartments are not altered after *MYC* overexpression, as shown in a representative view of chromosome 8. Pearson's correlation heatmaps were drawn using first principal component (PC1) at 100kb resolution, shown as tracks above the heatmaps. (B-C) Overlap between topological associated domains (TADs) identified in Low *MYC* and High *MYC* cells at (B) 50kb and (C) 100kb resolution.

Figure 22 | MYC overexpression appears to decrease chromatin interaction frequency without normalization.

(A) Overlap between chromatin loops identified in Low *MYC* and High *MYC* cells. (B) Fold change of aggregate peak signal enrichment (P) between High *MYC* and Low *MYC* cells. (C) Aggregate peak analysis (APA) heatmaps at 5kb resolution showing aggregate signal of loops within the “Lost”, “Common”, “Gained” and “All” loop sets, as well as the 105kb x 105kb contact matrices around loop centroids. Loop sets were filtered to remove short loops near the diagonal (shown above each heatmap; numerator: filtered loops; denominator: total loops). P: Aggregate peak signal enrichment.

Figure 23 | Normalization of APA heatmaps by SIQHiC and by local background signal reveal increased chromatin interaction frequency after *MYC* overexpression, particularly at gained loops.

(A-B) Aggregate peak analysis (APA) heatmaps at 5kb resolution normalized by (A) SIQHiC and (B) local background signal, showing aggregate signal of loops within the “Lost”, “Common”, “Gained” and “All” loop sets, as well as the 105kb x 105kb contact matrices around loop centroids. Loop sets were filtered to remove short loops near the diagonal (shown above each heatmap; numerator: filtered loops; denominator: total loops). P: Aggregate peak signal enrichment. LL: Average signal at the lower left corner, representing local background signal. P2LL: Peak to lower left ratio. (C) Fold change of SIQHiC normalized aggregate peak signal enrichment (P) between High *MYC* and Low *MYC* cells. (D) Fold change of P2LL ratio between High *MYC* and Low *MYC* cells.

I performed a Gene Ontology (GO) term analysis of the c-Myc bound genes at lost and gained chromatin loops to ascertain whether these chromatin interaction alterations were functional. Lost chromatin loops were enriched for general GO terms, including protein binding, poly(A) RNA binding and negative regulation of transcription (**Figure 24A**). Gained loops enriched for previously identified *MYC*-regulated pathways shown in **Table 6** and **Table 7**, such as transcription, translation, splicing and cell adhesion (**Figure 24B**), suggesting that the increased chromatin interaction frequency at gained loops contributes to gene regulation by *MYC* overexpression. Gained and lost chromatin loops did not associate with a general direction of gene regulation (**Figure 25A-C**). However, gained chromatin loops were associated with more significantly regulated transcripts compared to lost and common loops (**Figure 25A-C**), suggesting that chromatin interactions contributes to gene regulation in both directions.

Figure 24 | Gained chromatin loops are enriched for genes differentially regulated by c-Myc.

(A-B) Top significantly enriched gene ontology (GO) terms of c-Myc bound genes at (A) lost and (B) gained chromatin loops, based on GO enrichment test FDR value.

Figure 25 | Gained loops are associated with more significantly regulated transcripts.

(A-C) Significantly regulated transcripts associated with (A) lost, (B) common and (C) gained chromatin loops ($|\beta| > 1$, FDR < 0.05 , Wald test).

Next, I looked at the chromatin interactions around the *MYC* gene locus, because *MYC* gene expression is known to be highly regulated by nearby superenhancers through chromatin interactions. Using SIQHiC normalized and non-normalized data, I generated virtual 4C plots showing chromatin interactions at this locus

with the *MYC* gene promoter as the viewpoint. Consistent with the global APA analysis, non-normalized virtual 4C showed a general decrease in the frequency of chromatin interactions at the *MYC* gene promoter, but SIQHiC normalization revealed a general increase in chromatin interaction frequency instead (**Figure 26**).

Although Hi-C allows for genome-wide identification of chromatin interactions, it requires very high sequencing depth to obtain adequate resolution for comprehensive loop detection. 4C-seq focuses only on chromatin interactions at a specific viewpoint, thereby reducing the amount of sequencing required. Hence, I performed 4C-seq to visualize chromatin interactions emanating from a randomly selected c-Myc bound superenhancer near the *PROC* gene which gained a chromatin loop to the *CYP27C1* gene after *MYC* overexpression. Consistent with the Hi-C data, 4C-seq detected this gained chromatin loop in Low *MYC* cells and showed an increase in chromatin interaction frequency after *MYC* overexpression (**Figure 27**). 4C-seq also identified additional gained loops (**Figure 27, blue arrows**) that were not significant in the Hi-C data.

Collectively, these results indicate that *MYC* overexpression increases chromatin contact frequency as a whole, and strengthens a set of gained chromatin loops in particular.

Figure 26 | *MYC* overexpression increases chromatin contact frequency at the *MYC* gene locus

UCSC Genome browser view of High *MYC* and Low *MYC* cells at the *MYC* locus. Tracks show Hi-C chromatin loops, virtual 4C signal of interactions at the *MYC* promoter (red and blue tracks), H3K27ac and c-Myc ChIP-seq signal. The difference between the virtual 4C signals of High *MYC* and Low *MYC* cells are shown with and without SIQHiC normalization (green tracks). Superenhancers (red) and typical enhancers (black) are shown as bars below the H3K27ac ChIP-seq tracks. c-Myc binding peaks are shown as bars below the c-Myc ChIP-seq tracks.

Figure 27 | MYC overexpression increases chromatin contact frequency and chromatin loops at superenhancers
 4C-seq of chromatin interactions at a c-Myc bound superenhancer near the *PROC* gene (highlighted in pink) on chromosome 2. Tracks show Hi-C chromatin loops, 4C-seq signal (red and blue tracks), difference between 4C-seq signal of High *MYC* and Low *MYC* cells (green), H3K27ac and c-Myc ChIP-seq signal. The Hi-C gained chromatin loop is highlighted in orange. Blue arrows show additional gained chromatin interactions identified by 4C-seq.

5.4 Chromatin interactions are particularly enriched at superenhancers and tend to connect c-Myc binding sites

In the previous section, I have shown that *MYC* overexpression increases chromatin contact frequency at chromatin loops. I proceeded to characterize these chromatin loops to determine whether this increase is directly facilitated by c-Myc superenhancer invasion. I intersected chromatin loop anchors with c-Myc binding sites and found that 21.8% of all chromatin loops were occupied by c-Myc in Low *MYC* cells, increasing to 32.5% after *MYC* overexpression (**Figure 28A**). Gained loops were exquisitely enriched for c-Myc binding even prior to *MYC* overexpression (34.3%), compared to lost (14.5%) and common loops (16.8%) (**Figure 28A**). The proportion of c-Myc bound gained loops was significantly higher than randomly shuffled loop coordinates, indicating that these chromatin loops preferentially occur at c-Myc binding sites (**Figure 28A**). I had previously shown in **Figure 8** that chromatin interactions connected c-Myc bound loci together in SNU16 and K562 cells. In U2OS cells, 15.7% of gained loops were bound by c-Myc at both loop anchors after *MYC* overexpression, with 7.6% of gained loops connecting genomic loci that were already bound by c-Myc at physiological *MYC* levels (**Figure 28B**). Additionally, gained c-Myc binding sites were more associated with chromatin loops compared to lost c-Myc binding sites, while common c-Myc binding sites were most associated with chromatin (**Figure 28C**). These results indicate that c-Myc bound genomic loci interact more frequently after *MYC* overexpression. In support of this, immunofluorescence staining showed that c-Myc forms discrete punctate spots within the nucleus after *MYC* overexpression (**Figure 29**). These c-Myc foci overlap regions with high RNA polymerase II signal (**Figure 29**), consistent with c-Myc binding at active gene promoters.

Figure 28 | Gained chromatin loops are enriched for c-Myc and tend to connect c-Myc binding sites together

(A) Percentage of Hi-C chromatin loops occupied by c-Myc compared to randomly shuffled loop coordinates. Percentage of c-Myc bound chromatin loops are shown as orange dots. Box plots show 1000 iterations of c-Myc occupancy at random genomic loci of the same size and on the same chromosome as the actual chromatin loops. (B) Percentage of chromatin loop anchors occupied by c-Myc. (C) Percentage of c-Myc binding sites associated with Hi-C loops.

Figure 29 | Overexpressed c-Myc formed discrete punctate spots overlapping RNA polymerase II within the nucleus

Immunofluorescence staining of c-Myc and RNA polymerase II in High MYC cells. From left to right, immunofluorescence images show merged signal, DAPI staining of DNA in blue, c-Myc staining in green and RNA polymerase II staining in red. RNA Pol II: RNA Polymerase II.

Next, I annotated chromatin loops according to their overlap with promoters and stitched enhancers. Around 10% of lost and common loops connected promoters together, 13% connected stitched enhancers together and 15%-19% connected promoters to stitched enhancers (**Figure 30A**). In contrast, gained loops were more enriched in chromatin loops connecting CREs together, with 30% connecting promoters, 23% connecting stitched enhancers and 26% connecting promoters to stitched enhancers (**Figure 30A**), demonstrating preferential strengthening of interactions between CREs. Previous research had shown that superenhancers were highly associated with chromatin interactions (Cao et al., 2017). In U2OS cells, 42.0% of superenhancers were associated with significant chromatin loops and further increased to 66.2% after *MYC* overexpression, while typical enhancers and promoters were less associated with chromatin loops (**Figure 30B**).

Figure 30 | c-Myc bound cis-regulatory elements preferentially gain chromatin loops

(A) Percentage of lost, common and gained chromatin loops connecting cis-regulatory elements. (B) Percentage of cis-regulatory elements with chromatin loops. (C) Percentage of c-Myc bound and unbound cis-regulatory elements with chromatin loops.

In particular, c-Myc bound CREs were more associated with chromatin loops than non c-Myc bound CREs, especially c-Myc bound superenhancers (**Figure 31A**). *MYC* overexpression particularly increased chromatin looping at typical enhancers and promoters occupied by c-Myc, although superenhancers gained chromatin loops regardless of c-Myc occupancy (**Figure 31A**). Focusing on promoter-enhancer (P-E) loops, *MYC* overexpression increased c-Myc binding at all promoter-proximal loop anchors, but gained P-E loops were markedly enriched for c-Myc binding at the distal enhancer loop anchor (23%) compared to common loops (9%) and lost loops (8%), and 9% of gained P-E loops connected c-Myc binding sites present prior to *MYC* overexpression (**Figure 31B**). Examples of gained P-E loop associated oncogenes include *DDX18*, an RNA helicase involved in translation and splicing, and *HK2*, which participates in aerobic glycolysis. *MYC* overexpression connected *DDX18* and *HK2* to neighbouring enhancers that had gained c-Myc binding (**Figure 32A-B**), leading

to a 1.5 fold and 4 fold increase in gene expression respectively (**Figure 32C-D**). Taken together, these results suggest that increased c-Myc expression strengthens chromatin interactions between c-Myc binding sites at enhancers and promoters.

Figure 31 | MYC overexpression increases chromatin interactions between c-Myc bound enhancers and promoters

(A) Percentage of c-Myc unbound (blue) and bound (red) typical enhancers, superenhancers and promoters with Hi-C chromatin loops in Low *MYC* and High *MYC* cells. (B) Percentage of lost (left), common (middle) and gained (right) promoter-enhancer (P-E) chromatin loops occupied by c-Myc at the promoter proximal anchor or at both anchors, as illustrated above the bar graphs.

Figure 32 | *MYC* overexpression connected cancer-associated genes to neighbouring c-Myc bound enhancers.

(A-B) UCSC Genome browser view of High *MYC* and Low *MYC* cells at the (A) *DDX18* and (B) *HK2* locus. Tracks show Hi-C chromatin loops, H3K27ac and c-Myc ChIP-seq signal. Typical enhancers are shown as bars below the H3K27ac ChIP-seq tracks. c-Myc binding peaks are shown as bars below the c-Myc ChIP-seq tracks. (C-D) RNA-seq normalized transcripts per million (TPM) expression of (C) *DDX18* and (D) *HK2* in U2OS low *MYC* and high *MYC* cells. Error bar shows standard error across 2 biological replicates.

5.5 *MYC* overexpression increased chromatin interaction frequency regardless of transcription activity

Transcriptional activity may play a role in maintaining the chromatin interaction

landscape. For example, coding and non-coding transcription can reposition cohesin to direct chromatin looping (Busslinger et al., 2017; Heinz et al., 2018; Isoda et al., 2017), and increases the mobility of CREs (Gu et al., 2018). Since c-Myc overexpression increases transcriptional activity, I questioned whether the observed increase in chromatin interaction frequency was a consequence of increased transcription. Firstly, I compared chromatin looping at transcripts significantly regulated by c-Myc ($P < 0.05$, $\text{abs}(\beta) > 1$) with a set of non-regulated transcripts ($P > 0.95$, transcripts per million > 1). Chromatin loops were not preferentially gained or lost at the up-regulated and down-regulated transcripts compared to the non-regulated transcripts (**Figure 33A**). However, more c-Myc bound transcripts gained chromatin loops (**Figure 33B**) compared to transcripts lacking c-Myc binding (**Figure 33C**), regardless of their differential transcript expression.

Figure 33 | c-Myc bound transcripts gain chromatin loops regardless of differential transcript expression

(A-C) Percentage of (A) all, (B) c-Myc bound and (C) non c-Myc bound transcripts with Hi-C chromatin loops. Transcripts were stratified into upregulated (orange), downregulated (blue) and non-regulated (green) transcripts.

Next, I categorized transcripts into tertiles based on transcript expression after *MYC* overexpression. More top tertile transcripts overlapped with chromatin loops compared to bottom tertile transcripts (**Figure 34A**), indicating that transcription activity is associated with chromatin interactions. However, more c-Myc bound transcripts gained chromatin loops (**Figure 34B**) compared to transcripts without c-Myc binding (**Figure 34C**). Although transcription activity is associated with increased chromatin looping, these results indicate that direct binding of c-Myc increases chromatin interaction frequency regardless of transcript expression levels or differential expression.

Figure 34 | c-Myc binding increases chromatin interaction frequency regardless of transcript expression levels

(A-C) Percentage of (A) all transcripts, (B) c-Myc bound transcripts and (C) non c-Myc bound transcripts with Hi-C chromatin loops. Transcripts were stratified into tertiles based on transcript expression after *MYC* overexpression.

5.6 Summary

Using SIQHIC to normalize Hi-C samples by cell count revealed that *MYC* overexpression increased the global chromatin contact frequency, particularly strengthening interactions between c-Myc binding sites at enhancers and promoters. Superenhancers were especially enriched for c-Myc binding and chromatin interactions. Gained promoter-enhancer chromatin loops tended to be

bound by c-Myc at the promoter loop anchor and acquired c-Myc at the enhancer loop anchor after *MYC* overexpression. In other words, overexpressed c-Myc concentrated at localized regions of high transcriptional activity, binding to enhancers that were linearly distant but spatially proximal to c-Myc bound promoters (**Figure 35**). Gained loops overlap better with c-Myc bound promoters regardless of their transcriptional activity, indicating that the binding of c-Myc directly contributes to the increase in contact frequency.

Figure 35 | c-Myc concentrates at localized regions of high transcriptional activity

Yellow spheres represent transcription factors. Blue spheres represent c-Myc molecules. Red lines indicate superenhancers along the black lines representing chromatin.

Chapter 6: Conclusions and future directions

In this thesis, I have shown that almost all superenhancers are bound by c-Myc in high *MYC* expressing cancer cell lines. Overexpressing c-Myc increases c-Myc binding sites and constituent H3K27ac peaks within superenhancers, potentially increasing regulatory activity. Spike-In Quantitative Hi-C (SIQHiC) revealed a global increase in chromatin interaction frequency, especially between enhancers and promoters bound by c-Myc. In this chapter, I will discuss how these findings provide a better understanding of the interplay between *MYC* overexpression and the chromatin interaction landscape in cancer, and propose future directions for further research.

6.1 c-Myc preferentially invades into active superenhancers

MYC overexpression is frequently observed in multiple cancer cell types and the general consensus is that overexpressed c-Myc binds to enhancers and increases expression of active genes. However, it is not clear why different oncogenes are selectively upregulated by c-Myc in different cancers. In cancer, superenhancers form cancer-specific signatures that aberrantly drive oncogene expression. In this work, I found that almost all superenhancers were occupied by c-Myc in 3 high *MYC* expressing cancers, and 4C-seq showed that superenhancers formed chromatin interactions with promoters of cancer-associated genes, potentially providing the avenue by which c-Myc selectively regulates cancer-specific oncogenes. *MYC* shRNA knockdown reduced c-Myc binding sites at superenhancers in LNCaP cells while *MYC* overexpression in U2OS cells increased c-Myc binding sites at superenhancers, indicating that c-Myc binding at superenhancers is concentration dependent. c-Myc bound adjacent to H3K27ac peaks and modulated H3K27 acetylation levels, increasing both c-Myc binding sites as well as constituent H3K27ac peaks within superenhancers, potentially increasing the regulatory activity of superenhancers.

Although *MYC* overexpression modulated H3K27 acetylation and increased H3K27ac constituent peaks within superenhancers, *MYC* overexpression by itself was unable to activate novel superenhancers in U2OS cells. This is in line with

previous studies which showed that c-Myc does not bind to closed chromatin (Guccione et al., 2006; Lin et al., 2012; Soufi et al., 2012). In a separate study on *MYC* overexpression in hTERT-immortalized human mammary epithelial (IMEC) cells, activation of *de novo* enhancers was observed after 2 rounds of clonal growth from single cells but not in the original *MYC*-overexpressed IMEC cells (Poli et al., 2018). This indicates that the acquisition of oncogenic superenhancers may occur downstream of *MYC* overexpression as part of the cell's adaptation to sustained high *MYC* expression. For example, *MYC* overexpression activates the hypoxic response (Li et al., 2020b) and replication stress response pathways (Kuzyk and Mai, 2014), which have been shown to modulate superenhancer activity (Moreau et al., 2018) and increase chromatin accessibility (Murai et al., 2020). Taken together, these results support my 1st hypothesis that c-Myc preferentially invades into superenhancers, but also indicate that c-Myc does not activate silent oncogenic superenhancers by itself.

6.2 Spike-in Quantitative Hi-C normalizes Hi-C samples by cell count to detect global changes in chromatin contact frequency

In this thesis, I introduced Spike-In Quantitative Hi-C (SIQHiC), which incorporates a mouse cell spike-in strategy to normalize Hi-C samples by cell count. Traditional Hi-C techniques perform cross-sample normalization based on the assumption that chromatin interactions are stable between samples, thereby detecting relative changes in chromatin interactions. By normalizing for cell count, SIQHiC allows for the detection of absolute changes in genome wide chromatin contact frequency, which can occur when chromatin architectural proteins or epigenetic modifiers are perturbed (Gryder et al., 2019; Rao et al., 2017; Rosencrance et al., 2020; Zuin et al., 2014).

Since SIQHiC relies on spiking in mouse cells to human cells at a known ratio, cell counting is a critical step. Therefore, mouse and human cell counts were obtained using an average of 3 counts using an automated cell counter. Given that chromatin interactions are largely stable between untreated Hi-C replicates, HMR of control replicates (containing untreated mouse and human cells) can serve as a measure of cell counting consistency.

In this study, although mouse cells make up 20% of the sample, mouse chromatin contacts constitute only 1% to 3% of the total chromatin contacts in each sample. This discrepancy may be due to the following reasons. Firstly, the mouse genome is around 14% smaller than the human genome (Chinwalla et al., 2002), resulting in fewer pairwise interactions that can be obtained per cell. Secondly, in-nucleus restriction enzyme digestion and blunt end ligation efficiencies may differ between human and mouse cells. Thirdly, although human and mouse genomes have similar overall G/C content, the local distribution of G/C content across the human genome is much more varied than in mice genome (Chinwalla et al., 2002), possibly contributing to differences in PCR amplification bias. Nevertheless, these differences between mouse and human cells should not affect SIQHiC normalization across samples because the same number of human and mouse cells are processed together in each sample, balancing out these biases.

6.3 SIQHiC revealed increased chromatin contact frequency between c-Myc binding sites at enhancers and promoters

c-Myc binds at gene promoters at physiological conditions, invading into enhancers and superenhancers at supraphysiological expression levels. Using SIQHiC, I showed that *MYC* overexpression increased global chromatin contact frequency, especially between c-Myc binding sites at enhancers and promoters. Gained promoter-enhancer chromatin loops tended to be occupied by c-Myc at the promoter loop anchor and gain c-Myc binding at the distal loop anchor after *MYC* overexpression, supporting the hypothesis that chromatin interactions facilitate c-Myc enhancer invasion by bringing enhancers close to c-Myc bound promoters. Together, these results suggest that c-Myc accumulates at localized regions of high transcriptional activity.

Previous studies have shown that active transcription may contribute to the maintenance of the genome architecture (Busslinger et al., 2017; Gu et al., 2018; Heinz et al., 2018; Isoda et al., 2017). Although *MYC* overexpression increases transcriptional activity, c-Myc bound promoters consistently associate better with increased chromatin interactions than non c-Myc bound promoters,

regardless of transcription levels. This points towards the direct participation of c-Myc in bringing promoters and enhancers together.

Recent studies have provided evidence supporting liquid-liquid phase separation (LLPS) as a mechanism for transcription regulation by superenhancers (Hnisz et al., 2017). In this model, transcription factors such as RNA polymerase II and Mediator accumulate at regions of high transcriptional activity, where superenhancers connect with gene promoters, and interact with each other to form local assemblies of high protein concentration (Boija et al., 2018; Cho et al., 2018; Sabari et al., 2018). These assemblies further accumulate transcription factors that can participate in these interactions, forming a membraneless condensate that constrains these participants at superenhancers and promoters. Transcription factors that interact in LLPS condensates frequently contain intrinsically disordered regions with low complexity, and c-Myc has been shown to participate in such condensates *in vitro* (Boija et al., 2018). In this work, immunofluorescence staining of c-Myc and RNA polymerase II in High *MYC* cells showed co-localization at discrete puncta within the nucleus, supporting this model.

The LLPS model of transcription regulation can explain the observations in this work. LLPS sequesters and accumulates overexpressed c-Myc molecules at promoters and interacting superenhancers, allowing c-Myc to bind at lower affinity binding sites at superenhancers after saturating canonical binding sites at promoters. This would additionally explain the loss of c-Myc binding sites in U2OS cells after *MYC* overexpression (**Figure 15**) and the gain of c-Myc binding sites in LNCaP cells after *MYC* shRNA knockdown (**Figure 9**), since the sequestration of overexpressed c-Myc at highly transcribed regions would deplete c-Myc at less active regions in the nucleus in U2OS cells. Supporting this, lost c-Myc binding sites were less associated with promoters, enhancers (**Figure 15B**) and chromatin loops (**Figure 28C**) compared to gained c-Myc binding sites in U2OS cells. At the same time, the increase in protein-protein interactions between c-Myc molecules and other transcription factors restricts chromatin mobility and prolongs chromatin contact between superenhancers and promoters, thereby increasing chromatin contact frequency. Further

investigations are required to confirm the presence of LLPS condensates at c-Myc enriched promoter-superenhancer foci in cancer.

Cancers often acquire oncogenic superenhancers that drive oncogene expression and exhibit altered chromatin conformations. These results suggest that c-Myc may be playing a role in maintaining the altered genome architecture in cancer by stabilizing the transient changes in chromatin interactions that were acquired along the way. Strengthened chromatin interactions were enriched for transcription, translation, splicing and DNA replication genes – pathways that promote cancer cell growth and proliferation. High *MYC* expressing cancers may then be reliant on these chromatin interactions. In section 6.5, I propose future work in targeting chromatin interactions as a therapeutic strategy against high *MYC* expressing cancers.

6.4 Significance

Overexpressed c-Myc had been shown to invade into enhancers in cancer, but the underlying mechanisms remain unclear. Collectively, my findings revealed the accumulation of overexpressed c-Myc at transcriptionally active regions of superenhancers and promoters, shedding new light on the mechanisms underlying c-Myc enhancer invasion. These results also suggest that high *MYC* expressing cancers may be reliant on c-Myc for the maintenance of the cancer genome architecture, and lay the groundwork for further research on targeting chromatin interactions as a therapeutic strategy against high *MYC* expressing cancers.

6.5 Future directions

In future work, I intend to further investigate c-Myc superenhancer invasion in cancer, especially in the maintenance of the cancer genomic architecture, in hopes of formulating new therapeutic strategies against high *MYC* expressing cancers.

Superenhancers are composed of multiple H3K27ac peaks, but these constituent

enhancers may not contribute equally to regulate transcription. In some superenhancers, selected constituent enhancers are particularly enriched for chromatin interactions with other constituent enhancers and gene promoters (Huang et al., 2018). These “hub enhancers” are associated with higher regulatory activity and their removal profoundly impacts transcription (Huang et al., 2018; Huang et al., 2016; Shin et al., 2016). At the same time, some promoters act as hubs to form extensive chromatin interactions with multiple distal enhancers that function redundantly to maintain gene transcription (Osterwalder et al., 2018). Given that c-Myc binds at enhancers and promoters and increases chromatin contact frequency at these sites, it will be interesting to investigate whether these c-Myc occupied enhancers and promoters act as regulatory hubs to control transcriptional activity. To test this hypothesis, CRISPR-Cas9/dCas9 based genetic screens can be performed on constituent enhancers within superenhancers to compare the regulatory function of enhancers with and without c-Myc binding (Korkmaz et al., 2016; Li et al., 2020a). Specifically, dCas9 fusions with repressor domains such as KRAB and LSD1 can be targeted to individual constituent enhancers to screen for enhancer function. c-Myc binding sites at constituent enhancers can also be deleted using CRISPR-Cas9 to determine the importance of c-Myc to enhancer function.

In this work, I introduced SIQHiC, a method that incorporates cell count normalization into Hi-C. Here, I performed SIQHiC on *MYC* overexpressed U2OS cells to uncover an increase in global chromatin contact frequency. This method will be a valuable tool to understand the global changes in chromatin architecture that occur in cancer. SIQHiC will also enable the analysis of wide scale genome reorganization that occurs during stem cell differentiation. SIQHiC will also be suitable for understanding the action of chromatin remodellers on genome organization at the global level.

One limitation of SIQHiC is that the method only assays for pairwise interactions and is inadequate for investigating the multivalent chromatin interactions (simultaneous interactions between 3 or more genomic loci) that are predicted to occur at hub enhancers and promoters. New chromatin conformation techniques have been developed in recent years to tackle the problem of capturing

multivalent chromatin interactions, such as split-pool recognition of interactions by tag extension (SPRITE) (Quinodoz et al., 2018) and ChIA-Drop (Zheng et al., 2019). Continuing this work, the mouse cell spike-in strategy employed in SIQHiC can be applied to these techniques to characterize the multivalent chromatin interactions occurring at enhancers and promoters with and without c-Myc binding.

Having shown that *MYC* overexpression promotes chromatin interactions, an important area for further investigation is whether these chromatin interactions are integral to c-Myc function. If so, perturbing chromatin interactions could be a novel method to target high *MYC* expressing cancers. In a recent study, degradation of CTCF disrupted enhancer-promoter chromatin interactions and immediately reduced the expression of *MYC* and its downstream target genes, without significantly affecting global gene transcription (Hyle et al., 2019). Prolonged CTCF degradation in B-cell acute lymphoblastic leukaemia cells led to cell death, although it is not clear whether this was due to the depletion of c-Myc or disruption of chromatin organization (Hyle et al., 2019). However, mouse embryonic stem cells could tolerate the degradation of CTCF, and reintroduction of CTCF restored chromatin loops that were lost after CTCF degradation (Nora et al., 2017). This raises hopes that cancers can be uniquely targeted by titrating the duration and extent of CTCF degradation without adversely affecting normal cells. Besides the removal of CTCF, a class of anti-cancer drugs known as curaxins can disrupt chromatin looping by intercalating with DNA and interfering with CTCF binding (Kantidze et al., 2019), without causing DNA damage (Gasparian Alexander et al., 2011; Kantidze et al., 2019; Lu et al., 2021). Curaxins were well tolerated in xenograft models across multiple cancer cell types (Barone et al., 2017; Dermawan et al., 2016; Kim et al., 2016) and recently completed phase I clinical trials (Sarantopoulos et al., 2020), making it a promising drug. Further studies will be required to determine whether *MYC*-addicted cancers are uniquely susceptible to CTCF depletion or inhibition, compared to low *MYC* cancers and normal cells.

YY1 is another potential chromatin architectural protein that can be targeted. YY1 was proposed to facilitate enhancer-promoter chromatin interactions

through homodimer formation (Beagan et al., 2017; Weintraub et al., 2017). Importantly, YY1 had been shown to bind co-ordinately with c-Myc to activate gene expression (Shrivastava et al., 1996; Vella et al., 2012; Vernon and Gaston, 2000), likely playing a role in transcription regulation by c-Myc.

Through these future investigations, I hope to build on the findings of this thesis to further improve our understanding of c-Myc and its role in cancer chromatin organization and provide novel strategies for cancer therapy.

References

- Affer, M., Chesi, M., Chen, W.D., Keats, J.J., Demchenko, Y.N., Tamizhmani, K., Garbitt, V.M., Riggs, D.L., Brents, L.A., Roschke, A.V., Van Wier, S., Fonseca, R., Bergsagel, P.L., Kuehl, W.M., 2014. Promiscuous MYC locus rearrangements hijack enhancers but mostly super-enhancers to dysregulate MYC expression in multiple myeloma. *Leukemia* 28, 1725-1735.
- Ahmadiyeh, N., Pomerantz, M.M., Grisanzio, C., Herman, P., Jia, L., Almendro, V., He, H.H., Brown, M., Liu, X.S., Davis, M., Caswell, J.L., Beckwith, C.A., Hills, A., MacConaill, L., Coetzee, G.A., Regan, M.M., Freedman, M.L., 2010. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proceedings of the National Academy of Sciences* 107, 9742.
- Akhtar-Zaidi, B., Cowper-Sal-lari, R., Corradin, O., Saiakhova, A., Bartels, C.F., Balasubramanian, D., Myeroff, L., Lutterbaugh, J., Jarrar, A., Kalady, M.F., Willis, J., Moore, J.H., Tesar, P.J., Laframboise, T., Markowitz, S., Lupien, M., Scacheri, P.C., 2012. Epigenomic enhancer profiling defines a signature of colon cancer. *Science* 336, 736-739.
- Akincilar, S.C., Khattar, E., Boon, P.L.S., Unal, B., Fullwood, M.J., Tergaonkar, V., 2016. Long-Range Chromatin Interactions Drive Mutant TERT Promoter Activation. *Cancer Discovery* 6, 1276.
- Allevato, M., Bolotin, E., Grossman, M., Mane-Padros, D., Sladek, F.M., Martinez, E., 2017. Sequence-specific DNA binding by MYC/MAX to low-affinity non-E-box motifs. *PLOS ONE* 12, e0180147.
- Alshaker, H., Mills, R., Hunter, E., Salter, M., Ramadass, A., Skinner, B.M., Westra, W., Green, J., Akoulitchev, A., Winkler, M., Pchejetski, D., 2021. Chromatin conformation changes in peripheral blood can detect prostate cancer and stratify disease risk groups. *Journal of Translational Medicine* 19, 46.

- An, O., Tan, K.T., Li, Y., Li, J., Wu, C.S., Zhang, B., Chen, L., Yang, H., 2020. CSI NGS Portal: An Online Platform for Automated NGS Data Analysis and Sharing. *Int J Mol Sci* 21.
- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., Ntini, E., Arner, E., Valen, E., Li, K., Schwarzfischer, L., Glatz, D., Raithel, J., Lilje, B., Rapin, N., Bagger, F.O., Jørgensen, M., Andersen, P.R., Bertin, N., Rackham, O., Burroughs, A.M., Baillie, J.K., Ishizu, Y., Shimizu, Y., Furuhashi, E., Maeda, S., Negishi, Y., Mungall, C.J., Meehan, T.F., Lassmann, T., Itoh, M., Kawaji, H., Kondo, N., Kawai, J., Lennartsson, A., Daub, C.O., Heutink, P., Hume, D.A., Jensen, T.H., Suzuki, H., Hayashizaki, Y., Müller, F., Forrest, A.R.R., Carninci, P., Rehli, M., Sandelin, A., 2014. An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455-461.
- Aukema, S.M., Siebert, R., Schuurin, E., van Imhoff, G.W., Kluin-Nelemans, H.C., Boerma, E.-J., Kluin, P.M., 2011. Double-hit B-cell lymphomas. *Blood* 117, 2319-2331.
- Babu, D., Fullwood, M.J., 2015. 3D genome organization in health and disease: emerging opportunities in cancer translational medicine. *Nucleus* 6, 382-393.
- Bahram, F., von der Lehr, N., Cetinkaya, C., Larsson, L.G., 2000. c-Myc hot spot mutations in lymphomas result in inefficient ubiquitination and decreased proteasome-mediated turnover. *Blood* 95, 2104-2110.
- Banerji, J., Olson, L., Schaffner, W., 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33, 729-740.
- Barone, T.A., Burkhart, C.A., Safina, A., Haderski, G., Gurova, K.V., Pural, A.A., Gudkov, A.V., Plunkett, R.J., 2017. Anticancer drug candidate CBL0137, which inhibits histone chaperone FACT, is efficacious in preclinical orthotopic models of temozolomide-responsive and -resistant glioblastoma. *Neuro Oncol* 19, 186-196.
- Bastonini, E., Jeznach, M., Field, M., Juszczak, K., Corfield, E., Dezfouli, M., Ahmat, N., Smith, A., Womersley, H., Jordan, P., Ramadass, A., Akoulitchev, A., Goding, C.R., 2014. Chromatin barcodes as

- biomarkers for melanoma. *Pigment Cell & Melanoma Research* 27, 788-800.
- Beagan, J.A., Duong, M.T., Titus, K.R., Zhou, L., Cao, Z., Ma, J., Lachanski, C.V., Gillis, D.R., Phillips-Cremins, J.E., 2017. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome research*.
- Beer, S., Zetterberg, A., Ihrle, R.A., McTaggart, R.A., Yang, Q., Bradon, N., Arvanitis, C., Attardi, L.D., Feng, S., Ruebner, B., Cardiff, R.D., Felsher, D.W., 2004. Developmental Context Determines Latency of MYC-Induced Tumorigenesis. *PLOS Biology* 2, e332.
- Bell, J.S.K., Vertino, P.M., 2017. Orphan CpG islands define a novel class of highly active enhancers. *Epigenetics* 12, 449-464.
- Beroukhi, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., Mc Henry, K.T., Pinchback, R.M., Ligon, A.H., Cho, Y.J., Haery, L., Greulich, H., Reich, M., Winckler, W., Lawrence, M.S., Weir, B.A., Tanaka, K.E., Chiang, D.Y., Bass, A.J., Loo, A., Hoffman, C., Prensner, J., Liefeld, T., Gao, Q., Yecies, D., Signoretti, S., Maher, E., Kaye, F.J., Sasaki, H., Tepper, J.E., Fletcher, J.A., Tabernero, J., Baselga, J., Tsao, M.S., Demichelis, F., Rubin, M.A., Janne, P.A., Daly, M.J., Nucera, C., Levine, R.L., Ebert, B.L., Gabriel, S., Rustgi, A.K., Antonescu, C.R., Ladanyi, M., Letai, A., Garraway, L.A., Loda, M., Beer, D.G., True, L.D., Okamoto, A., Pomeroy, S.L., Singer, S., Golub, T.R., Lander, E.S., Getz, G., Sellers, W.R., Meyerson, M., 2010. The landscape of somatic copy-number alteration across human cancers. *Nature* 463, 899-905.
- Bhatia, K., Spangler, G., Hamdy, N., Neri, A., Brubaker, G., Levin, A., Magrath, I., 1995. Mutations in the Coding Region of c-myc Occur Independently of Mutations in the Regulatory Regions and are Predominantly Associated with myc/Ig Translocation. In: Potter, M., Melchers, F. (Eds.), *Mechanisms in B-Cell Neoplasia 1994*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 389-398.

- Blackwell, T.K., Kretzner, L., Blackwood, E.M., Eisenman, R.N., Weintraub, H., 1990. Sequence-specific DNA binding by the c-Myc protein. *Science* 250, 1149-1151.
- Blackwell, T.K., Huang, J., Ma, A., Kretzner, L., Alt, F.W., Eisenman, R.N., Weintraub, H., 1993. Binding of myc proteins to canonical and noncanonical DNA sequences. *Mol Cell Biol* 13, 5216-5224.
- Blackwood, E.M., Eisenman, R.N., 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* 251, 1211-1217.
- Blancato, J., Singh, B., Liu, A., Liao, D.J., Dickson, R.B., 2004. Correlation of amplification and overexpression of the c-myc oncogene in high-grade breast cancer: FISH, in situ hybridisation and immunohistochemical analyses. *British Journal of Cancer* 90, 1612-1619.
- Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P., Darzacq, X., Zweckstetter, M., 2018. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nature Structural & Molecular Biology*.
- Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., Abraham, B.J., Afeyan, L.K., Guo, Y.E., Rimel, J.K., Fant, C.B., Schuijers, J., Lee, T.I., Taatjes, D.J., Young, R.A., 2018. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* 175, 1842-1855.e1816.
- Bradner, J.E., Hnisz, D., Young, R.A., 2017. Transcriptional Addiction in Cancer. *Cell* 168, 629-643.
- Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 34, 525-527.
- Bretones, G., Delgado, M.D., León, J., 2015. Myc and cell cycle control. *Biochim Biophys Acta* 1849, 506-516.
- Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., Peters, J.M., 2017. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* 544, 503-507.

- Campone, M., Noël, B., Couriaud, C., Grau, M., Guillemin, Y., Gautier, F., Gouraud, W., Charbonnel, C., Campion, L., Jézéquel, P., Braun, F., Barré, B., Coqueret, O., Barillé-Nion, S., Juin, P., 2011. c-Myc dependent expression of pro-apoptotic Bim renders HER2-overexpressing breast cancer cells dependent on anti-apoptotic Mcl-1. *Mol Cancer* 10, 110.
- 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., 2017. Super-Enhancers and Broad H3K4me3 Domains Form Complex Gene Regulatory Circuits Involving Chromatin Interactions. *Sci Rep* 7.
- Cheng, S.W., Davies, K.P., Yung, E., Beltran, R.J., Yu, J., Kalpana, G.V., 1999. c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. *Nat Genet* 22, 102-105.
- Chinwalla, A.T., Cook, L.L., Delehaunty, K.D., Fewell, G.A., Fulton, L.A., Fulton, R.S., Graves, T.A., Hillier, L.W., Mardis, E.R., McPherson, J.D., Miner, T.L., Nash, W.E., Nelson, J.O., Nhan, M.N., Pepin, K.H., Pohl, C.S., Ponce, T.C., Schultz, B., Thompson, J., Trevaskis, E., Waterston, R.H., Wendl, M.C., Wilson, R.K., Yang, S.-P., An, P., Berry, E., Birren, B., Bloom, T., Brown, D.G., Butler, J., Daly, M., David, R., Deri, J., Dodge, S., Foley, K., Gage, D., Gnerre, S., Holzer, T., Jaffe, D.B., Kamal, M., Karlsson, E.K., Kells, C., Kirby, A., Kulbokas, E.J., Lander, E.S., Landers, T., Leger, J.P., Levine, R., Lindblad-Toh, K., Mauceli, E., Mayer, J.H., McCarthy, M., Meldrim, J., Meldrim, J., Mesirov, J.P., Nicol, R., Nusbaum, C., Seaman, S., Sharpe, T., Sheridan, A., Singer, J.B., Santos, R., Spencer, B., Stange-Thomann, N., Vinson, J.P., Wade, C.M., Wierzbowski, J., Wyman, D., Zody, M.C., Birney, E., Goldman, N., Kasprzyk, A., Mongin, E., Rust, A.G., Slater, G., Stabenau, A., Ureta-Vidal, A., Whelan, S., Ainscough, R., Attwood, J., Bailey, J., Barlow, K., Beck, S., Burton, J., Clamp, M., Clee, C., Coulson, A., Cuff, J., Curwen, V., Cutts, T., Davies, J., Eyra, E., Grafham, D., Gregory, S., Hubbard, T., Hunt, A., Jones, M., Joy, A., Leonard, S., Lloyd, C., Matthews, L., McLaren, S., McLay, K., Meredith, B., Mullikin, J.C., Ning, Z., Oliver, K., Overton-

Larty, E., Plumb, R., Potter, S., Quail, M., Rogers, J., Scott, C., Searle, S., Shownkeen, R., Sims, S., Wall, M., West, A.P., Willey, D., Williams, S., Abril, J.F., Guigó, R., Parra, G., Agarwal, P., Agarwala, R., Church, D.M., Hlavina, W., Maglott, D.R., Sapojnikov, V., Alexandersson, M., Pachter, L., Antonarakis, S.E., Dermitzakis, E.T., Reymond, A., Ucla, C., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Hsu, F., Karolchik, D., Kent, W.J., Roskin, K.M., Schwartz, M.S., Sugnet, C., Weber, R.J., Bork, P., Letunic, I., Suyama, M., Torrents, D., Zdobnov, E.M., Botcherby, M., Brown, S.D., Campbell, R.D., Jackson, I., Bray, N., Couronne, O., Dubchak, I., Poliakov, A., Rubin, E.M., Brent, M.R., Flicek, P., Keibler, E., Korf, I., Batalov, S., Bult, C., Frankel, W.N., Carninci, P., Hayashizaki, Y., Kawai, J., Okazaki, Y., Cawley, S., Kulp, D., Wheeler, R., Chiaromonte, F., Collins, F.S., Felsenfeld, A., Guyer, M., Peterson, J., Wetterstrand, K., Copley, R.R., Mott, R., Dewey, C., Dickens, N.J., Emes, R.D., Goodstadt, L., Ponting, C.P., Winter, E., Dunn, D.M., von Niederhausern, A.C., Weiss, R.B., Eddy, S.R., Johnson, L.S., Jones, T.A., Elnitski, L., Kolbe, D.L., Eswara, P., Miller, W., O'Connor, M.J., Schwartz, S., Gibbs, R.A., Muzny, D.M., Glusman, G., Smit, A., Green, E.D., Hardison, R.C., Yang, S., Haussler, D., Hua, A., Roe, B.A., Kucherlapati, R.S., Montgomery, K.T., Li, J., Li, M., Lucas, S., Ma, B., McCombie, W.R., Morgan, M., Pevzner, P., Tesler, G., Schultz, J., Smith, D.R., Tromp, J., Worley, K.C., Lander, E.S., Abril, J.F., Agarwal, P., Alexandersson, M., Antonarakis, S.E., Baertsch, R., Berry, E., Birney, E., Bork, P., Bray, N., Brent, M.R., Brown, D.G., Butler, J., Bult, C., Chiaromonte, F., Chinwalla, A.T., Church, D.M., Clamp, M., Collins, F.S., Copley, R.R., Couronne, O., Cawley, S., Cuff, J., Curwen, V., Cutts, T., Daly, M., Dermitzakis, E.T., Dewey, C., Mouse Genome Sequencing, C., Genome Sequencing, C., Whitehead Institute, M.I.T.C.f.G.R., European Bioinformatics, I., Wellcome Trust Sanger, I., Research Group in Biomedical, I., Bioinformatics, National Center for Biotechnology, I., Department of, M., Division of Medical, G., Center for Biomolecular, S., Engineering, Embl, Consortium, U.M.M.S., Lawrence Berkeley National, L.,

Department of Computer, S., School of Computer, S., The Jackson, L., Laboratory for Genome, E., Affymetrix, I., Departments of, S., Health Evaluation, S., National Human Genome Research, I., Wellcome Trust Centre for Human, G., Department of Electrical, E., Department of Human, A., Genetics, Department of Human, G., Howard Hughes Medical, I., Department of, G., Departments of, B., Molecular, B., Computer, S., Engineering, Department of Computer, S., Engineering, Baylor College of, M., The Institute for Systems, B., Department of, B., Molecular, B., Howard Hughes Medical, I., Department of, C., Biochemistry, Departments of, G., Medicine, Harvard-Partners Center for, G., Genomics, Department of, S., Institute, U.D.J.G., Cold Spring Harbor, L., Wellcome, T., Max Planck Institute for Molecular, G., Genome Therapeutics, C., Bioinformatics Solutions, I., Department of, M., Human, G., Department of, B., Members of the Mouse Genome Analysis, G., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520-562.

Chipumuro, E., Marco, E., Christensen, Camilla L., Kwiatkowski, N., Zhang, T., Hatheway, Clark M., Abraham, Brian J., Sharma, B., Yeung, C., Altabef, A., Perez-Atayde, A., Wong, K.-K., Yuan, G.-C., Gray, Nathanael S., Young, Richard A., George, Rani E., 2014. CDK7 Inhibition Suppresses Super-Enhancer-Linked Oncogenic Transcription in MYCN-Driven Cancer. *Cell* 159, 1126-1139.

Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., Cisse, I.I., 2018. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* 361, 412-415.

Christensen, Camilla L., Kwiatkowski, N., Abraham, Brian J., Carretero, J., Al-Shahrour, F., Zhang, T., Chipumuro, E., Herter-Sprue, Grit S., Akbay, Esra A., Altabef, A., Zhang, J., Shimamura, T., Capelletti, M., Reibel, Jakob B., Cavanaugh, Jillian D., Gao, P., Liu, Y., Michaelsen, Signe R., Poulsen, Hans S., Aref, Amir R., Barbie, David A., Bradner, James E., George, Rani E., Gray, Nathanael S., Young, Richard A., Wong, K.-K., 2014. Targeting Transcriptional Addictions in Small Cell Lung Cancer with a Covalent CDK7 Inhibitor. *Cancer Cell* 26, 909-922.

- Chroma, K., Mistrik, M., Moudry, P., Gursky, J., Liptay, M., Strauss, R., Skrott, Z., Vrtel, R., Bartkova, J., Kramara, J., Bartek, J., 2016. Tumors overexpressing RNF168 show altered DNA repair and responses to genotoxic treatments, genomic instability and resistance to proteotoxic stress. *Oncogene* 36, 2405.
- Clegg, N.J., Couto, S.S., Wongvipat, J., Hieronymus, H., Carver, B.S., Taylor, B.S., Ellwood-Yen, K., Gerald, W.L., Sander, C., Sawyers, C.L., 2011. MYC Cooperates with AKT in Prostate Tumorigenesis and Alters Sensitivity to mTOR Inhibitors. *PLOS ONE* 6, e17449.
- Cole, K.A., Huggins, J., Laquaglia, M., Hulderman, C.E., Russell, M.R., Bosse, K., Diskin, S.J., Attiyeh, E.F., Sennett, R., Norris, G., Laudenslager, M., Wood, A.C., Mayes, P.A., Jagannathan, J., Winter, C., Mosse, Y.P., Maris, J.M., 2011. RNAi screen of the protein kinome identifies checkpoint kinase 1 (CHK1) as a therapeutic target in neuroblastoma. *Proceedings of the National Academy of Sciences of the United States of America* 108, 3336-3341.
- Cowling, V.H., Cole, M.D., 2007. The Myc transactivation domain promotes global phosphorylation of the RNA polymerase II carboxy-terminal domain independently of direct DNA binding. *Mol Cell Biol* 27, 2059-2073.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A., Jaenisch, R., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences* 107, 21931.
- Dang, C.V., 2013. MYC, metabolism, cell growth, and tumorigenesis. *Cold Spring Harb Perspect Med* 3.
- de Pretis, S., Kress, T.R., Morelli, M.J., Sabò, A., Locarno, C., Verrecchia, A., Doni, M., Campaner, S., Amati, B., Pelizzola, M., 2017. Integrative analysis of RNA polymerase II and transcriptional dynamics upon MYC activation. *Genome research* 27, 1658-1664.
- Dekker, J., Rippe, K., Dekker, M., Kleckner, N., 2002. Capturing chromosome conformation. *Science* 295, 1306-1311.

- Dermawan, J.K.T., Hitomi, M., Silver, D.J., Wu, Q., Sandlesh, P., Sloan, A.E., Purmal, A.A., Gurova, K.V., Rich, J.N., Lathia, J.D., Stark, G.R., Venere, M., 2016. Pharmacological Targeting of the Histone Chaperone Complex FACT Preferentially Eliminates Glioblastoma Stem Cells and Prolongs Survival in Preclinical Models. *Cancer research* 76, 2432.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., Ren, B., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376-380.
- Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., Diao, Y., Liang, J., Zhao, H., Lobanenko, V.V., Ecker, J.R., Thomson, J.A., Ren, B., 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331.
- 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., Green, R.D., Dekker, J., 2006. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research* 16, 1299-1309.
- Drier, Y., Cotton, M.J., Williamson, K.E., Gillespie, S.M., Ryan, R.J.H., Kluk, M.J., Carey, C.D., Rodig, S.J., Sholl, L.M., Afrogheh, A.H., Faquin, W.C., Queimado, L., Qi, J., Wick, M.J., El-Naggar, A.K., Bradner, J.E., Moskaluk, C.A., Aster, J.C., Knoechel, B., Bernstein, B.E., 2016. An oncogenic MYB feedback loop drives alternate cell fates in adenoid cystic carcinoma. *Nature Genetics* 48, 265-272.
- Dryden, N.H., Broome, L.R., Dudbridge, F., Johnson, N., Orr, N., Schoenfelder, S., Nagano, T., Andrews, S., Wingett, S., Kozarewa, I., Assiotis, I., Fenwick, K., Maguire, S.L., Campbell, J., Natrajan, R., Lambros, M., Perrakis, E., Ashworth, A., Fraser, P., Fletcher, O., 2014. Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. *Genome research* 24, 1854-1868.
- Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., Aiden, E.L., 2016. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Systems* 3, 95-98.

- Eilers, M., Eisenman, R.N., 2008. Myc's broad reach. *Genes Dev* 22, 2755-2766.
- Eliades, P., Abraham, B.J., Ji, Z., Miller, D.M., Christensen, C.L., Kwiatkowski, N., Kumar, R., Njauw, C.N., Taylor, M., Miao, B., Zhang, T., Wong, K.-K., Gray, N.S., Young, R.A., Tsao, H., 2018. High MITF Expression Is Associated with Super-Enhancers and Suppressed by CDK7 Inhibition in Melanoma. *Journal of Investigative Dermatology* 138, 1582-1590.
- Encode Project Consortium, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74.
- Evan, G.I., Wyllie, A.H., Gilbert, C.S., Littlewood, T.D., Land, H., Brooks, M., Waters, C.M., Penn, L.Z., Hancock, D.C., 1992. Induction of apoptosis in fibroblasts by c-myc protein. *Cell* 69, 119-128.
- 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., Chew, E.G., Huang, P.Y., Welboren, W.J., Han, Y., Ooi, H.S., Ariyaratne, P.N., Vega, V.B., Luo, Y., Tan, P.Y., Choy, P.Y., Wansa, K.D., Zhao, B., Lim, K.S., Leow, S.C., Yow, J.S., Joseph, R., Li, H., Desai, K.V., Thomsen, J.S., Lee, Y.K., Karuturi, R.K., Herve, T., Bourque, G., Stunnenberg, H.G., Ruan, X., Cacheux-Rataboul, V., Sung, W.K., Liu, E.T., Wei, C.L., Cheung, E., Ruan, Y., 2009. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* 462, 58-64.
- Gasparian Alexander, V., Burkhart Catherine, A., Purmal Andrei, A., Brodsky, L., Pal, M., Saranadasa, M., Bosykh Dmitry, A., Commane, M., Guryanova Olga, A., Pal, S., Safina, A., Sviridov, S., Koman Igor, E., Veith, J., Komar Anton, A., Gudkov Andrei, V., Gurova Katerina, V., 2011. Curaxins: Anticancer Compounds That Simultaneously Suppress NF- κ B and Activate p53 by Targeting FACT. *Science Translational Medicine* 3, 95ra74-95ra74.
- Gillies, S.D., Morrison, S.L., Oi, V.T., Tonegawa, S., 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33, 717-728.

- Goga, A., Yang, D., Tward, A.D., Morgan, D.O., Bishop, J.M., 2007. Inhibition of CDK1 as a potential therapy for tumors over-expressing MYC. *Nature Medicine* 13, 820-827.
- Greenwald, W.W., Li, H., Smith, E.N., Benaglio, P., Nariiai, N., Frazer, K.A., 2017. Pgltools: a genomic arithmetic tool suite for manipulation of Hi-C peak and other chromatin interaction data. *BMC Bioinformatics* 18, 207.
- Gryder, B.E., Pomella, S., Sayers, C., Wu, X.S., Song, Y., Chiarella, A.M., Bagchi, S., Chou, H.-C., Sinniah, R.S., Walton, A., Wen, X., Rota, R., Hathaway, N.A., Zhao, K., Chen, J., Vakoc, C.R., Shern, J.F., Stanton, B.Z., Khan, J., 2019. Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma. *Nature Genetics* 51, 1714-1722.
- Gu, B., Swigut, T., Spencley, A., Bauer, M.R., Chung, M., Meyer, T., Wysocka, J., 2018. Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* 359, 1050-1055.
- Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L., Amati, B., 2006. Myc-binding-site recognition in the human genome is determined by chromatin context. *Nature cell biology* 8, 764-770.
- Guo, J., Parise, R.A., Joseph, E., Egorin, M.J., Lazo, J.S., Prochownik, E.V., Eiseman, J.L., 2009. Efficacy, pharmacokinetics, tissue distribution, and metabolism of the Myc-Max disruptor, 10058-F4 [Z,E]-5-[4-ethylbenzylidene]-2-thioxothiazolidin-4-one, in mice. *Cancer Chemother Pharmacol* 63, 615-625.
- Guo, J., Li, T., Schipper, J., Nilson, K.A., Fordjour, F.K., Cooper, J.J., Gordân, R., Price, D.H., 2014. Sequence specificity incompletely defines the genome-wide occupancy of Myc. *Genome Biology* 15, 482.
- Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., Ching, K.A., Antosiewicz-Bourget, J.E., Liu, H., Zhang, X., Green, R.D., Lobanenko, V.V., Stewart, R., Thomson, J.A., Crawford, G.E., Kellis, M., Ren, B., 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108.

- Heinz, S., Texari, L., Hayes, M.G.B., Urbanowski, M., Chang, M.W., Givarkes, N., Rialdi, A., White, K.M., Albrecht, R.A., Pache, L., Marazzi, I., García-Sastre, A., Shaw, M.L., Benner, C., 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell* 174, 1522-1536.e1522.
- Hermeking, H., Eick, D., 1994. Mediation of c-Myc-induced apoptosis by p53. *Science* 265, 2091-2093.
- Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., Sharp, P.A., 2017. A phase separation model predicts key features of transcriptional control. *Cell* 169, 13-23.
- Hnisz, D., Abraham, Brian J., Lee, Tong I., Lau, A., Saint-André, V., Sigova, Alla A., Hoke, Heather A., Young, Richard A., 2013. Super-Enhancers in the Control of Cell Identity and Disease. *Cell* 155, 934-947.
- Hoffman, B., Liebermann, D.A., 2008. Apoptotic signaling by c-MYC. *Oncogene* 27, 6462-6472.
- Höglund, A., Nilsson, L.M., Muralidharan, S.V., Hasvold, L.A., Merta, P., Rudelius, M., Nikolova, V., Keller, U., Nilsson, J.A., 2011. Therapeutic implications for the induced levels of Chk1 in Myc-expressing cancer cells. *Clin Cancer Res* 17, 7067-7079.
- Horiuchi, D., Kusdra, L., Huskey, N.E., Chandriani, S., Lenburg, M.E., Gonzalez-Angulo, A.M., Creasman, K.J., Bazarov, A.V., Smyth, J.W., Davis, S.E., Yaswen, P., Mills, G.B., Esserman, L.J., Goga, A., 2012. MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. *The Journal of experimental medicine* 209, 679-696.
- Hsieh, C.L., Fei, T., Chen, Y., Li, T., Gao, Y., Wang, X., Sun, T., Sweeney, C.J., Lee, G.S., Chen, S., Balk, S.P., Liu, X.S., Brown, M., Kantoff, P.W., 2014. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. *Proceedings of the National Academy of Sciences of the United States of America* 111, 7319-7324.
- Hsu, F., Kent, W.J., Clawson, H., Kuhn, R.M., Diekhans, M., Haussler, D., 2006. The UCSC Known Genes. *Bioinformatics (Oxford, England)* 22, 1036-1046.

- Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S.H., Xu, J., Yuan, G.-C., 2018. Dissecting super-enhancer hierarchy based on chromatin interactions. *Nature Communications* 9, 943.
- Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T.V., Zon, L.I., Yuan, G.-C., Orkin, S.H., Xu, J., 2016. Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. *Developmental Cell* 36, 9-23.
- Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., Higgs, D.R., 2014. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nature Genetics* 46, 205-212.
- Hydbring, P., Bahram, F., Su, Y., Tronnorsjö, S., Högstrand, K., von der Lehr, N., Sharifi, H.R., Lilischkis, R., Hein, N., Wu, S., Vervoorts, J., Henriksson, M., Grandien, A., Lüscher, B., Larsson, L.-G., 2010. Phosphorylation by Cdk2 is required for Myc to repress Ras-induced senescence in cotransformation. *Proceedings of the National Academy of Sciences* 107, 58.
- Hyle, J., Zhang, Y., Wright, S., Xu, B., Shao, Y., Easton, J., Tian, L., Feng, R., Xu, P., Li, C., 2019. Acute depletion of CTCF directly affects MYC regulation through loss of enhancer–promoter looping. *Nucleic Acids Research* 47, 6699-6713.
- Isoda, T., Moore, A.J., He, Z., Chandra, V., Aida, M., Denholtz, M., Piet van Hamburg, J., Fisch, K.M., Chang, A.N., Fahl, S.P., Wiest, D.L., Murre, C., 2017. Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell* 171, 103-119.e118.
- Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Várnai, C., Thiecke, M.J., Burden, F., Farrow, S., Cutler, A.J., Rehnström, K., Downes, K., Grassi, L., Kostadima, M., Freire-Pritchett, P., Wang, F., Martens, J.H., Kim, B., Sharifi, N., Janssen-Megens, E.M., Yaspo, M.-L., Linser, M., Kovacsovics, A., Clarke, L., Richardson, D., Datta, A., Flicek, P., Stunnenberg, H.G., Todd, J.A., Zerbino, D.R., Stegle, O., Ouwehand,

- W.H., Frontini, M., Wallace, C., Spivakov, M., Fraser, P., 2016. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell* 167, 1369-1384.e1319.
- Jiang, Y.Y., Lin, D.C., Mayakonda, A., Hazawa, M., Ding, L.W., Chien, W.W., Xu, L., Chen, Y., Xiao, J.F., Senapedis, W., Baloglu, E., Kanojia, D., Shang, L., Xu, X., Yang, H., Tyner, J.W., Wang, M.R., Koeffler, H.P., 2017. Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma. *Gut* 66, 1358-1368.
- 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., Taatjes, D.J., Dekker, J., Young, R.A., 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430.
- Kalkat, M., Resetca, D., Lourenco, C., Chan, P.-K., Wei, Y., Shiah, Y.-J., Vitkin, N., Tong, Y., Sunnerhagen, M., Done, S.J., Boutros, P.C., Raught, B., Penn, L.Z., 2018. MYC Protein Interactome Profiling Reveals Functionally Distinct Regions that Cooperate to Drive Tumorigenesis. *Molecular Cell* 72, 836-848.e837.
- Kang, J., Sergio, C.M., Sutherland, R.L., Musgrove, E.A., 2014. Targeting cyclin-dependent kinase 1 (CDK1) but not CDK4/6 or CDK2 is selectively lethal to MYC-dependent human breast cancer cells. *BMC Cancer* 14, 32-32.
- Kantidze, O.L., Luzhin, A.V., Nizovtseva, E.V., Safina, A., Valieva, M.E., Golov, A.K., Velichko, A.K., Lyubitelev, A.V., Feofanov, A.V., Gurova, K.V., Studitsky, V.M., Razin, S.V., 2019. The anti-cancer drugs curaxins target spatial genome organization. *Nature Communications* 10, 1441.
- Kasibhatla, S., Beere, H.M., Brunner, T., Echeverri, F., Green, D.R., 2000. A 'non-canonical' DNA-binding element mediates the response of the Fas-ligand promoter to c-Myc. *Current Biology* 10, 1205-1208.
- Kauffmann-Zeh, A., Rodriguez-Viciana, P., Ulrich, E., Gilbert, C., Coffey, P., Downward, J., Evan, G., 1997. Suppression of c-Myc-induced apoptosis by Ras signalling through PI(3)K and PKB. *Nature* 385, 544-548.

- Ke, L., Zhou, H., Wang, C., Xiong, G., Xiang, Y., Ling, Y., Khabir, A., Tsao, G.S., Zeng, Y., Zeng, M., Busson, P., Kieff, E., Guo, X., Zhao, B., 2017. Nasopharyngeal carcinoma super-enhancer–driven ETV6 correlates with prognosis. *Proceedings of the National Academy of Sciences* 114, 9683.
- Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The human genome browser at UCSC. *Genome research* 12, 996-1006.
- Kessler Jessica, D., Kahle Kristopher, T., Sun, T., Meerbrey Kristen, L., Schlabach Michael, R., Schmitt Earlene, M., Skinner Samuel, O., Xu, Q., Li Mamie, Z., Hartman Zachary, C., Rao, M., Yu, P., Dominguez-Vidana, R., Liang Anthony, C., Solimini Nicole, L., Bernardi Ronald, J., Yu, B., Hsu, T., Golding, I., Luo, J., Osborne, C.K., Creighton Chad, J., Hilsenbeck Susan, G., Schiff, R., Shaw Chad, A., Elledge Stephen, J., Westbrook Thomas, F., 2012. A SUMOylation-Dependent Transcriptional Subprogram Is Required for Myc-Driven Tumorigenesis. *Science* 335, 348-353.
- Khattar, E., Tergaonkar, V., 2017. Transcriptional Regulation of Telomerase Reverse Transcriptase (TERT) by MYC. *Frontiers in cell and developmental biology* 5, 1-1.
- Kieffer-Kwon, K.-R., Nimura, K., Rao, S.S.P., Xu, J., Jung, S., Pekowska, A., Dose, M., Stevens, E., Mathe, E., Dong, P., Huang, S.-C., Ricci, M.A., Baranello, L., Zheng, Y., Tomassoni Ardori, F., Resch, W., Stavreva, D., Nelson, S., McAndrew, M., Casellas, A., Finn, E., Gregory, C., St. Hilaire, B.G., Johnson, S.M., Dubois, W., Cosma, M.P., Batchelor, E., Levens, D., Phair, R.D., Misteli, T., Tessarollo, L., Hager, G., Lakadamyali, M., Liu, Z., Floer, M., Shroff, H., Aiden, E.L., Casellas, R., 2017. Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. *Molecular Cell* 67, 566-578.e510.
- Kim, M., Neznanov, N., Wilfong, C.D., Fleyshman, D.I., Purmal, A.A., Haderski, G., Stanhope-Baker, P., Burkhart, C.A., Gurova, K.V., Gudkov, A.V., Skitzki, J.J., 2016. Preclinical Validation of a Single-

- Treatment Infusion Modality That Can Eradicate Extremity Melanomas. *Cancer research* 76, 6620.
- Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., Markenscoff-Papadimitriou, E., Kuhl, D., Bito, H., Worley, P.F., Kreiman, G., Greenberg, M.E., 2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182-187.
- Klattenhoff, Carla A., Scheuermann, Johanna C., Surface, Lauren E., Bradley, Robert K., Fields, Paul A., Steinhäuser, Matthew L., Ding, H., Butty, Vincent L., Torrey, L., Haas, S., Abo, R., Tabebordbar, M., Lee, Richard T., Burge, Christopher B., Boyer, Laurie A., 2013. Braveheart, a Long Noncoding RNA Required for Cardiovascular Lineage Commitment. *Cell* 152, 570-583.
- Koch, H.B., Zhang, R., Verdoodt, B., Bailey, A., Zhang, C.D., Yates, J.R., 3rd, Menssen, A., Hermeking, H., 2007. Large-scale identification of c-MYC-associated proteins using a combined TAP/MudPIT approach. *Cell cycle (Georgetown, Tex.)* 6, 205-217.
- Koh, C.M., Sabo, A., Guccione, E., 2016. Targeting MYC in cancer therapy: RNA processing offers new opportunities. *Bioessays* 38, 266-275.
- Kohl, N.E., Legouy, E., DePinho, R.A., Nisen, P.D., Smith, R.K., Gee, C.E., Alt, F.W., 1986. Human N-myc is closely related in organization and nucleotide sequence to c-myc. *Nature* 319, 73.
- Korkmaz, G., Lopes, R., Ugalde, A.P., Nevedomskaya, E., Han, R., Myacheva, K., Zwart, W., Elkon, R., Agami, R., 2016. Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nature Biotechnology* 34, 192-198.
- Kress, T.R., Sabo, A., Amati, B., 2015. MYC: connecting selective transcriptional control to global RNA production. *Nat Rev Cancer* 15, 593-607.
- Kulakovskiy, I.V., Vorontsov, I.E., Yevshin, I.S., Sharipov, R.N., Fedorova, A.D., Rumynskiy, E.I., Medvedeva, Y.A., Magana-Mora, A., Bajic, V.B., Papatsenko, D.A., Kolpakov, F.A., Makeev, V.J., 2018. HOCOMOCO: towards a complete collection of transcription factor

- binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Research* 46, D252-D259.
- Küppers, R., Dalla-Favera, R., 2001. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* 20, 5580.
- Kuzyk, A., Mai, S., 2014. c-MYC-Induced Genomic Instability. *Cold Spring Harbor Perspectives in Medicine* 4.
- Kwiatkowski, N., Zhang, T., Rahl, P.B., Abraham, B.J., Reddy, J., Ficarro, S.B., Dastur, A., Amzallag, A., Ramaswamy, S., Tesar, B., Jenkins, C.E., Hannett, N.M., McMillin, D., Sanda, T., Sim, T., Kim, N.D., Look, T., Mitsiades, C.S., Weng, A.P., Brown, J.R., Benes, C.H., Marto, J.A., Young, R.A., Gray, N.S., 2014. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* 511, 616-620.
- Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., Shiekhhattar, R., 2013. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* 494, 497-501.
- Lajoie, B.R., Dekker, J., Kaplan, N., 2015. The Hitchhiker's Guide to Hi-C Analysis: Practical guidelines. *Methods (San Diego, Calif.)* 72, 65-75.
- Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., Narlikar, G.J., 2017. Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* 547, 236-240.
- Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. eprint arXiv:1303.3997, arXiv:1303.3997.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* 25, 2078-2079.
- Li, K., Liu, Y., Cao, H., Zhang, Y., Gu, Z., Liu, X., Yu, A., Kaphle, P., Dickerson, K.E., Ni, M., Xu, J., 2020a. Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing. *Nature Communications* 11, 485.

- Li, S., Lin, P., Young, K.H., Kanagal-Shamanna, R., Yin, C.C., Medeiros, L.J., 2013. MYC/BCL2 Double-Hit High-Grade B-Cell Lymphoma. *Advances in Anatomic Pathology* 20, 315-326.
- Li, Y., Sun, X.-X., Qian, D.Z., Dai, M.-S., 2020b. Molecular Crosstalk Between MYC and HIF in Cancer. *Frontiers in Cell and Developmental Biology* 8.
- Liao, H., Ji, F., Geng, X., Xing, M., Li, W., Chen, Z., Shen, H., Ying, S., 2017. CDK1 promotes nascent DNA synthesis and induces resistance of cancer cells to DNA-damaging therapeutic agents. *Oncotarget* 8, 90662-90673.
- Liao, Y., Smyth, G.K., Shi, W., 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research* 47, e47-e47.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S., Dekker, J., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-293.
- Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., Young, R.A., 2012. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* 151, 56-67.
- Liu, Q., Wang, M., Hu, Y., Xing, H., Chen, X., Zhang, Y., Zhu, P., 2014. Significance of CD71 expression by flow cytometry in diagnosis of acute leukemia. *Leukemia & Lymphoma* 55, 892-898.
- Liu, X., Tesfai, J., Evrard, Y.A., Dent, S.Y.R., Martinez, E., 2003. c-Myc Transformation Domain Recruits the Human STAGA Complex and Requires TRRAP and GCN5 Acetylase Activity for Transcription Activation*. *Journal of Biological Chemistry* 278, 20405-20412.
- Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods (San Diego, Calif.)* 25, 402-408.
- Lorenzin, F., Benary, U., Baluapuri, A., Walz, S., Jung, L.A., von Eyss, B., Kisker, C., Wolf, J., Eilers, M., Wolf, E., 2016. Different promoter

- affinities account for specificity in MYC-dependent gene regulation. *eLife* 5, e15161.
- 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.
- Loven, 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.
- Lovén, J., Orlando, David A., Sigova, Alla A., Lin, Charles Y., Rahl, Peter B., Burge, Christopher B., Levens, David L., Lee, Tong I., Young, Richard A., 2012. Revisiting Global Gene Expression Analysis. *Cell* 151, 476-482.
- Lu, H., Yu, D., Hansen, A.S., Ganguly, S., Liu, R., Heckert, A., Darzacq, X., Zhou, Q., 2018. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* 558, 318-323.
- Lu, K., Liu, C., Liu, Y., Luo, A., Chen, J., Lei, Z., Kong, J., Xiao, X., Zhang, S., Wang, Y.-Z., Ma, L., Dou, S.-X., Wang, P.-Y., Li, M., Li, G., Li, W., Chen, P., 2021. Curaxin-Induced DNA Topology Alterations Trigger the Distinct Binding Response of CTCF and FACT at the Single-Molecule Level. *Biochemistry* 60, 494-499.
- Lun, A.T.L., Smyth, G.K., 2015. diffHic: a Bioconductor package to detect differential genomic interactions in Hi-C data. *BMC Bioinformatics* 16, 258.
- Mack, S.C., Pajtlér, K.W., Chavez, L., Okonechnikov, K., Bertrand, K.C., Wang, X., Erkek, S., Federation, A., Song, A., Lee, C., Wang, X., McDonald, L., Morrow, J.J., Saiakhova, A., Sin-Chan, P., Wu, Q., Michaelraj, K.A., Miller, T.E., Hubert, C.G., Ryzhova, M., Garzia, L., Donovan, L., Dombrowski, S., Factor, D.C., Luu, B., Valentim, C.L.L., Gimple, R.C., Morton, A., Kim, L., Prager, B.C., Lee, J.J.Y., Wu, X., Zuccaro, J., Thompson, Y., Holgado, B.L., Reimand, J., Ke, S.Q., Tropper, A., Lai, S., Vijayarajah, S., Doan, S., Mahadev, V., Miñan, A.F., Gröbner, S.N., Lienhard, M., Zapatka, M., Huang, Z., Aldape, K.D., Carcaboso, A.M., Houghton, P.J., Keir, S.T., Milde, T., Witt, H., Li, Y., Li, C.-J., Bian, X.-W., Jones, D.T.W., Scott, I., Singh,

- S.K., Huang, A., Dirks, P.B., Bouffet, E., Bradner, J.E., Ramaswamy, V., Jabado, N., Rutka, J.T., Northcott, P.A., Lupien, M., Lichter, P., Korshunov, A., Scacheri, P.C., Pfister, S.M., Kool, M., Taylor, M.D., Rich, J.N., 2017. Therapeutic targeting of ependymoma as informed by oncogenic enhancer profiling. *Nature* 553, 101.
- Madsen, J.G.S., Madsen, M.S., Rauch, A., Traynor, S., Van Hauwaert, E.L., Haakonsson, A.K., Javierre, B.M., Hyldahl, M., Fraser, P., Mandrup, S., 2020. Highly interconnected enhancer communities control lineage-determining genes in human mesenchymal stem cells. *Nature Genetics* 52, 1227-1238.
- Malynn, B.A., de Alboran, I.M., O'Hagan, R.C., Bronson, R., Davidson, L., DePinho, R.A., Alt, F.W., 2000. N-myc can functionally replace c-myc in murine development, cellular growth, and differentiation. *Genes & Development* 14, 1390-1399.
- Martinato, F., Cesaroni, M., Amati, B., Guccione, E., 2008. Analysis of Myc-induced histone modifications on target chromatin. *PloS one* 3, e3650-e3650.
- McKeown, M.R., Bradner, J.E., 2014. Therapeutic strategies to inhibit MYC. *Cold Spring Harb Perspect Med* 4.
- 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., Lovén, J., Fritz, C.C., Majeti, R., 2017. Super-Enhancer analysis defines novel epigenomic subtypes of non-APL AML including an RAR α dependency targetable by SY-1425, a potent and selective RAR α agonist. *Cancer discovery* 7, 1136-1153.
- Mifsud, B., Tavares-Cadete, F., Young, A.N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S.W., Andrews, S., Grey, W., Ewels, P.A., Herman, B., Happe, S., Higgs, A., LeProust, E., Follows, G.A., Fraser, P., Luscombe, N.M., Osborne, C.S., 2015. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nature Genetics* 47, 598-606.
- Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I.E., Males, M., Viales, R.R., Furlong, E.E.M., 2018. The degree of enhancer or

- promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes & Development* 32, 42-57.
- Moreau, P.R., Örd, T., Downes, N.L., Niskanen, H., Bouvy-Liivrand, M., Aavik, E., Ylä-Herttuala, S., Kaikkonen, M.U., 2018. Transcriptional Profiling of Hypoxia-Regulated Non-coding RNAs in Human Primary Endothelial Cells. *Frontiers in Cardiovascular Medicine* 5.
- Muhar, M., Ebert, A., Neumann, T., Umkehrer, C., Jude, J., Wieshofer, C., Rescheneder, P., Lipp, J.J., Herzog, V.A., Reichholf, B., Cisneros, D.A., Hoffmann, T., Schlapansky, M.F., Bhat, P., von Haeseler, A., Köcher, T., Obenauf, A.C., Popow, J., Ameres, S.L., Zuber, J., 2018. SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science* 360, 800.
- Murai, J., Zhang, H., Pongor, L., Tang, S.-W., Jo, U., Moribe, F., Ma, Y., Tomita, M., Pommier, Y., 2020. Chromatin Remodeling and Immediate Early Gene Activation by SLFN11 in Response to Replication Stress. *Cell Reports* 30, 4137-4151.e4136.
- Muthalagu, N., Junttila, M.R., Wiese, K.E., Wolf, E., Morton, J., Bauer, B., Evan, G.I., Eilers, M., Murphy, D.J., 2014. BIM is the primary mediator of MYC-induced apoptosis in multiple solid tissues. *Cell Rep* 8, 1347-1353.
- Narendra, V., Bulajić, M., Dekker, J., Mazzoni, E.O., Reinberg, D., 2016. CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes & Development* 30, 2657-2662.
- Natoli, G., Andrau, J.-C., 2012. Noncoding Transcription at Enhancers: General Principles and Functional Models. *Annual Review of Genetics* 46, 1-19.
- Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., Zhao, K., Levens, D., 2012. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* 151, 68-79.
- 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.e922.
- Nord, Alex S., Blow, Matthew J., Attanasio, C., Akiyama, Jennifer A., Holt, A., Hosseini, R., Phouanenvong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., Rubenstein, John L.R., Rubin, Edward M., Pennacchio, Len A., Visel, A., 2013. Rapid and Pervasive Changes in Genome-wide Enhancer Usage during Mammalian Development. *Cell* 155, 1521-1531.
- 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.
- Ooi, W.F., Xing, M., Xu, C., Yao, X., Ramlee, M.K., Lim, M.C., Cao, F., Lim, K., Babu, D., Poon, L.-F., Lin Suling, J., Qamra, A., Irwanto, A., Qu Zhengzhong, J., Nandi, T., Lee-Lim, A.P., Chan, Y.S., Tay, S.T., Lee, M.H., Davies, J.O.J., Wong, W.K., Soo, K.C., Chan, W.H., Ong, H.S., Chow, P., Wong, C.Y., Rha, S.Y., Liu, J., Hillmer, A.M., Hughes, J.R., Rozen, S., Teh, B.T., Fullwood, M.J., Li, S., Tan, P., 2016. Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. *Nature Communications* 7, 12983.
- Orlando, David A., Chen, Mei W., Brown, Victoria E., Solanki, S., Choi, Yoon J., Olson, Eric R., Fritz, Christian C., Bradner, James E., Guenther, Matthew G., 2014. Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. *Cell Reports* 9, 1163-1170.
- Osterwalder, M., Barozzi, I., Tissires, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Plajzer-Frick, I., Pickle, C.S., Kato, M., Garvin, T.H., Pham, Q.T., Harrington, A.N., Akiyama, J.A., Afzal, V., Lopez-Rios, J., Dickel, D.E., Visel, A., Pennacchio, L.A., 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* 554, 239-243.
- Pachano, T., Snchez-Gaya, V., Ealo, T., Mariner-Faul, M., Bleckwehl, T., Asenjo, H.G., Respuela, P., Cruz-Molina, S., Muoz-San Martn, M., Haro, E., van Ijcken, W.F.J., Landeira, D., Rada-Iglesias, A., 2021.

- Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nature Genetics* 53, 1036-1049.
- Pajic, A., Spitkovsky, D., Christoph, B., Kempkes, B., Schuhmacher, M., Staeger, M.S., Brielmeier, M., Ellwart, J., Kohlhuber, F., Bornkamm, G.W., Polack, A., Eick, D., 2000. Cell cycle activation by c-myc in a burkitt lymphoma model cell line. *Int J Cancer* 87, 787-793.
- Patel, B., Kang, Y., Cui, K., Litt, M., Riberio, M.S.J., Deng, C., Salz, T., Casada, S., Fu, X., Qiu, Y., Zhao, K., Huang, S., 2013. Aberrant TAL1 activation is mediated by an interchromosomal interaction in human T-cell acute lymphoblastic leukemia. *Leukemia* 28, 349.
- Pellacani, D., Bilenky, M., Kannan, N., Heravi-Moussavi, A., Knapp, David J.H.F., Gakkhar, S., Moksa, M., Carles, A., Moore, R., Mungall, Andrew J., Marra, Marco A., Jones, Steven J.M., Aparicio, S., Hirst, M., Eaves, Connie J., 2016. Analysis of Normal Human Mammary Epigenomes Reveals Cell-Specific Active Enhancer States and Associated Transcription Factor Networks. *Cell Reports* 17, 2060-2074.
- Pimentel, H., Bray, N.L., Puente, S., Melsted, P., Pachter, L., 2017. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods* 14, 687-690.
- Poli, V., Fagnocchi, L., Fasciani, A., Cherubini, A., Mazzoleni, S., Ferrillo, S., Miluzio, A., Gaudio, G., Vaira, V., Turdo, A., Gaggianesi, M., Chinnici, A., Lipari, E., Bicciato, S., Bosari, S., Todaro, M., Zippo, A., 2018. MYC-driven epigenetic reprogramming favors the onset of tumorigenesis by inducing a stem cell-like state. *Nature Communications* 9, 1024.
- Pott, S., Lieb, J.D., 2015. What are super-enhancers? *Nat Genet* 47, 8-12.
- Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* 26, 841-842.
- Quinodoz, S.A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J.M., Detmar, E., Lai, M.M., Shishkin, A.A., Bhat, P., Takei, Y., Trinh, V., Aznauryan, E., Russell, P., Cheng, C., Jovanovic, M., Chow, A., Cai, L., McDonel, P., Garber, M., Guttman, M., 2018. Higher-Order Inter-chromosomal

- Hubs Shape 3D Genome Organization in the Nucleus. *Cell* 174, 744-757.e724.
- Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., Young, R.A., 2010. c-Myc regulates transcriptional pause release. *Cell* 141, 432-445.
- Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., Manke, T., 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Research* 42, W187-W191.
- Rao, Suhas S.P., Huntley, Miriam H., Durand, Neva C., Stamenova, Elena K., Bochkov, Ivan D., Robinson, James T., Sanborn, Adrian L., Machol, I., Omer, Arina D., Lander, Eric S., Aiden, Erez L., 2014. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159, 1665-1680.
- Rao, S.S.P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., Huang, X., Shamim, M.S., Shin, J., Turner, D., Ye, Z., Omer, A.D., Robinson, J.T., Schlick, T., Bernstein, B.E., Casellas, R., Lander, E.S., Aiden, E.L., 2017. Cohesin Loss Eliminates All Loop Domains. *Cell* 171, 305-320.e324.
- Rodríguez-Carballo, E., Gámez, B., Ventura, F., 2016. p38 MAPK Signaling in Osteoblast Differentiation. *Frontiers in Cell and Developmental Biology* 4.
- Rosencrance, C.D., Ammouri, H.N., Yu, Q., Ge, T., Rendleman, E.J., Marshall, S.A., Eagen, K.P., 2020. Chromatin Hyperacetylation Impacts Chromosome Folding by Forming a Nuclear Subcompartment. *Molecular cell* 78, 112-126.e112.
- Russell, M.R., Levin, K., Rader, J., Belcastro, L., Li, Y., Martinez, D., Pawel, B., Shumway, S.D., Maris, J.M., Cole, K.A., 2013. Combination therapy targeting the Chk1 and Wee1 kinases shows therapeutic efficacy in neuroblastoma. *Cancer research* 73, 776-784.
- 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., Li, C.H., Guo, Y.E., Day, D.S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T.I., Cisse, I.I., Roeder, R.G., Sharp, P.A., Chakraborty, A.K.,

- Young, R.A., 2018. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361, 379-+.
- Sabo, A., Kress, T.R., Pelizzola, M., de Pretis, S., Gorski, M.M., Tesi, A., Morelli, M.J., Bora, P., Doni, M., Verrecchia, A., Tonelli, C., Faga, G., Bianchi, V., Ronchi, A., Low, D., Muller, H., Guccione, E., Campaner, S., Amati, B., 2014. Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature* 511, 488-492.
- Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., Geeting, K.P., Gnirke, A., Melnikov, A., McKenna, D., Stamenova, E.K., Lander, E.S., Aiden, E.L., 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.
- Santamaría, D., Barrière, C., Cerqueira, A., Hunt, S., Tardy, C., Newton, K., Cáceres, J.F., Dubus, P., Malumbres, M., Barbacid, M., 2007. Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* 448, 811-815.
- Sarantopoulos, J., Mahalingam, D., Sharma, N., Iyer, R.V., Ma, W.W., Ahluwalia, M.S., Johnson, S., Purmal, A., Shpigotskaya, P., Hards, A., Leonov, A., Gurova, K., Gudkov, A., Zakurdaeva, K., Miller, L.L., Dowlati, A., 2020. Results of a completed phase I trial of CBL0137 administered intravenously (IV) to patients (Pts) with advanced solid tumors. *Journal of Clinical Oncology* 38, 3583-3583.
- Sauerwald, N., Singhal, A., Kingsford, C., 2019. Analysis of the structural variability of topologically associated domains as revealed by Hi-C. *NAR Genomics and Bioinformatics* 2.
- Sava, G.P., Fan, H., Coombes, R.C., Buluwela, L., Ali, S., 2020. CDK7 inhibitors as anticancer drugs. *Cancer Metastasis Rev* 39, 805-823.
- Schaub, F.X., Dhankani, V., Berger, A.C., Trivedi, M., Richardson, A.B., Shaw, R., Zhao, W., Zhang, X., Ventura, A., Liu, Y., Ayer, D.E., Hurlin, P.J., Cherniack, A.D., Eisenman, R.N., Bernard, B., Grandori, C., 2018. Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas. *Cell Systems* 6, 282-300.e282.

- 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., 2016. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep* 17, 2042-2059.
- Schmitz, R., Hansmann, M.-L., Bohle, V., Martin-Subero, J.I., Hartmann, S., Mechttersheimer, G., Klapper, W., Vater, I., Giefing, M., Gesk, S., Stanelle, J., Siebert, R., Küppers, R., 2009. TNFAIP3 (A20) is a tumor suppressor gene in Hodgkin lymphoma and primary mediastinal B cell lymphoma. *The Journal of Experimental Medicine* 206, 981.
- Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L., Spitz, F., 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 551, 51-56.
- 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.
- Shin, H.Y., Willi, M., Yoo, K.H., Zeng, X., Wang, C., Metser, G., Hennighausen, L., 2016. Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nature Genetics* 48, 904-911.
- Shorstova, T., Foulkes, W.D., Witcher, M., 2021. Achieving clinical success with BET inhibitors as anti-cancer agents. *British Journal of Cancer* 124, 1478-1490.
- Shrivastava, A., Yu, J., Artandi, S., Calame, K., 1996. YY1 and c-Myc associate in vivo in a manner that depends on c-Myc levels. *Proceedings of the National Academy of Sciences* 93, 10638.
- 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.
- Smith, Emily M., Lajoie, Bryan R., Jain, G., Dekker, J., 2016. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *The American Journal of Human Genetics* 98, 185-201.

- Soucek, L., Jucker, R., Panacchia, L., Ricordy, R., Tatò, F., Nasi, S., 2002. Omomyc, a potential Myc dominant negative, enhances Myc-induced apoptosis. *Cancer research* 62, 3507-3510.
- Soufi, A., Donahue, G., Zaret, K.S., 2012. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151, 994-1004.
- Stansfield, J.C., Cresswell, K.G., Vladimirov, V.I., Dozmorov, M.G., 2018. HiCcompare: an R-package for joint normalization and comparison of HI-C datasets. *BMC Bioinformatics* 19, 279.
- Stanton, B., Khan, J., Gryder, B., 2018. Absolute Quantification of Architecture (AQuA-HiChIP) Enables Measurement of Differential Chromatin Interactions.
- Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., Karpen, G.H., 2017. Phase separation drives heterochromatin domain formation. *Nature* 547, 241-245.
- Struhl, K., 2007. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nature Structural & Molecular Biology* 14, 103-105.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* 102, 15545.
- Szabo, Q., Bantignies, F., Cavalli, G., 2019. Principles of genome folding into topologically associating domains. *Science Advances* 5, eaaw1668.
- Thng, D.K.H., Toh, T.B., Chow, E.K.-H., 2021. Capitalizing on Synthetic Lethality of MYC to Treat Cancer in the Digital Age. *Trends in Pharmacological Sciences* 42, 166-182.
- Thongjuea, S., Stadhouders, R., Grosveld, F.G., Soler, E., Lenhard, B., 2013. r3Cseq: an R/Bioconductor package for the discovery of long-range genomic interactions from chromosome conformation capture and next-generation sequencing data. *Nucleic Acids Research* 41, e132-e132.
- Toyoshima, M., Howie, H.L., Imakura, M., Walsh, R.M., Annis, J.E., Chang, A.N., Frazier, J., Chau, B.N., Loboda, A., Linsley, P.S., Cleary, M.A.,

- Park, J.R., Grandori, C., 2012. Functional genomics identifies therapeutic targets for MYC-driven cancer. *Proceedings of the National Academy of Sciences* 109, 9545.
- Troppan, K., Hofer, S., Wenzl, K., Lassnig, M., Pursche, B., Steinbauer, E., Wiltgen, M., Zulus, B., Renner, W., Beham-Schmid, C., Deutsch, A., Neumeister, P., 2015. Frequent Down Regulation of the Tumor Suppressor Gene A20 in Multiple Myeloma. *PLOS ONE* 10, e0123922.
- van de Werken, H.J.G., Landan, G., Holwerda, S.J.B., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Öz, Y., Bouwman, B.A.M., Verstegen, M.J.A.M., de Wit, E., Tanay, A., de Laat, W., 2012. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nature Methods* 9, 969.
- Vella, P., Barozzi, I., Cuomo, A., Bonaldi, T., Pasini, D., 2012. Yin Yang 1 extends the Myc-related transcription factors network in embryonic stem cells. *Nucleic Acids Research* 40, 3403-3418.
- Vernon, E.G., Gaston, K., 2000. Myc and YY1 mediate activation of the Surf-1 promoter in response to serum growth factors. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1492, 172-179.
- Walz, S., Lorenzin, F., Morton, J., Wiese, K.E., von Eyss, B., Herold, S., Rycak, L., Dumay-Odelot, H., Karim, S., Bartkuhn, M., Roels, F., Wustefeld, T., Fischer, M., Teichmann, M., Zender, L., Wei, C.L., Sansom, O., Wolf, E., Eilers, M., 2014. Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. *Nature* 511, 483-487.
- Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., Rando, O.J., Birney, E., Myers, R.M., Noble, W.S., Snyder, M., Weng, Z., 2012. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome research* 22, 1798-1812.
- 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, Y., Zhang, T., Kwiatkowski, N., Abraham, Brian J., Lee, Tong I., Xie, S., Yuzugullu, H., Von, T., Li, H., Lin, Z., Stover, Daniel G., Lim, E., Wang, Zhigang C., Iglehart, J.D., Young, Richard A., Gray, Nathanael S., Zhao, Jean J., 2015. CDK7-Dependent Transcriptional Addiction in Triple-Negative Breast Cancer. *Cell* 163, 174-186.
- Weinhold, N., Jacobsen, A., Schultz, N., Sander, C., Lee, W., 2014. Genome-wide analysis of noncoding regulatory mutations in cancer. *Nature Genetics* 46, 1160-1165.
- Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., Guo, Y.E., Hnisz, D., Jaenisch, R., Bradner, J.E., Gray, N.S., Young, R.A., 2017. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* 171, 1573-1588.e1528.
- Whitfield, J.R., Beaulieu, M.-E., Soucek, L., 2017. Strategies to Inhibit Myc and Their Clinical Applicability. *Frontiers in Cell and Developmental Biology* 5, 10.
- Whitlock, N.C., Trostel, S.Y., Wilkinson, S., Terrigino, N.T., Hennigan, S.T., Lake, R., Carrabba, N.V., Atway, R., Walton, E.D., Gryder, B.E., Capaldo, B.J., Ye, H., Sowalsky, A.G., 2020. MEIS1 down-regulation by MYC mediates prostate cancer development through elevated HOXB13 expression and AR activity. *Oncogene* 39, 5663-5674.
- Whyte, Warren A., Orlando, David A., Hnisz, D., Abraham, Brian J., Lin, Charles Y., Kagey, Michael H., Rahl, Peter B., Lee, Tong I., Young, Richard A., 2013. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* 153, 307-319.
- Wu, K.-J., Grandori, C., Amacker, M., Simon-Vermot, N., Polack, A., Lingner, J., Dalla-Favera, R., 1999. Direct activation of TERT transcription by c-MYC. *Nature Genetics* 21, 220-224.
- Yang, Y., Su, Z., Song, X., Liang, B., Zeng, F., Chang, X., Huang, D., 2016. Enhancer RNA-driven looping enhances the transcription of the long noncoding RNA DHRS4-AS1, a controller of the DHRS4 gene cluster. *Scientific Reports* 6, 20961.

- Yao, Y., Ni, Y., Zhang, J., Wang, H., Shao, S., 2017. The role of Notch signaling in gastric carcinoma: molecular pathogenesis and novel therapeutic targets. *Oncotarget* 8, 53839-53853.
- Yin, X., Giap, C., Lazo, J.S., Prochownik, E.V., 2003. Low molecular weight inhibitors of Myc-Max interaction and function. *Oncogene* 22, 6151-6159.
- Yuan, C., Hu, H., Kuang, M., Chen, Z., Tao, X., Fang, S., Sun, Y., Zhang, Y., Chen, H., 2017a. Super enhancer associated RAI14 is a new potential biomarker in lung adenocarcinoma. *Oncotarget* 8, 105251-105261.
- Yuan, J., Jiang, Y.Y., Mayakonda, A., Huang, M., Ding, L.W., Lin, H., Yu, F., Lu, Y., Loh, T.K.S., Chow, M., Savage, S., Tyner, J.W., Lin, D.C., Koeffler, H.P., 2017b. Super-Enhancers Promote Transcriptional Dysregulation in Nasopharyngeal Carcinoma. *Cancer research* 77, 6614-6626.
- Zeid, R., Lawlor, M.A., Poon, E., Reyes, J.M., Fulciniti, M., Lopez, M.A., Scott, T.G., Nabet, B., Erb, M.A., Winter, G.E., Jacobson, Z., Polaski, D.R., Karlin, K.L., Hirsch, R.A., Munshi, N.P., Westbrook, T.F., Chesler, L., Lin, C.Y., Bradner, J.E., 2018. Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nature Genetics* 50, 515-523.
- 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., 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9, R137.
- Zhao, Z., Tavoosidana, G., Sjölander, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Singh Sandhu, K., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., Ohlsson, R., 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature Genetics* 38, 1341.
- Zheng, M., Tian, S.Z., Capurso, D., Kim, M., Maurya, R., Lee, B., Piecuch, E., Gong, L., Zhu, J.J., Li, Z., Wong, C.H., Ngan, C.Y., Wang, P., Ruan, X., Wei, C.-L., Ruan, Y., 2019. Multiplex chromatin interactions with single-molecule precision. *Nature* 566, 558-562.

- Zhu, J., Ding, H., Wang, X., Lu, Q., 2015. PABPC1 exerts carcinogenesis in gastric carcinoma by targeting miR-34c. *International Journal of Clinical and Experimental Pathology* 8, 3794-3802.
- Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van Ijcken, W.F.J., Grosveld, F.G., Ren, B., Wendt, K.S., 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences* 111, 996.