# The role of MYB in MLL-AF9 leukaemia




I-Jun Lau

Jesus College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Michaelmas 2019

# Abstract

MYB is a DNA-binding transcription factor with a critical role in the establishment of definitive haematopoiesis. In haematological malignancies, *MYB* is a direct downstream target common to several MLL-fusion proteins (MLL-FP), including MLL-AF9, the most common MLL-FP in acute myeloid leukaemia. MYB is necessary for the maintenance of MLL-AF9 leukaemia, and even partial and transient loss of MYB has been shown to abrogate MLL-AF9 leukaemia *in vivo*. While MYB inhibition is able to attenuate leukaemogenesis, its specific requirement in the processes of leukaemia initiation and transformation is still unclear. Furthermore, despite its obvious importance in normal and malignant haematopoiesis, much is still unknown regarding its molecular function. Recent evidence from disease models suggests that MYB functions through establishing aberrant enhancer activity, however this hypothesis has yet to be rigorously tested. In this thesis, to investigate the requirement for MYB in leukaemia initiation, we engineered a conditional miRNA mediated *Myb* repression circuit in a mouse model system, which represses *Myb* only in the presence of the MLL-AF9 oncoprotein. *Myb* repression during leukaemia initiation was correlated with a trend in decreased clonogenic potential *in vitro*, as well as a convincing reduction in leukaemia fitness *in vivo*, suggesting that a threshold level of *Myb* is required to propagate an aggressive leukaemia phenotype. Using an artificial chromatin recruitment system, we demonstrated that the MYB transcriptional activation domain was sufficient to generate an enhancer-like element and induce transcription of distal sequences, suggesting that MYB may be competent to initiate enhancer activity *in vivo*. Further to this, we showed that MYB dependent gene activation appears to be mediated through its maintenance of enhancer activity, as demonstrated by the loss of H3K27ac, enhancer RNA transcription and chromatin accessibility at the cognate intergenic enhancers of specific MYB sensitive genes with MYB knockdown. Importantly, we found that MYB dependent gene activation can be disrupted with minimal changes in enhancer-promoter interaction frequency, suggesting that these are functionally and temporally separable events.

**To my dearest family**

"We love because He first loved us"

– John 4:19

# Acknowledgements

I would like to begin by thanking my supervisors Tom Milne and Paresh Vyas. My unreserved thanks go to Tom, for fostering a co-operative and challenging lab environment in which I could learn and grow as a scientist, and also for his kind patience, constant encouragement and words of wisdom. I am also grateful to Paresh, for giving me the freedom to pursue my research interests and for his invaluable sponsorship and guidance in navigating the rigors of academic life.

I would like to extend my gratitude to all the members of their laboratories, both past and present. Especial thanks go to Laura Godfrey, who has steadfastly weathered the ups and downs of DPhil life with me, who is a "kindred spirit" and dearest friend. My sincerest thanks also go to Nick Crump, for always having the time to listen, help and teach, and for bravely enduring the first drafts of this thesis. I would also like to thank Phill North, for introducing me to the art of growing both cell lines and garden plants, and for our memorable conversations on the commute to and from work. Thanks also to Joe Harman, for attempting to educate me in the "dark art" of bioinformatics, and for persevering in sharing with me his love of the horror genre (to no avail). My DPhil years were all the richer for Lynn Quek, thanks to her acerbic wit and readiness to share a laugh, a natter and the prize-winning produce from her allotment. Special mention also goes to Jon Kerry for first showing me how to clone and for indulging my fascination with colour blindness, to Cat Garnett for being a great friend and colleague, and to Hedia Chagraoui for all her help and advice.

I have benefited immensely from the supportive and collaborative environment of the WIMM. In particular, I would like to thank Tudor Fulga, for admitting me into the exclusive "Circuit Club", as well as Alex Thomas and Simona Valletta, for all their expertise and instruction on mouse biology. I would also like to acknowledge Claus Nerlov and Adam Mead for their perceptive advice and support as part of my thesis committee. In addition, my heartfelt thanks go to Sarah Gooding, Debbie Hay, James Davies, Damien Downes, Jelena Telenius, Martin Larke, Catherine Porcher, Jim Hughes and Doug Higgs - all of whom have taken time out of their busy schedules to offer practical assistance, meaningful discussions and valuable insight. I am privileged to have been able to train in such a stimulating and progressive environment.

Finally, I would like to thank my wonderful family: my parents Soei Foan and Andrew, for their unconditional love and unwavering belief that I can achieve anything I set out to accomplish; my late Uncle Kai, who was the scientific trailblazer of the family and in whose footsteps I have aspired to follow; my godmother Laura, for her generous support through medical school; my extended family and my wider family of friends, for always being there to cheer me on.

Most importantly, I am thankful for my husband, Chris, and my daughter, Ava, who keep me grounded and every day give me the strength and inspiration to never give up. You both make everything worthwhile. Thank you for the gift of your love.

# Declaration

This work was made possible by generous funding from the Medical Research Council and the Anya Sturdy Charitable Trust For Medical Research for which I am very grateful.

I declare that the work presented in this thesis is entirely my own, and except where otherwise stated describes my own research. This thesis has not been submitted, either partially or in full, for another qualification at this University, or for a qualification at any other institution.

**Collaborators and their contributions to this work:**
**Figure 3.2.** The *Myb*$^{\mathrm{MRE/MRE}}$ mouse line was generated in collaboration with the Fulga (MRC Weatherall Institute of Molecular Medicine, Oxford) and Davies labs (Wellcome Trust Centre for Human Genetics, Oxford). sgRNA and ssODN design and optimisation was performed by Yale Michaels (Fulga Lab), and CRISPR-Cas9 editing experimental work was performed by the Davies lab.

**Figure 3.3.** Nanostring data was generated by the Fulga lab and analysed by Nicholas Crump.

**Figure 3.4.** Primers for genotyping of transgenic mESCs and mice were designed by Yale Michaels.

**Figures 3.5-3.8, 3.10-3.17.** Antibody staining panels for immunophenotypic analysis were designed by Alex Thomas and Simona Valletta (Nerlov lab, MRC Weatherall Institute of Molecular Medicine, Oxford) who also assisted with initial data acquisition and analysis.

**Figures 3.9-3.16.** Tail vein injections were performed by Alex Thomas.

**Figures 5.1** BRD4, H3K79me2, EZH2, MLL-N and CTCF ChIP-seq data used for analysis was generated by E.Ballabio, N.Crump, J.Kerry and T.Milne (Milne lab, MRC Weatherall Institute of Molecular Medicine, Oxford).

**Figures 5.2-5.5, 5.8-5.15.** Bioinformatic analysis downstream of mapping to the genome was performed by Nicholas Crump and Joe Harman (Milne lab, MRC Weatherall Institute of Molecular Medicine, Oxford).

# List of publications

Godfrey L, Crump NT, Thorne R, Lau IJ, Repapi E, Dimou D, Smith AL, Harman JR, Telenius JM, Oudelaar A, Downes DJ, Vyas P, Hughes JR & Milne TA. DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation. Nature Communications. 10, 2803 (2019). doi: https://doi.org/10.1038/s41467-019-10844-3.

Quek L, Otto GW, Garnett C, Lhermitte L, Karamitros D, Stoilova B, Lau IJ et al. Genetically distinct leukemic stem cells in human CD34-acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. J Exp Med . 2016 Jul 1;213(8):1513-1535. doi: 10.1084/jem.20151775.

Stavropoulou V, Kaspar S, Brault L, Sanders MA, Juge S, Morettini S, Tzankov A, Iacovino M, Lau IJ, Milne TA, Royo H, Kyba M, Valk PJ, Peters AH, Schwaller J. MLL-AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT Related Genes Linked to Poor Outcome. Cancer Cell. 2016 Jun 23. pii: S1535-6108(16)30216-1. doi: 10.1016/j.ccell.2016.05.011.

Eyre TA, Lau IJ, Mackillop L, Collins GP. Management and controversies of classical Hodgkin lymphoma in pregnancy. Br J Haematol. 2015 Jun;169(5):613-30. doi:10.1111/bjh.13327.

Lau IJ, Smith D, Aitchison R, Blesing N, Roberts P, Peniket A, Yong K, Rabin N, Ramasamy K. Bendamustine in combination with thalidomide and dexamethasone is a viable salvage option in myeloma relapsed and/or refractory to bortezomib and lenalidomide. Ann Hematol. 2015 Apr;94(4):643- 9. doi: 10.1007/s00277-014-2238-2.

Gooding S, Lau IJ, Sheikh M, Roberts P, Wong J, Dickens E, Bullement A, Elvidge J, Lee D, Ramasamy K Double Relapsed and/or Refractory Multiple Myeloma: Clinical Outcomes and Real World Healthcare Costs. PLoS One. 2015 Sep 14;10(9):e0136207. doi:10.1371/journal.pone.0136207.

Voiculescu O, Bodenstein L, Lau IJ, Stern CD. Local cell interactions and self- • -amplifying individual cell ingression drive amniote gastrulation. Elife. 2014 May 21;3:e01817. doi:10.7554/eLife.018172015;94:643–9.

# Contents

**List of Figures** **xi**

**List of Tables** **xiv**

**List of Abbreviations** **xv**

**1 Introduction** **1**

1.1 Transcription factors in gene regulation . . . . . 1

1.1.1 Chromatin architecture and organisation . . . . . 2

1.1.2 Transcription factor function . . . . . 2

1.1.3 Enhancer function . . . . . 5

1.1.4 Transcriptional dysregulation in disease . . . . . 9

1.2 *MLL* rearranged leukaemia . . . . . 10

1.2.1 *MLL* rearrangements in acute myeloid leukaemia . . . . . 11

1.2.2 The mutational landscape of *MLL*r leukaemia . . . . . 12

1.2.3 *MLL*r leukaemia - a model for leukaemogenesis . . . . . 12

1.2.4 *MLL*r leukaemia - a transcriptional disease . . . . . 14

1.3 MYB - a key regulator in normal and malignant haematopoiesis . . 17

1.3.1 The MYB family of transcription factors . . . . . 17

1.3.2 MYB in normal haematopoiesis . . . . . 18

1.3.3 MYB in leukaemia . . . . . 21

1.4 Summary and aims . . . . . 23

**2 Materials and methods** **24**

2.1 Mouse studies . . . . . 25

2.1.1 Generation of $Myb^{\mathrm{MRE/MRE}}$ mouse line . . . . . 25

2.1.2 Genotyping of mESCs and mice . . . . . 26

2.1.3 Peripheral blood analysis . . . . . 26

2.1.4 Dissection and preparation of tissue . . . . . 27

2.1.5 Immunophenotypic analysis and cell sorting . . . . . 27

2.1.6 Cell purification using magnetic-activated cell sorting . . . . 28

2.1.7 Retroviral transduction and serial replating . . . . . 29

2.1.8 Functional assays . . . . . 30

2.2 Cell lines and cell culture techniques . . . . . 31

2.2.1 Cell line culture conditions . . . . . 31

2.2.2 Generation of TetR protein expressing mESC lines . . . . . 31

2.2.3 Colony forming assays . . . . . 32

2.3 RNAi assays . . . . . 32

2.4 Protein assays . . . 33
2.4.1 Western blotting . . . 33
2.4.2 Intracellular FACS . . . 34
2.5 Chromatin immunoprecipitation (ChIP) assays . . . 34
2.5.1 ChIP qPCR . . . 36
2.5.2 ChIP-sequencing . . . 36
2.5.3 Reference normalised ChIP-sequencing (ChIP-rx) . . . 36
2.6 RNA assays . . . 37
2.6.1 Gene expression analysis . . . 37
2.6.2 Nascent RNA-sequencing . . . 37
2.6.3 Poly(A) minus RNA-sequencing . . . 38
2.7 ATAC-seq . . . 38
2.8 Capture-C . . . 39
2.9 Bioinformatic analysis . . . 40
2.9.1 Sequence analysis . . . 40
2.9.2 Locus Overlap Analysis . . . 41
2.9.3 RNA-seq and gene expression analysis . . . 41
2.9.4 Capture-C analysis . . . 41

**3 Myb repression during leukaemogenesis results in an attenuated leukaemia phenotype 43**
3.1 Introduction . . . 44
3.1.1 MicroRNA mediated regulation of gene expression . . . 44
3.1.2 Engineering an endogenous miRNA mediated *Myb* repression circuit . . . 46
3.2 Aims . . . 48
3.3 Results . . . 48
3.3.1 Generation of an endogenous *Myb* repression circuit in mice 48
3.3.2 $Myb^{MRE/MRE}$ mice exhibit no major deficits in steady state haematopoiesis . . . 51
3.3.3 $Myb^{MRE/MRE}$ mice have preserved bone marrow reconstitution ability . . . 56
3.3.4 *In vitro* assay optimisation . . . 59
3.3.5 $Myb^{MRE/MRE}$ cells show a decreased trend in transformation potential *in vitro* . . . 61
3.3.6 *In vivo* experimental approach to test the requirement for Myb in leukaemic transformation . . . 65
3.3.7 *MLL-AF9* transduced $Myb^{MRE/MRE}$ and WT cells are immunophenotypically similar prior to transplantation . . . 66

3.3.8 MLL-AF9 transduced *Myb*$^{MRE/MRE}$ cells give rise to an attenuated leukaemia phenotype *in vivo* . . . . . 68
3.4 Discussion . . . . . 71

**4 MYB is sufficient for establishing an enhancer-like element 74**
4.1 Introduction . . . . . 74
4.1.1 MYB structure, regulation and function . . . . . 74
4.1.2 Understanding the role of MYB in establishing enhancer function . . . . . 79
4.1.3 The mouse embryonic stem cell TetR system . . . . . 80
4.2 Aims . . . . . 82
4.3 Results . . . . . 83
4.3.1 Myb binding is sufficient for the recruitment of functional CBP and P300 . . . . . 83
4.3.2 The Myb transactivation domain is sufficient for stable binding of its known interaction partners . . . . . 86
4.3.3 The Myb transactivation domain is sufficient for the recruitment of transcription activation complexes and RNAPII . . 90
4.3.4 Myb is sufficient for establishing an enhancer-like element at chromatin . . . . . 91
4.3.5 A single TetR binding site is sufficient for Myb dependent enhancer-like activity . . . . . 95
4.4 Discussion . . . . . 98

**5 MYB dependent enhancer function is distinct from the maintenance of enhancer-promoter proximity 101**
5.1 Introduction . . . . . 102
5.2 Aims . . . . . 103
5.3 Results . . . . . 103
5.3.1 MYB binding correlates with CBP and P300 at enhancers and promoters . . . . . 103
5.3.2 Identifying a set of MYB sensitive gene targets in MLL-AF9 leukaemia . . . . . 110
5.3.3 MYB dependent gene targets have higher levels of MYB binding and are important for cell proliferation and replication 114
5.3.4 H3K27ac at enhancers and promoters is sensitive to MYB knockdown . . . . . 115
5.3.5 Loss of MYB binding results in differential changes to chromatin accessibility . . . . . 117

5.3.6 Loss of MYB binding does not affect TAD structure or enhancer-promoter looping . . . . . 119
5.3.7 MYB is necessary for eRNA transcription at the enhancers of MYB dependent gene targets . . . . . 123
5.4 Discussion . . . . . 127

**6 Discussion and future work** **132**
6.1 General discussion . . . . . 132
6.2 An *in vivo* model for understanding disease resistance . . . . . 133
6.3 A model for MYB function at enhancers . . . . . 133
6.3.1 MYB function in establishing enhancer activity . . . . . 133
6.3.2 MYB function in maintaining enhancer activity . . . . . 135
6.4 Future work . . . . . 136

**7 Appendices** **138**

**References** **145**

# List of Figures


1.1 MLL-FPs drive aberrant gene expression . . . . . 15

3.1 MicroRNA mediated gene regulation. . . . . 46
3.2 Schematic of engineering an endogenous *Myb* repression circuit. . . 47
3.3 *MLL-AF9* expression in mouse Lin- bone marrow cells results in upregulation of *Myb* and *miR-196-b*. . . . . 49
3.4 Engineered *miR-196-b* MREs at the *Myb* 3' UTR are functional in mESCs. . . . . 50
3.5 Analysis of HSC and multipotent progenitor populations. . . . . 52
3.6 Analysis of myeloerythroid progenitor populations. . . . . 53
3.7 Analysis of lymphoid progenitor populations. . . . . 54
3.8 Analysis of peripheral blood lineages. . . . . 55
3.9 Experimental design for competitive repopulation assay. . . . . 57
3.10 $Myb^{\mathrm{MRE/MRE}}$ mice have increased PB reconstitution ability. . . . . 57
3.11 $Myb^{\mathrm{MRE/MRE}}$ mice have preserved BM reconstitution ability. . . . . 58
3.12 Mouse GMP and Lin- cells have comparable transformation efficiency *in vitro* when transduced with *MLL-AF9*. . . . . 60
3.13 Experimental design for *in vitro* assay . . . . . 62
3.14 No significant difference in the clonogenic potential of *MLL-AF9* transduced $Myb^{\mathrm{MRE/MRE}}$ and WT Lin- cells . . . . . 63
3.15 *MLL-AF9* transduced $Myb^{\mathrm{MRE/MRE}}$ cells have lower Myb expression compared to WT controls. . . . . 64
3.16 Experimental design for *in vivo* assay . . . . . 65
3.17 MLL-AF9 transduced $Myb^{\mathrm{MRE/MRE}}$ and WT cells are immunophenotypically similar prior to transplantation . . . . . 67
3.18 Mice transplanted with MLL-AF9 transduced $Myb^{\mathrm{MRE/MRE}}$ cells have a survival advantage. . . . . 69

4.1 Schematic of MYB protein structure and conserved domains . . . . 75
4.2 The TetR system. . . . . 81
4.3 Validation of the TetR-Myb system. . . . . 84
4.4 Cbp and p300 recruitment is dependent on TetR-Myb. . . . . 85
4.5 H3K27ac is dependent on TetR-Myb. . . . . 86
4.6 Validation of the TetR-TA system. . . . . 87
4.7 The Myb transactivation domain is sufficient for stable binding of its known interaction partners . . . . . 89
4.8 The Myb transactivation domain is sufficient for the recruitment of Brd4, Med1 and RNAPII. . . . . 90

4.9 Myb binding results in increased chromatin accessibility and active transcription. . . . 92
4.10 The Myb transactivation domain is sufficient for the recruitment of H3K4me1 but not H3K4me3. . . . 94
4.11 Myb-TA recruitment of Cbp, H3K27ac, H3K4me1, Brd4 and RNAPII at a single endogenous TetO site in mouse ESCs. . . . 96
4.12 Myb-TA binding at a single endogenous *TetO* site results in increased chromatin accessibility and active transcription. . . . 97

5.1 ChIP-seq tracks at the *BCL2* locus. . . . 104
5.2 Genome-wide distribution and correlations of MYB ChIP-seq peaks with other datasets. . . . 105
5.3 Correlational metaprofiles and heatmaps of MYB, CBP, P300 and H3K27ac at gene regulatory elements. . . . 107
5.4 Correlational Violin plots of MYB, CBP, P300 and H3K4me3 at gene regulatory elements. . . . 108
5.5 LOLA analysis at MYB-bound vs non-MYB-bound peaks. . . . 109
5.6 Optimisation of MYB knockdown using targeting siRNA. . . . 111
5.7 Time course analysis of MYB KD. . . . 112
5.8 Loss of MYB binding results in changes in transcription. . . . 113
5.9 Loss of MYB binding results in transcriptional changes affecting leukaemia maintenance. . . . 114
5.10 Loss of MYB binding results in decreased H3K27ac at gene regulatory elements. . . . 116
5.11 *BCL2* is an example of a MYB dependent gene target. . . . 117
5.12 Loss of MYB binding results in sporadic ATAC changes. . . . 118
5.13 Capture-C traces at genes which are transcriptionally downregulated following MYB knockdown. . . . 120
5.14 Capture-C traces at genes which are transcriptionally upregulated following MYB knockdown. . . . 121
5.15 Capture-C traces at genes which are transcriptionally unchanged following MYB knockdown. . . . 122
5.16 MYB is required for eRNA transcription at *BCL2*. . . . 124
5.17 MYB is required for eRNA transcription at *MYC*. . . . 125
5.18 MYB is required for eRNA transcription at *EZH2*. . . . 126

6.1 Model for the role of MYB in establishing enhancer activity. . . . 134
6.2 Model for the role of MYB in maintaining enhancer activity. . . . 135

7.1 MSCV MLL-AF9/mCherry plasmid map . . . 141

7.2 Domain view of Capture-C traces at downregulated genes . . . . . 142
7.3 Domain view of Capture-C traces at upregulated genes . . . . . 143
7.4 Domain view of Capture-C traces at unchanged genes . . . . . 144

# List of Tables


1.1 Transcription elongation complexes related to MLL-FP function . . 16

2.1 Immunophenotypic markers used to define HSC and multipotent progenitor populations . . . . . 28
2.2 Immunophenotypic markers used to define myeloerythroid progenitors 28
2.3 Immunophenotypic markers used to define lymphoid progenitors . . 29
2.4 Immunophenotypic markers used to define peripheral blood lineages 29
2.5 Antibodies used in Western blot and ChIP experiments . . . . . 36

4.1 MYB post-translational modifications and effect on function . . . . 77
4.2 MYB direct protein interactions. . . . . 78
4.3 Conservation of human MYB vs mouse Myb interacting proteins . . 83

7.1 Antibodies used for flow cytometry analysis and sorting . . . . . 139

# List of Abbreviations

**3C** . . . . . Chromosome conformation capture

**4-SU** . . . . . 4-Thiouridine

**7AAD** . . . . . 7-aminoactinomycin D

**AGO** . . . . . Argonaut protein

**ALL** . . . . . Acute lymphoid leukaemia

**ALP** . . . . . All lymphoid progenitors

**AML** . . . . . Acute myeloid leukaemia

**ATAC** . . . . . Assay for Transposase-Accessible Chromatin

**BLP** . . . . . B-lymphoid progenitors

**BM** . . . . . Bone marrow

**bp** . . . . . Base pair

**ChIP** . . . . . Chromatin immunoprecipitation

**CDB** . . . . . ChIP dilution buffer

**cDNA** . . . . . Complementary DNA

**CFU** . . . . . Colony forming unit

**CFU-E** . . . . Erythroid colony forming unit

**CLP** . . . . . Common lymphoid progenitor

**$CO_2$** . . . . . Carbon dioxide

**CRISPR** . . . Clustered regularly interspaced short palindromic repeats

**CTD** . . . . . C-terminal domain

**DB(D)** . . . . . DNA binding (domain)

**dCas9** . . . . . Cas9 Endonuclease Dead, aka dead Cas9

**DNA** . . . . . Deoxyribonucleic acid

**DNase I** . . . . Deoxyribonuclease I

**DSG** . . . . . Disuccinimidyl glutarate

**DTT** . . . . . 1,4-Dithiothreitol

**ECL** . . . . . Enhanced chemiluminescence

**EDTA** . . . . . Ethylenediaminetetraacetic acid

**EPO** . . . . . Erythropoietin

**eRNA** . . . . . Enhancer ribonucleic acid

**FA** . . . . . Formaldehyde

**FACS** . . . . . Fluorescence-activated cell sorting

**FCS** . . . . . Foetal calf serum

**FMO** . . . . . Fluorescence minus one

**GMP** . . . . . Granulocyte macrophage progenitor

**gRNA** . . . . . Guide ribonucleic acid

**HSC** . . . . . Haematopoietic stem cell

**HSPC** . . . . . Haematopoietic stem progenitor cell

**IDD** . . . . . Intrinsically disordered domain

**IL-3** . . . . . Interleukin 3

**IL-6** . . . . . Interleukin 6

**IMDM** . . . . Iscove Modified Dulbecco Media

**kb** . . . . . Kilo base

**KD** . . . . . Knockdown

**KEE** . . . . . H3K79me2/3-dependent enhancer element

**LIC** . . . . . Leukaemia initiating cell

**Lin** . . . . . Lineage

**LK** . . . . . Lin- c-Kit+

**lncRNA** . . . . Long non-coding ribonucleic acid

**LOLA** . . . . . Locus overlap analysis

**LSC** . . . . . Leukaemia stem cell

**LT-HSC** . . . . Long-term haematopoietic stem cell

**LMPP** . . . . . Lymphoid-primed multipotent progenitor

**LSK** . . . . . Lin- Sca-1+ c-Kit+

**MACS** . . . . . Magnetic-activated cell sorting

**MCV** . . . . . Mean cell volume

**MDS** . . . . . Myelodysplastic syndrome

**MEP** . . . . . Megakaryocyte-erythroid progenitor

**mESC** . . . . . Mouse embryonic stem cell

**miRISC** . . . . miRNA-mediated RNA-induced silencing complex

**miRNA** . . . . Micro ribonucleic acid

**MkP** . . . . . Megakaryocyte progenitor

**MLL** . . . . . Mixed lineage leukaemia

**MLL-FP** . . . MLL-fusion protein

**MLLr** . . . . . Mixed lineage leukaemia gene rearranged/rearrangement

**MPP** . . . . . Multipotent progenitor

**MRE** . . . . . miRNA response element

**MRE:MA9** . . $Myb^{\mathrm{MRE/MRE}}$ cells transduced with *MLL-AF9*

**MSCV** . . . . . Murine Stem Cell Virus

**ncRNA** . . . . Non-coding ribonucleic acid

**NGS** . . . . . Next generation sequencing

**NK/T** . . . . . Natural killer T cell

**NMR** . . . . . Nuclear magnetic resonance

**NR(D)** . . . . Negative regulatory (domain)

**nRNA** . . . . . Nascent ribonucleic acid

**ORF** . . . . . Open reading frame

**PB** . . . . . Peripheral blood

**PBS** . . . . . Phosphate-buffered saline

**PCA** . . . . . Principal component analysis

**PCR** . . . . . Polymerase chain reaction

**PGK** . . . . . 3-phosphoglycerate kinase

**PIC** . . . . . Protease inhibitor cocktail

**PRC** . . . . . Polycomb Repressive Complex

**Pre CFU-E** . . Pre-erythroid colony forming unit

**Pre GM** . . . . Pre–granulocyte-macrophage progenitor

**Pre MegE** . . Pre-megakaryocyte erythroid progenitor

**Pro Ery** . . . . Proerythroblast

**PTD** . . . . . Partial tandem duplication

**P-TEFb** . . . . Positive transcription elongation factor B

**PTD** . . . . . Partial tandem duplication

**PVDF** . . . . . Polyvinylidene fluoride

**qPCR** . . . . . Quantitative polymerase chain reaction

**qRT-PCR** . . . Quantitative reverse transcriptase polymerase chain reaction

**RISC** . . . . . RNA-induced silencing complex

**RNA** . . . . . Ribonucleic acid

**RNAi** . . . . . Ribonucleic acid inhibition

**RNAPII** . . . Ribonucleic acid polymerase II

**RPMI** . . . . . Roswell Park Memorial Institute

**rRNA** . . . . . Ribosomal ribonucleic acid

**SCF** . . . . . Stem cell factor

**SDS** . . . . . Sodium dodecyl sulfate

**SE** . . . . . Super-enhancer

**SEC** . . . . . Super elongation complex

**SEM** . . . . . Standard Error of the Mean

**-seq** . . . . . Sequencing

**sgRNA** . . . . Single guide ribonucleic acid

**snoRNA** . . . Small nucleolar ribonucleic acid

**SNP** . . . . . Single nucleotide polymorphism

**ssODN** . . . . Single-stranded oligo DNA nucleotides

**TA(D)** . . . . . Transactivation (domain)

**TAD** . . . . . Topologically associating domain

**TAF** . . . . . TBP-associated factor

**T-ALL** . . . . . T-cell acute lymphoblastic leukaemia

**TBP** . . . . . TATA-box binding protein

**TBS** . . . . . Tris-buffered saline

**TetO** . . . . . Tetracycline operator

**TetR** . . . . . Tetracycline repressor

**TF** . . . . . Transcription factor

**Tris-HCl** . . . Tris(hydroxymethyl)aminomethane hydrochloride

**tRNA** . . . . . Transfer ribonucleic acid

**TSS** . . . . . Transcription start site

**UTR** . . . . . Untranslated region

**WBC** . . . . . While blood cell

**WT** . . . . . Wild type

**WTLM** . . . . Wild type litter mate

**WT:MA9** . . . WT cells transduced with *MLL-AF9*

# 1 Introduction

## Contents


**1.1 Transcription factors in gene regulation . . . . . 1**
1.1.1 Chromatin architecture and organisation . . . . . 2
1.1.2 Transcription factor function . . . . . 2
1.1.3 Enhancer function . . . . . 5
1.1.4 Transcriptional dysregulation in disease . . . . . 9
**1.2 *MLL* rearranged leukaemia . . . . . 10**
1.2.1 *MLL* rearrangements in acute myeloid leukaemia . . . . 11
1.2.2 The mutational landscape of *MLL*r leukaemia . . . . . 12
1.2.3 *MLL*r leukaemia - a model for leukaemogenesis . . . . . 12
1.2.4 *MLL*r leukaemia - a transcriptional disease . . . . . 14
**1.3 MYB - a key regulator in normal and malignant haematopoiesis . . . . . 17**
1.3.1 The MYB family of transcription factors . . . . . 17
1.3.2 MYB in normal haematopoiesis . . . . . 18
1.3.3 MYB in leukaemia . . . . . 21
**1.4 Summary and aims . . . . . 23**


## 1.1 Transcription factors in gene regulation

The accurate, precise, and robust regulation of gene expression is the cornerstone of maintaining homeostasis and normal function. Eukaryotic gene expression begins with the specific step wise regulation of transcription initiation and elongation by ribonucleic acid polymerase II (RNAPII), and the subsequent assembly of ribonucleic acid (RNA) or protein product. DNA-binding transcription factors (TFs) ensure this specificity by occupying defined DNA sequences at distal and proximal cis-regulatory elements. Disruption of this process can lead to disease phenotypes.

### 1.1.1 Chromatin architecture and organisation

In eukaryotes, TF access to regulatory regions in the genome is tightly controlled by chromatin structure. DNA is wrapped around histone octamers to form nucleosomes and this is further arranged into spatially-segregated open (accessible) and closed (inaccessible) chromatin regions within the nucleus (Lieberman-Aiden *et al.* 2009). Thus, gene regulation begins with the structural organisation of the genome.

Within regions of euchromatin, chromosome conformation capture (3C) techniques such as Hi-C have identified higher order chromatin structures, commonly referred to as topologically associating domains (TADs) (Lieberman-Aiden *et al.* 2009; Dixon *et al.* 2012; Mifsud *et al.* 2015). Emerging work suggests that TADs are demarcated by the combined effects of cohesin and CTCF-mediated boundary regions, through a process described as loop extrusion (Sanborn *et al.* 2015; Fudenberg *et al.* 2016). In addition, global 3C techniques, and more recently, higher resolution and targeted next generation techniques such as Capture Hi-C and Capture-C respectively, have also revealed the genome-wide presence of further looping structures within these domains, occurring most commonly between enhancers and promoters (Sanyal *et al.* 2012; Mifsud *et al.* 2015; Schoenfelder *et al.* 2015; Davies *et al.* 2016). The assembly of chromatin into TADs is therefore thought to contribute to gene regulation by creating structural compartments, thereby facilitating and limiting the activity of cis-regulatory elements to genes within the same region.

### 1.1.2 Transcription factor function

Transcription factors (TFs) are a class of proteins that are responsible for complex temporal and tissue-specific control of gene expression. They activate (or repress) their gene targets by binding to short, specific DNA sequences at enhancers and promoters (Small *et al.* 1992; Yuh *et al.* 1994; Sandmann *et al.* 2007; Wunderlich *et al.* 2009; Spitz *et al.* 2012).

In general, TFs typically consist of at least one DNA binding (DB) domain and at least one functionally separate transactivation (TA) domain (Brent *et al.* 1985;

Keegan *et al.* 1986). DB domains fall into distinct structural groups, by which TFs are classified (Fulton *et al.* 2009; Vaquerizas *et al.* 2009). TF binding to DNA is mediated by the DB domain, leading to specific recognition of their cognate DNA sequences that are usually 4-12bp in length (Vaquerizas *et al.* 2009; Hume *et al.* 2015). TA domains are usually intrinsically disordered domains (IDDs) and are classified according to their amino acid profile or predicted structure (Mitchell *et al.* 1989; Sigler 1988; Triezenberg 1995; Roberts 2000; Staby *et al.* 2017). Importantly, IDDs are thought to be crucial for the formation of phase condensates, which are nuclear structures that are considered to be key for gene regulation and will discussed in more detail below (1.1.3.1). Most TA domains interact with an overlapping set of co-activating complexes and it is via these interactions that the TFs are able to regulate transcription (Dai *et al.* 1996; Eckner 1996; Oliner *et al.* 1996; Yuan *et al.* 1996; Yin *et al.* 2014; B. L. Allen *et al.* 2015). Both the DB and TA domains are a prerequisite for gene activation (Guarente *et al.* 1982; Hochschild *et al.* 1983).

The majority of TFs have been observed to bind only at sites of DNAse I hypersensitivity (Thurman *et al.* 2012) which indicates regions of accessible DNA which are not nucleosome bound and are thus unprotected from digestion. It was therefore previously believed that a pre-existing level of chromatin accessibility was required for TF binding to occur in eukaryotes (Guertin *et al.* 2010; Biddie *et al.* 2011). There is now reasonable evidence to suggest that TFs can contribute directly to both initiating and maintaining chromatin accessibility. Firstly, it has been proposed that TFs can initiate accessibility remodelling by displacing nucleosomes through passive competition (Svaren *et al.* 1994). TFs exploit short periods of DNA accessibility during nucleosome turnover events e.g. shortly after replication (Ramachandran *et al.* 2016; Petruk *et al.* 2017a; Petruk *et al.* 2017b), to gain a foothold on histone-bound DNA. The probability of a binding event taking place increases with the concentration of the competing TF, as demonstrated for CEBPa and NFIB, where transcriptional upregulation of these proteins induces accessibility of their target genes during cellular reprogramming (Di Stefano *et al.* 2016; Denny *et al.* 2016). Secondly, TFs can bind to internucleosomal (linker) DNA,

and either through passive displacement of architectural proteins alone (Mirny 2010), or together with chromatin remodellers, destabilise proximal nucleosomes, thereby establishing access to previously inaccessible DNA (Fan *et al.* 2005; McBryant *et al.* 2006; Nalabothula *et al.* 2014; Gomez-Diaz *et al.* 2014). Thirdly, some TFs have been shown to bind to an already accessible regulatory element and initiate distal chromatin remodelling in trans, for example MYOD1 binding to its already accessible enhancer results in nucleosome depletion at its promoter (Taberlay *et al.* 2011). Finally, certain 'pioneer TFs' including FOXA1, PU.1 and OCT4, have been shown to bind directly to nucleosomal DNA and either independently, or through the recruitment of chromatin remodellers, generate an accessible state (Cirillo *et al.* 2002; Holmqvist *et al.* 2005; Iwafuchi-Doi *et al.* 2016b; Iwafuchi-Doi *et al.* 2016a; King *et al.* 2017). In all the above models, following initial nucleosome eviction, subsequent binding of secondary TFs and other proteins co-operate to stabilise and maintain the open chromatin state (Tillo *et al.* 2010; Spitz *et al.* 2012). Evidence along multiple lines suggests that that the binding of other TFs nearby can help stabilise TF:DNA interactions. The mechanisms of co-operative TF binding are complex but briefly, they may occur via direct physical TF:TF association before or concurrent with DNA binding, or simply by TFs binding independent of each other at the same accessible regulatory element (Arnosti *et al.* 1996; Walhout 2006; Wilson *et al.* 2010; Voss *et al.* 2011; Jolma *et al.* 2013). Related to this, combinatorial TF binding has also been shown to be important in activating tissue specific genes, where target gene activation only occurs when particular combinations of TFs are bound to specific gene regulatory elements (Ness *et al.* 1994). Examples include co-operation of Myb and Cebpb to transactivate *Mim-1* (Ness *et al.* 1993) and Myb and Ets2 in *CD13/APN* transactivation (Shapiro 1995).

When bound at target loci, TFs facilitate the recruitment and assembly of protein complexes. These include co-activating complexes made up of proteins such as the lysine acetyltransferases CBP/P300, chromatin proteins such as BRD4, which recruits the positive transcription elongation factor P-TEFb, and Mediator which forms part of the RNAPII pre-initiation complex (Cho *et al.* 2018; Sabari

*et al.* 2018; Y. Shin *et al.* 2018). Subsequently, active histone modifications – including H3K27ac, which incidentally is deposited by CBP/P300 – keep chromatin open for binding of secondary TFs and also create additional affinities to further stabilise complexes, for example BRD4 binding to H3K27ac (Zeng *et al.* 2002; Wu *et al.* 2007; Filippakopoulos *et al.* 2012; Gilan *et al.* 2016). In this way, TFs create a permissive environment for gene activation by recruiting co-activating complexes and mediating epigenetic modifications, which together serve to promote transcription initiation and elongation by RNAPII.

### 1.1.3 Enhancer function

Regulation of gene expression requires efficient assembly of the transcription apparatus at specific genomic sites. TFs contribute to this specificity by occupying specific DNA sequences at enhancers and promoter elements. Enhancers are canonically defined as short (100-1000bp), non-coding DNA sequences that regulate transcription independent of the relative distance, location or orientation to their cognate promoter/s (Banerji *et al.* 1981; Benoist *et al.* 1981; Maniatis *et al.* 1987; Lettice *et al.* 2002; Lettice *et al.* 2003). Enhancers provide a platform for transcription factor activity and thus, transcription factor and enhancer function is inextricably linked.

Two broad models have been proposed to explain how TF binding at enhancers may serve to regulate gene expression (Arnosti *et al.* 2005). In the "enhanceosome" model, the binding sites within an enhancer facilitate the coordinated activity of TFs to stabilise the assembly of co-activating complexes, leading to gene activation. Thus, alterations in the sequence or location of a single binding site, or the absence of one regulatory protein, results in overall loss of enhancer activity. In contrast, the "billboard" model proposes that each binding site or binding site cluster within an enhancer exists as an individual functional unit, such that the entire enhancer element is an ensemble of multiple functional units, each of which are able to independently regulate gene expression.

Several general characteristics have been associated with transcriptionally active enhancers. Although not universal, these include 1) concentrated clusters of TF binding, 2) DNase I hypersensitivity as a proxy for chromatin accessibility, 3) the presence of activating complexes and transcriptional co-factors such as BRD4 and Mediator, 4) histone modifications such as H3K27ac and H3K4me1, 5) physical proximity between the enhancer and promoter, and more recently, 6) the production of non-coding enhancer RNA transcripts, termed eRNAs (Heintzman *et al.* 2007; Xi *et al.* 2007; Heintzman *et al.* 2009; Creyghton *et al.* 2010; Thurman *et al.* 2012; T.-K. Kim *et al.* 2015; Y. Liu *et al.* 2018; Schoenfelder *et al.* 2019).

#### 1.1.3.1 Enhancer-promoter interactions

Enhancers may be situated at great linear genomic distances away from the genes they regulate. Through the three dimensional organisation of chromatin, enhancers and the promoters of their gene targets may be brought into close physical proximity, bypassing other gene promoters located in the intervening genomic sequence (Furlong *et al.* 2018; Schoenfelder *et al.* 2019). Indeed, at many gene loci, proximity between the enhancer and the promoter is thought to be essential for enhancer function and gene activation (H. Chen *et al.* 2018; Lim *et al.* 2018), although some notable exceptions have been described (Alexander *et al.* 2019; Benabdallah *et al.* 2019).

Super-enhancers (SEs) are large enhancers with enriched binding of TFs and activating complexes such as BRD4 and Mediator and are associated with high levels of transcriptional activity of their target genes (Whyte *et al.* 2013; Hnisz *et al.* 2013; Sabari *et al.* 2018). In cancer cells, important oncogenes are often associated with SEs (Hnisz *et al.* 2013; Loven *et al.* 2013) at which the formation of phase-separated condensates have been observed. Phase separation is partially mediated by weak affinities between the IDDs of proteins, for example the TA domains of TFs and other co-activators (Boija *et al.* 2018; Sabari *et al.* 2018; Shrinivas *et al.* 2019). Recent work on phase separated activation complexes that include proteins such as Mediator (e.g. MED1) and BRD4 has provided an interesting model which links the assembly of activation complexes at enhancers (particularly SEs)

and promoter sequences to the initiation and maintenance of interactions between enhancer and promoter. (Boija *et al.* 2018; Sabari *et al.* 2018; Shrinivas *et al.* 2019). Other possible models include CTCF/cohesin stabilised enhancer-promoter interactions (Schuijers *et al.* 2018; El Khattabi *et al.* 2019; Hyle *et al.* 2019), or potentially protein/RNA complexes bound to both the enhancer and promoter that interact with one another (W. Li *et al.* 2013; Hsieh *et al.* 2014; Weintraub *et al.* 2017). The presence of specific histone modifications at the enhancer is thought to contribute to one or all of these models, mainly by stabilising the presence of specific protein complexes, such as cohesin binding to H3K4me1 or BRD4 binding to H3K27ac (Gilan *et al.* 2016; Yan *et al.* 2017; B. Lai *et al.* 2017). Recent work also suggests that at H3K79me2/3-dependent enhancer elements (KEEs), the presence of H3K79me2/3 can help maintain open chromatin regions to facilitate the binding of sequence specific transcription factors, and is required for enhancer-promoter interaction (Godfrey *et al.* 2019). This could constitute a more general principle where histone modifications help regulate DNA accessibility and TF binding, and ultimately the formation of enhancer-promoter loops.

#### 1.1.3.2 Enhancer RNAs

RNAPII had previously been observed to generate non-coding RNAs at locus control regions (Tuan *et al.* 1992; Ashe *et al.* 1997; Masternak *et al.* 2003; Rogan *et al.* 2004; Ling *et al.* 2005; Ho *et al.* 2006), but it is only recently recognised that mammalian enhancers are broadly transcribed and generate eRNAs (De Santa *et al.* 2010; J. Kim *et al.* 2010; Koch *et al.* 2011).

A large proportion of RNAPII activity is found at non-coding regions of the genome (Carninci *et al.* 2005; Birney *et al.* 2007; Kapranov *et al.* 2007), resulting in non-coding RNA species. eRNAs are one such group and are distinct from other better characterised, evolutionarily conserved non-coding RNAs (ncRNAs), such as transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), small nucleolar RNAs (snoRNAs), and microRNAs (miRNAs).

Characteristically, eRNAs are capped, mostly short, unspliced transcripts (∼500-2000bp), and are not usually subject to polyadenylation, although polyadenylated eRNAs have been observed (T.-K. Kim *et al.* 2010; Koch *et al.* 2011; R. Andersson *et al.* 2014a). Thus they are relatively short-lived and often subject to rapid degradation by the exosome (Preker *et al.* 2008; R. Andersson *et al.* 2014b). eRNAs have been found to be generated from putative enhancer regions defined as having H3K27ac and higher levels of H3K4me1/2 compared with the H3K4me3 (Heintzman *et al.* 2009). They typically arise from bidirectional transcription from the enhancer, although unidirectional transcription has also been observed (W. Li *et al.* 2013; R. Andersson *et al.* 2014b; Mikhaylichenko *et al.* 2018). They also exist in much lower abundance compared to mRNAs (19-34-fold lower) and are not exported from the nucleus (R. Andersson *et al.* 2014a; R. Andersson *et al.* 2014b).

eRNA expression is highly correlated with changes in the expression of nearby genes, which suggests a functional link between eRNA and gene expression (T.-K. Kim *et al.* 2010; Wang *et al.* 2011; Lam *et al.* 2013; Raisner *et al.* 2018). It has been proposed that eRNAs are not simply a bystander effect of enhancer activity, but are in fact able to contribute to enhancer function via several mechanisms.

Firstly, there is evidence to suggest that eRNAs play a role in establishing chromatin accessibility. Loss of eRNA expression at distal regulatory elements has been shown to correlate with decreased DNase I hypersensitivity at the promoters of myogenic genes (MYOD1 and MYOG) and reduction in their gene expression (Mousavi *et al.* 2013). Also, as it is well documented that acetylation of H3K27 by CBP/P300 helps maintain open chromatin states, the observation that eRNAs can directly interact with CBP and stimulate its histone acetyltransferase (HAT) activity may provide a possible mechanism for their role in chromatin accessibility (Bose *et al.* 2017a). Secondly, although there is no evidence that enhancer-promoter looping is affected by loss of eRNA expression (Schaukowitch *et al.* 2014; Raisner *et al.* 2018), eRNAs have been shown to stabilise enhancer-promoter interactions through their interaction with Cohesin, CTCF and Mediator (W. Li *et al.* 2013; F. Lai *et al.* 2013; Hansen *et al.* 2019). Thirdly, several eRNAs have also been

shown to facilitate transcription elongation, inducing negative elongation factor (NELF) release from paused RNAPII (Schaukowitch *et al.* 2014).

It has previously been observed that G-quadruplexes are enriched at TF binding sites, and specifically at enhancer and promoter elements (Hegyi 2015; Hou *et al.* 2019). More recently, short G-tract and G-quadruplex RNAs have been shown to interact with PRC2 via its subunit SUZ12 (Kanhere *et al.* 2010; Beltran *et al.* 2016). This interaction is mutually antagonistic with chromatin and results in the removal of PRC2 from chromatin (Beltran *et al.* 2016; Beltran *et al.* 2019). While this phenomenon has been observed predominantly at gene promoters, it suggests another possible role for eRNAs in enhancer activation, where G-tract eRNAs result in PRC2 eviction from enhancers and thus facilitate the activation of target genes.

### 1.1.4 Transcriptional dysregulation in disease

It is clear that TFs play a key role in gene regulation and this is most apparent when their normal function is disrupted, as illustrated by the many diseases that are associated with regulatory element sequence-variations, as well as the overexpression or mutation of TFs.

When expressed at physiological levels, a TF typically binds at regulatory regions for which it has high affinity. However, in overexpressed states, the increased concentration of TF molecules increases its probability of binding not only to high affinity sites, but also to low affinity sites (X. Liu *et al.* 2002). This results in changes in its genome wide occupancy and the activity of its target genes. For example, when MYC expression was induced in a B cell lymphoma cell line, its binding was observed to increase at previously occupied loci. Additionally, in different cancers, MYC overexpression was associated with binding at low affinity enhancer regions. In both cases, increased MYC occupancy was associated with amplification of target gene expression (C. Y. Lin *et al.* 2012). Therefore, simply overexpressing oncogenic TFs can alter the transcriptional circuitry of the cell through aberrant targeting. Furthermore, disease-associated variants in regulatory DNA have been shown to affect transcription factor occupancy across the genome (Reddy *et al.* 2012; Maurano

*et al.* 2015). While acquisition of new TF binding motifs in non-coding regions have been shown to create aberrant enhancers that upregulate key oncogenic targets, as observed in T cell acute lymphoblastic leukaemia (T-ALL) patient samples and cell lines at different loci (Mansour *et al.* 2014; Abraham *et al.* 2017).

The transactivating ability of TFs can also be affected indirectly by increases in co-factor concentration, and directly by acquired activating mutations. Diverse examples have been observed. In the example of MYB, truncating mutations that result in deletions at both the 5' and 3' end of the gene, lead to disruption of specific regions in the DB and negative regulatory (NR) domains responsible for attenuating its activity, such as truncating mutations induced by avian viruses resulting in the transforming v-Myb oncoprotein in chicken leukaemia (Rosson 1990). Similarly, alternative splicing of MYB has been shown to generate proteins with increased transactivation potential and these isoforms have been found in patient leukaemia samples, as well as different haematopoietic populations during differentiation (Shen-Ong *et al.* 1990; O'Rourke *et al.* 2008). Finally, chromosomal translocations involving multiple TFs have been described that result in constitutively active fusion proteins that drive the disease phenotype. Examples include MYB-NKIB in solid tumours (Persson *et al.* 2009) and TCRB-MYB in a subset of T-ALLs (Clappier *et al.* 2007). The effect of aberrant TF activity is likely to be further exacerbated by interactions with co-factors and chromatin regulators.

Therefore, as TFs play such a pivotal role in the execution of transcription programs, dissecting their function is key to understanding disease processes and develop targeted therapies.

## 1.2 *MLL* rearranged leukaemia

The *Mixed-Lineage Leukaemia* (*MLL*) gene (also known as *KMT2A*) is the human homologue of Drosophila *Trithorax*. It encodes a methyltransferase responsible for methylation of histone H3 lysine residue 4 and is particularly important in the positive regulation of gene expression during development (Parry *et al.* 1993; Mbangkollo *et al.* 1995; Marschalek *et al.* 1997). *MLL* is a recurring site for

cytogenetic alterations found in human acute leukaemia and indeed, was first identified as the gene that spanned the breakpoint in 11q23 translocated leukaemias (Ziemin-van der Poel *et al.* 1991; Cimino *et al.* 1991; Djabali *et al.* 1992; Tkachuk *et al.* 1992; Gu *et al.* 1992).

The most frequently occurring *MLL* mutations are chromosome translocations that fuse the *MLL* gene in-frame with a second partner gene, creating novel MLL-fusion proteins (MLL-FPs). Over 121 partner genes have been identified, but the most common fusion partners are *AF9* (ALL-1 fused gene from chromosome 9), *AF10* (ALL-1 fused gene from chromosome 10), *ELL* (eleven-nineteen lysine-rich leukaemia), *AF6* (ALL-1 fused gene from chromosome 6) and *ENL* (eleven-nineteen leukaemia), which account for 80-90% of all *MLL* gene translocations (C. Meyer *et al.* 2018). As its name suggests, *MLL* gene rearrangements can give rise to acute leukaemias with a myeloid (AML), lymphoid (ALL), or mixed phenotype (MLL).

### 1.2.1 *MLL* rearrangements in acute myeloid leukaemia

Acute myeloid leukaemia (AML) is an aggressive, genetically heterogeneous disease characterised by the clonal expansion and accumulation of immature myeloid precursors, impaired haematopoiesis and bone marrow failure. Leukaemogenesis typically arises from the acquisition of a small number of mutations that ultimately lead to the execution of aberrant transcriptional programs (Thoms *et al.* 2019).

AML can be classified into distinct biological subsets based on the characterisation of karyotypic and molecular abnormalities, which is the strongest independent predictor of prognosis (Papaemmanuil *et al.* 2016; H. Döhner *et al.* 2017). In AML, mutations involving *MLL* occur in as many as 35-50% of infant, 18% of childhood and 10% of adult *de novo* cases (Pui *et al.* 2011; Muntean *et al.* 2012; Cancer Genome Atlas Research Network *et al.* 2013). *MLL* rearrangements (*MLL*r) are also implicated in secondary leukaemias that arise in patients following treatment with topoisomerase-II inhibitors for other malignancies (Huret *et al.* 2001; Krivtsov *et al.* 2007). Other *MLL* abnormalities in AML include partial tandem duplications (PTDs) of exons encoding the N-terminus of MLL in approximately 10% of AML

with normal karyotype (K. Döhner *et al.* 2002; Papaemmanuil *et al.* 2016). Wild type *MLL* is also occasionally duplicated or amplified in myelodysplastic syndrome (MDS) and AML, usually in the context of complex karyotypic abnormalities and therapy-related disease (Andersen *et al.* 2001; Zatkova *et al.* 2004; Zatkova *et al.* 2009).

*MLL* rearrangements and amplifications in AML are associated with intermediate and unfavourable prognoses, and despite risk-adapted treatment approaches, 5-year overall survival rates remain between 4 and 39% (Grimwade *et al.* 2010; Papaemmanuil *et al.* 2016; Estey 2018). Mortality is usually due to relapsed or refractory disease (H. Döhner *et al.* 2015).

### 1.2.2 The mutational landscape of *MLL*r leukaemia

In contrast to other malignancies, *MLL*r leukaemias are genetically simple diseases. Sensitive Next Generation Sequencing (NGS) of *MLL*r patient samples shows a low frequency of additional genetic mutations (from 1-4, mean 1.3), in 40-50% of cases (Cancer Genome Atlas Research Network *et al.* 2013; A. K. Andersson *et al.* 2015). The most frequently occurring co-operating mutations are activating *RAS* pathway and *FLT3* mutations, however, these are invariably subclonal, with variant allele frequencies consistently below 40% (Driessen *et al.* 2013; A. K. Andersson *et al.* 2015; Emerenciano *et al.* 2015; Lavallee *et al.* 2015). The redundancy for co-operating mutations suggests that MLL-FPs alone are potent oncogenes, likely to be the driver mutations in these leukaemias.

### 1.2.3 *MLL*r leukaemia - a model for leukaemogenesis

MLL-FPs alone can cause aggressive acute leukaemia in mouse model systems and xenograft assays. MLL-AF9 is the most common fusion protein in *MLL*r AML (C. Meyer *et al.* 2018) and is the most well studied in experimental models.

The first *in vivo* MLL-AF9 models were developed by the Rabbitts laboratory, and include an *Mll-AF9* knock-in model (Corral *et al.* 1996) and an elegant *in vivo* translocator model (Collins *et al.* 2000). In the translocator model, utilisation of engineered *loxP* sites were inserted into the introns of endogenous *Mll* and *Af9* and

corresponding to the most common translocation break points in patients. Cre-recombinase expression was driven by the haematopoietic lineage-specific promoters, *Lmo2* and *Lck*, resulting in tissue-targeted Mll-Af9 translocations. Both the knock-in and translocator systems were capable of producing AML, either from constitutive expression of knock-in Mll-AF9 in haematopoietic cells of chimeric or heterozygous mice (Corral *et al.* 1996; Dobson *et al.* 1999), or from primitive progenitor cells driven by *Lmo2-Cre* in the translocator model (Collins *et al.* 2000; Drynan *et al.* 2005; Forster *et al.* 2005). Of note, in the translocator model, *Mll-Af9* translocations in more lineage-restricted T cell progenitors (driven by *Lck-Cre*) did not to give rise to leukaemia (Dobson *et al.* 1999; Forster *et al.* 2005), while in the knock-in model, Mll-AF9 expression specifically in foetal liver cells could occasionally also give rise to ALL (W. Chen *et al.* 2011). This suggests that 1) the developmental stage (foetal vs adult) and 2) differentiation hierarchy (stem vs progenitor vs precursor) of the cell of origin plays an important role in leukaemic phenotype.

The impact of foetal vs adult cell of origin is somewhat addressed by xenograft models of retrovirally transduced human CD34+ cord blood (CB) cells (Barabe *et al.* 2007; Wei *et al.* 2008), as well as the experiments mentioned above on Mll-AF9 knock-in foetal liver cells. CB transduction has given rise to the full range of leukaemia phenotypes – AML, B-ALL (Barabe *et al.* 2007) and bi-phenotypic leukaemia (Wei *et al.* 2008). Recently, in human CD34+ CB cells, CRISPR-Cas9 mediated insertion of *AF9* complementary DNA (cDNA) to create an MLL-AF9 transgene (Buechele *et al.* 2015), as well as CRISPR-Cas9 mediated MLL-AF9 translocations (Jeong *et al.* 2019), were also able to generate the full range of leukaemia phenotypes. This is compelling evidence that the target cell environment plays a large role in disease phenotype, where more primitive cells either of foetal origin or high up the haematopoietic hierarchy are able to produce mixed phenotypes, while adult haematopoietic stem progenitor cells (HSPCs) produce more lineage-restricted leukaemia. This may explain the distribution of MLL-AF9 phenotypes observed in patients, where the frequency of MLL-AF9 in AML increases with age (C. Meyer *et al.* 2018).

Viral transduction models have been developed as an alternative strategy to better define the impact of haematopoietic hierarchy on MLL-AF9 driven leukaemogenesis. These experiments indicate that Lin- Sca-1+ c-Kit+ (LSK), granulocyte macrophage progenitor (GMP) (Krivtsov *et al.* 2006) and c-Kit+ bone marrow (BM) cells (Somervaille *et al.* 2009) are all capable of being transformed into AML, but interestingly, the target cell type determines disease severity. In comparison to GMPs, transformed LSKs give rise to a more aggressive AML and bear a gene expression signature that correlates with poor prognosis in patients. This suggests that the cell of origin can determine leukaemia prognosis (Krivtsov *et al.* 2013). These findings are complemented by an *in vivo* doxycycline inducible MLL-AF9 mouse system in which the fusion protein is expressed at physiological levels and consistently produces a more aggressive AML in long-term haematopoietic stem cells (LT-HSC) compared to GMPs. This again is supportive of the fact that a more immature target cell population correlates with a poorer prognosis (Stavropoulou *et al.* 2016).

Overall, while various MLL-AF9 models have not consistently captured the full repertoire of human disease phenotypes, showing an overall bias for AML, it has nonetheless been repeatedly observed that MLL-AF9 can faithfully transform adult HSPCs into AML and MLL-AF9 models are widely accepted as a well-characterised and powerful model for acute myeloid leukaemogenesis.

### 1.2.4 *MLL*r leukaemia - a transcriptional disease

MLL-FPs are thought to drive leukaemogenesis by binding to their target genes and increasing their expression via both transcriptional elongation and epigenetic mechanisms, thus disrupting normal transcriptional networks (Figure 1.1).

MLL-FPs may be recruited to their gene targets via the retained N-terminal portion of wild type MLL through several mechanisms. First, this region contains three AT hook motifs (Broeker *et al.* 1996; Slany *et al.* 1998) and a CXXC domain (Bach *et al.* 2009; Milne *et al.* 2010), both of which have been shown to contact DNA. Furthermore, the CXXC domain is likely stabilised at its target loci through

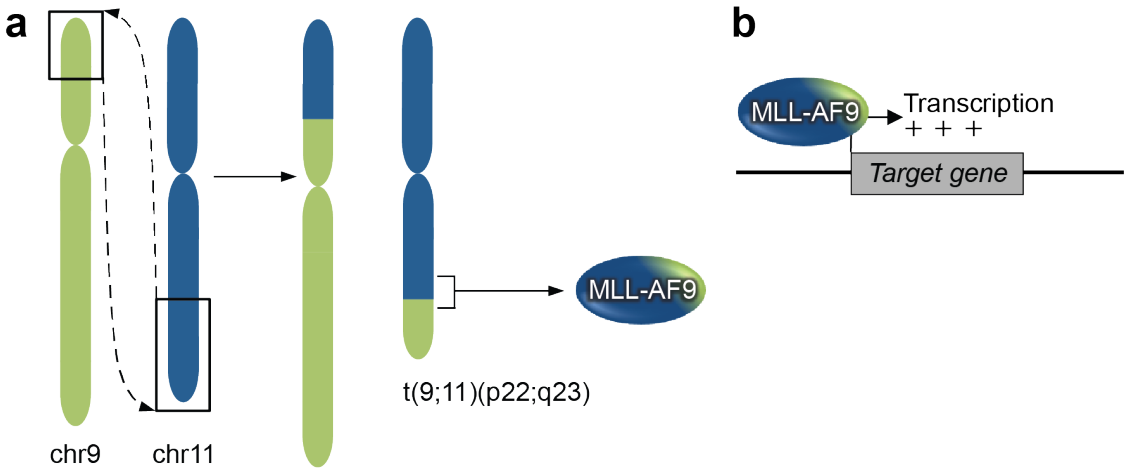



**Figure 1.1:** MLL-FPs drive aberrant gene expression. **a** Schematic of translocation between *MLL* and its fusion *AF9*, resulting in expression of the MLL-AF9 fusion protein. **b** Schematic of MLL-AF9 binding to and increasing expression of a gene target.

its interaction with the PAF1 complex which is involved in transcription elongation and is known to directly interact with RNAPII (Milne *et al.* 2010). Next, the N-terminus of MLL is also known to associate with LEDGF via direct interaction with Menin (Yokoyama *et al.* 2008), while LEDGF has been shown to bind to H3K36me2/3 via its PWWP domain (Hughes *et al.* 2004; Yokoyama *et al.* 2005; Milne *et al.* 2005; Y.-X. Chen *et al.* 2006; Caslini *et al.* 2007; Okuda *et al.* 2014). More recently, Menin has been shown to interact with the TA domain of MYB and thus mediate the association of MLL with chromatin binding transcription factors (Jin *et al.* 2010) (See Chapter 4.1.1). These interactions together are thought to facilitate MLL-FP binding to gene targets.

It is thought that MLL-FPs contribute to aberrant upregulation of their target genes predominantly via their fusion partners, which are able to facilitate the recruitment and stabilisation of 1) gene activation complexes and 2) chromatin proteins that positively regulate transcription.

A major mechanism of MLL-FP activation is to recruit the positive transcription elongation factor B (P-TEFb). P-TEFb has been shown to be responsible for phosphorylating serine-2 of the RNAPII C-terminal domain (CTD) heptapeptide repeats, as well as, DSIF and NELF, leading to the pause release of RNAPII and productive transcription elongation (Marshall *et al.* 1995; Marshall *et al.* 1996;

Wada *et al.* 1998; Yamaguchi *et al.* 1999). It has been purified together with many of the common MLL fusion partner proteins in several multimeric protein complexes, as listed in Table 1.1. The most well described are the ENL Associated Protein (EAP) complex (Mueller *et al.* 2007), the super elongation complex (SEC) (C. Lin *et al.* 2010) and its closely related AEP complex (AF4 family, ENL family, P-TEFb) (Yokoyama *et al.* 2010). Therefore, reciprocal stabilisation of elongation complex proteins via different MLL fusion partners may represent a common mechanism through which P-TEFb is constitutively stabilised, leading to aberrant upregulation of MLL-FP targets.

**Table 1.1:** Transcription elongation complexes related to MLL-FP function

| Complex | Components | Reference |
|---|---|---|
| AF4-associated | AFF1, AF9, DOT1L, ENL, MLLT10, P-TEFb | Bitoun *et al.* 2007 |
| EAP | AFF1/3/4, BCOR, CBX8, DOT1L, ENL, P-TEFb, RING1 | Mueller *et al.* 2007 |
| EAP core | AFF1, DOT1L, ENL, P-TEFb | Mueller *et al.* 2009 |
| AEP | AFF1/4, ENL, P-TEFb | Yokoyama *et al.* 2010 |
| SEC | AFF1/4, AF9, EAF1/2, ELL1/2/3, ENL, P-TEFb | C. Lin *et al.* 2010 |
| DotCom | AF9, CTNNB1, DOT1L, ENL, MLLT6, MLLT10, SKP1, TRRAP | Mohan *et al.* 2010 |
| AF9-associated | AF9, DOT1L (mutually exclusive with AF4/AFF4, AF9, P-TEFb) | Biswas *et al.* 2011 |

Underlined proteins are known MLL fusion partners.

In addition to P-TEFb recruitment, other fusion partner protein interactions that contribute to gene activation have been identified. During transcription initiation, the SEC/AEP has been shown to facilitate TBP loading (Okuda *et al.* 2015), likely via its interaction with AF9 (Biswas *et al.* 2011). Also, the Mediator subunit MED26, which is known to play a role in facilitating transition of RNAPII into the elongation stage of transcription, has been demonstrated to interact with AF9 and ENL, members of the SEC (Takahashi *et al.* 2011; Okuda *et al.* 2016). AF9, ENL and ELL individually have also been shown to contain C-terminal transcriptional activation (TA) domains (Prasad *et al.* 1995; Slany *et al.* 1998; Shilatifard *et al.*

1996) that are capable of promoting transcription elongation and in the cases of AF9 and ENL, are important for oncogenesis (Prasad *et al.* 1995; Slany *et al.* 1998).

Furthermore, accumulating evidence suggests that many MLL fusion partners positively regulate transcription through chromatin remodelling. The most well described being histone 3, lysine 79 methylation via the methyltransferase, DOT1L (Feng *et al.* 2002; Ng *et al.* 2002; Lacoste *et al.* 2002). DOT1L is able to interact directly with AF9, ENL, AF10 and AF17 (Okada *et al.* 2005; Reisenauer *et al.* 2009; Y. Li *et al.* 2014; Wan *et al.* 2017), and its presence in multiple elongation complexes suggests a mechanism for its recruitment in different MLL-FP leukaemias (Mueller *et al.* 2007) (see Table 1.1). Indeed, MLL-FP targets have been shown to have aberrantly high levels of H3K79me in the gene body relative to other non-MLL-FP targets with similar levels of gene expression (Krivtsov *et al.* 2008; Bernt *et al.* 2011). H3K79me is introduced during transcriptional elongation (Krogan *et al.* 2003) and has been shown to be essential for maintaining enhancer-promoter interactions at important MLL-FP targets (Godfrey *et al.* 2019).

Together, the evidence points to MLL-FP leukaemias being a transcriptional disease, while their comparative genetic simplicity means that they can be a powerful model to study how transcriptional dysregulation contributes to leukaemogenesis and leukaemia maintenance. Simultaneously, studying the direct downstream targets of MLL-FPs may yield therapeutic strategies to tackle this disease. One such downstream target is the master transcription factor, MYB, which has been shown to be indispensable for leukaemia maintenance (see Chapter 1.3).

## 1.3 MYB - a key regulator in normal and malignant haematopoiesis

### 1.3.1 The MYB family of transcription factors

*MYB* (also known as *c-Myb*) encodes a 75kDa DNA binding transcription factor, which acts as an activator as well as repressor of transcription, regulating genes important for lineage determination, cell proliferation, and differentiation. It was originally discovered as the normal cellular counterpart of the v-Myb transforming

proteins of avian leukaemia viruses, Avian myeloblastosis virus (Hall *et al.* 1941) and E26 (Ivanov *et al.* 1964; Nedyalkov *et al.* 1975) (as reviewed by Oh *et al.* 1999).

*MYB* is part of an evolutionarily conserved family of three transcription factors (Campanini *et al.* 2015) including *MYBL1* and *MYBL2* (also known as *A-Myb* and *B-Myb* respectively) (N. Nomura *et al.* 1988). They demonstrate highest homology in their DNA binding (DB) domains and appear to bind the canonical consensus sequence 5'-YAAC[GT]G-3' with overlapping sequence specificity (Biedenkapp *et al.* 1988; Mizuguchi *et al.* 1990; Howe *et al.* 1991; Golay *et al.* 1994). However, each gene exhibits distinct patterns of expression and function. In normal tissue development and homeostasis, besides the colon (Rosenthal *et al.* 1996; Zorbas *et al.* 1999) and brain (D. H. Shin *et al.* 2001; Malaterre *et al.* 2008), MYB is primarily expressed in the haematopoietic system (Mucenski *et al.* 1991), whereas MYBL1 expression is restricted to male germ cells and breast epithelial cells of pregnant mice (Mettus *et al.* 1994; Toscani *et al.* 1997), and at low levels in ovaries, developing brain (Trauth *et al.* 1994), as well as in germinal centre B-cells (Golay *et al.* 1994). In contrast to the tissue-specific expression of MYB and MYBL1, MYBL2 expression seems to be ubiquitous among proliferating cells (N. Nomura *et al.* 1988; Tanaka *et al.* 1999; Ness 2003).

### 1.3.2 MYB in normal haematopoiesis

*MYB* is a key regulator of the haematopoietic system. First and foremost, it is required for the development and maintenance of definitive haematopoiesis. Its importance is highlighted by the fact that *Myb* null mice do not survive beyond embryonic day 15 (E15) due to severe anaemia and the absence of most blood cell lineages (Mucenski *et al.* 1991).

Examination of constitutive and conditional *Myb* knock-out foetal liver at E15, confirms the presence of HSCs, LSK, and multipotent progenitor (MPP) populations. However, these cells are significantly reduced in number and are compromised in their ability to expand or give rise to mature colonies in culture (Mucenski *et al.* 1991; Sumner *et al.* 2000; Lieu *et al.* 2009). Furthermore, $Myb^{-/-}$ HSCs are unable to

contribute to the BM of lethally irradiated recipient mice in serial transplantation experiments (Lieu *et al.* 2009). This suggests that while Myb is not initially required for haematopoietic stem/progenitor development, it is required for their expansion, maintenance and subsequent multi-lineage differentiation. Interestingly, *Myb* heterozygous ($Myb^{+/-}$) mice are phenotypically indistinguishable from their wild-type litter mates and have similar BM populations and peripheral blood (PB) counts under steady-state conditions (Mucenski *et al.* 1991; Lieu *et al.* 2009). However, when exposed to γ-radiation or treated with 5-fluorouracil and required to proliferate, $Myb^{+/-}$ mice show delayed recovery as their haematopoietic stem cells are slower to re-enter into cycle, divide and generate colonies (Ramsay *et al.* 2004). Thus, for maintenance of normal haematopoiesis, we can conclude that 1) there is a threshold requirement for Myb, 2) reduced Myb levels (but not total absence) can be tolerated and 3) this threshold is higher under stress conditions.

From various experimental models, it is apparent that precise regulation of Myb expression is important for multi-lineage differentiation, and that the level required is lineage and stage specific. One such model utilises a knockdown allele of *Myb* ($Myb^{kd}$) that expresses ∼5-10% of the wild-type level of protein. When crossed onto a wild type or *Myb* null background, differential allelic segregation results in mice expressing distinct threshold levels of Myb, thus allowing for dosage dependent analysis of Myb function (Emambokus *et al.* 2003). Importantly, while $Myb^{-/kd}$ mice expressing ∼5-10% of the wild-type level of Myb only survived to E16 *in utero*, $Myb^{kd/kd}$ mice expressing approximately 20%, although not as long-lived as wild type or heterozygous mice, were able to survive into adulthood.

Compared to wild type, $Myb^{-/kd}$ and $Myb^{kd/kd}$ E15 foetal livers were hypocellular and demonstrated an abnormal accumulation of progenitor cells with accelerated terminal differentiation affecting specific lineages. In contrast to the absence of progenitor proliferation seen in *Myb* null cells, this indicates that low expression of Myb is sufficient for progenitor expansion but that a higher threshold is required for normal multi-lineage differentiation. Lower Myb expression also correlated with a loss of granulocytic potential, decrease in erythroid potential,

increase in monocyte/macrophage containing colonies and a significant skew towards megakaryopoiesis (Emambokus *et al.* 2003). Furthermore, data from $Myb^{kd/kd}$ mice, and that from other hypomorphic loss of function models, suggest that suboptimal levels of Myb are sufficient for megakaryopoiesis, and lead to a megakaryocytic and monocytic/macrophage differentiation bias at the expense of other cell lineages (Mucenski *et al.* 1991; Emambokus *et al.* 2003; Sandberg *et al.* 2005; Garcia *et al.* 2009). Additionally, low levels of Myb expression appear to be sufficient for erythroid progenitor expansion, but higher levels are required for their progression beyond the proerythroblast stage (Emambokus *et al.* 2003; Vegiopoulos *et al.* 2006).

Conditional knockout and dominant negative mouse models, as well as $Myb^{-/-}$ ESC/ $Rag1^{-/-}$ blastocyst chimeras have shown that discrete Myb expression is required at different stages of thymopoiesis (Badiani *et al.* 1994; R. D. Allen *et al.* 1999; Bender *et al.* 2004; Lieu *et al.* 2004). The most primitive T progenitors can develop in the absence of Myb (R. D. Allen *et al.* 1999). However, Myb is essential at three subsequent stages of T-cell development: progression from DN2 to DN3 cells, survival of pre-selection double-positive (DP) cells and generation of single-positive CD4+ cells (Bender *et al.* 2004; Lieu *et al.* 2004).

Of the various haematopoietic lineages, B cell development is most sensitive to loss of Myb. There is a near absence of B progenitors in $Myb^{-/kd}$ and $Myb^{kd/kd}$ mice, and those present display a largely immature phenotype (Emambokus *et al.* 2003). Targeted deletion of *Myb* has shown that it is required for pro-B to pre-B transition and for survival of splenic B cells (Thomas *et al.* 2005; Fahl *et al.* 2009).

Conversely, forced expression of Myb has been shown to prevent the terminal differentiation of erythrocytes and megakaryocytes, and lead to complete loss of B lymphopoiesis, clearly illustrating that stage appropriate downregulation of Myb is required for terminal differentiation (Sakamoto *et al.* 2006).

Consequently, it appears that precise regulation by Myb is required to effect haematopoietic homeostasis and it is unsurprising that perturbation of Myb expression is observed in many malignancies, including *MLL*r AML.

### 1.3.3 MYB in leukaemia

MYB overexpression has been observed in many malignancies, including solid tumours such as colorectal, breast and pancreatic cancers and is associated with a poor prognosis (Biroccio *et al.* 2001; Persson *et al.* 2009). *MYB* has also been shown to undergo rearrangement or translocation, leading to aberrant expression, and frequently constitutive activation in human lymphomas and leukaemias (Göran Stenman 2010; Pattabiraman *et al.* 2013). Furthermore, *MYB* has been shown to be a direct target of MLL-FPs and is overexpressed in this subset of leukaemias (Zuber *et al.* 2011).

MYB has been implicated as an important gene in *MLL*r leukaemia initiation. In a tamoxifen-inducible MLL-ENL system, *Myb* upregulation was directly related to the presence of MLL-ENL, and MLL-ENL-mediated transformation was inhibited by shRNA (short hairpin RNA) knockdown of *Myb* mRNA. Complementary to this, MLL-ENL transduced haematopoietic precursors from $Myb^{kd/kd}$ (expressing ~20% of the wild-type level of Myb) were unable to give rise to leukaemia *in vitro*, showing reduced and diffuse colonies after 3 rounds of serial replating (Hess *et al.* 2006). This is one of the few studies to link *Myb* to a pathway utilised during leukaemia transformation. However, while it provides some indication that leukaemic transformation is Myb dependent, the data is confounded by the fact that shRNA knockdown of *Myb* may be temporally disassociated with the onset of MLL-ENL mediated transformation, therefore it is not clear if the loss of Myb is affecting leukaemia initiation, or in fact leukaemia maintenance. Furthermore, Myb levels are already attenuated in $Myb^{kd/kd}$ cells prior to MLL-ENL expression, and these cells intrinsically possess aberrant clonogenic and differentiation potential (see Chapter 1.3.2). Therefore, we may simply be observing a lack of transformation due to an altered target cell environment. Consequently, the requirement for MYB in the process of leukaemia initiation and transformation is still unclear, in part because systems to ensure that robust MYB repression only occurs at the onset of leukaemia initiation have not been available.

In contrast, there is clear evidence that MYB is important for the maintenance of leukaemia. Early experiments using antisense oligonucleotides revealed sensitivity of leukaemia cell lines, and primary human AML and CML blasts to *MYB* inhibition (Anfossi *et al.* 1989; Calabretta *et al.* 1991), implying it is necessary for leukaemia maintenance. These were the first studies to suggest that myeloid leukaemias may be more sensitive to MYB inhibition than normal haematopoietic cells. These findings are further supported by shRNA experiments by Zuber *et al.* (2011), showing that MYB is necessary for the maintenance of multiple AML subtypes including MLL-AF9 leukaemia, and even partial and transient suppression of MYB leads to abrogation of leukaemia *in vivo*. Importantly, this same study described preserved myelopoiesis in normal haematopoietic cells despite a relative reduction of *Myb* mRNA levels by approximately 90% (Zuber *et al.* 2011). Taken together, these data suggest that leukaemia cells are more sensitive to decreased MYB compared with normal cells. Therefore a therapeutic window may be achieved through understanding MYB mediated pathways for leukaemia maintenance.

But how does MYB effect transcriptional dysregulation in disease? Despite its obvious importance in normal and malignant haematopoiesis, much is still unknown regarding the molecular function of MYB. Recent evidence from disease models suggests that MYB functions through establishing aberrant enhancer activity. In some T-ALLs, acquired somatic mutations in a precise non-coding site introduce a novel binding motif for MYB. This results in an active upstream enhancer that drives expression of the TAL1 oncogene. MYB binding at this locus is correlated with the presence of its acetylase-binding partner CBP, increased H3K27ac and core components of the TAL1 leukaemogenic transcriptional complex (Mansour *et al.* 2014). This suggests that MYB binding may be sufficient for enhancer activity at specific loci, functioning in part by mediating H3K27-acetylation and facilitating combinatorial transcription factor binding and raises the possibility of its requirement for establishing enhancer-promoter interactions.

## 1.4 Summary and aims

It is evident that MYB plays the role of master regulator in both normal and malignant haematopoiesis. However, key questions still remain in the field regarding its biological requirement in leukaemic transformation and its molecular function in gene regulation. A deeper understand of these processes will serve to further the field and hopefully contribute to the development of more targeted cancer therapies. Therefore, utilising a clinically relevant MLL-AF9 leukaemia model, this thesis aims to address the following questions:

1. Is MYB required for leukaemia initiation?
2. Is MYB binding sufficient for establishing an enhancer?
3. Is MYB binding necessary for maintaining enhancer function?

# 2

# Materials and methods

## Contents


**2.1 Mouse studies . . . . . 25**
2.1.1 Generation of *Myb*$^{MRE/MRE}$ mouse line . . . . . 25
2.1.2 Genotyping of mESCs and mice . . . . . 26
2.1.3 Peripheral blood analysis . . . . . 26
2.1.4 Dissection and preparation of tissue . . . . . 27
2.1.5 Immunophenotypic analysis and cell sorting . . . . . 27
2.1.6 Cell purification using magnetic-activated cell sorting . 28
2.1.7 Retroviral transduction and serial replating . . . . . 29
2.1.8 Functional assays . . . . . 30
**2.2 Cell lines and cell culture techniques . . . . . 31**
2.2.1 Cell line culture conditions . . . . . 31
2.2.2 Generation of TetR protein expressing mESC lines . . . 31
2.2.3 Colony forming assays . . . . . 32
**2.3 RNAi assays . . . . . 32**
**2.4 Protein assays . . . . . 33**
2.4.1 Western blotting . . . . . 33
2.4.2 Intracellular FACS . . . . . 34
**2.5 Chromatin immunoprecipitation (ChIP) assays . . . . 34**
2.5.1 ChIP qPCR . . . . . 36
2.5.2 ChIP-sequencing . . . . . 36
2.5.3 Reference normalised ChIP-sequencing (ChIP-rx) . . . . 36
**2.6 RNA assays . . . . . 37**
2.6.1 Gene expression analysis . . . . . 37
2.6.2 Nascent RNA-sequencing . . . . . 37
2.6.3 Poly(A) minus RNA-sequencing . . . . . 38
**2.7 ATAC-seq . . . . . 38**
**2.8 Capture-C . . . . . 39**
**2.9 Bioinformatic analysis . . . . . 40**
2.9.1 Sequence analysis . . . . . 40
2.9.2 Locus Overlap Analysis . . . . . 41
2.9.3 RNA-seq and gene expression analysis . . . . . 41
2.9.4 Capture-C analysis . . . . . 41

## 2.1 Mouse studies

All mice used in this study were bred and maintained according to UK Home Office regulations and standards. C57BL/6J mice were sourced from MRC Harwell/Charles River Laboratories. The transgenic mouse strains reported in this study were maintained on a C57BL/6J background and were generated and phenotyped in accordance with Animal Scientific Procedures Act 1986, with procedures reviewed by the Animal Welfare and Ethical Review Body (AWERB), and conducted under project licences PPL 30/2945 and PPL PA7C92A40.

### 2.1.1 Generation of *Myb*$^{\mathbf{MRE/MRE}}$ mouse line

For the generation of the *Myb* knock-in mouse embryonic stem cells (mESCs), sgRNA designed to target Cas9 to the sequence 5-acagagccggaactgcctcg-3' within the *Myb* 3'UTR, was cloned into the vector pX330 (Addgene plasmid, 42230). 1 x $10^6$ mESC were transfected with the sgRNA expression plasmid (5µg) and the ssODN donor (2µL, 100µM stock) using the Neon Transfection System (Thermo Fisher Scientific) in a 10µL tip (1150 V, 20 ms, 2 pulses). Following transfection, cells were plated on 0.1% gelatin and colonies picked 24-48 hours after transfection for genotyping.

For the generation of the *Myb* knock-in mouse models, C57BL/6J zygotes were microinjected with 20 ng/µl of sgRNA, 10 ng/µl NLS-Cas9 mRNA and 20ng of ssODN. sgRNA designed to target Cas9 to the sequence 5-acagagccggaactgcctcg-3' within the *Myb* 3'UTR, was synthesized by *in vitro* transcription using the MEGAshortscript$^{\text{TM}}$ T7 Transcription Kit (Thermo Fisher Scientific, Cat no. AM1354) from a DNA template. Capped mRNA for NLS-Cas9 was generated by firstly cloning the NLS-Cas9 cDNA from pX330 into pcDNA3.1, linearising the plasmid with XhoI and using this as a template for *in vitro* transcription using the mMESSAGE mMACHINE$^{\text{TM}}$ T7 Ultra Transcription Kit (Thermo Fisher Scientific, AM1345). *In vitro* transcribed RNAs were purified using the MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific, AM1908) and diluted prior to microinjection in 10 mM Tris.HCl pH7.5, 0.1 mM EDTA pH8.0. Microinjected

zygotes were cultured overnight to the two-cell stage and surgically implanted into recipient pseudopregnant CD1 females.

Founder mice were identified by genotyping using PCR (see Chapter 2.1.2 for details) amplifying a region of the 3' UTR, followed by Sanger sequencing to detect the incorporation of the template sequence.

### 2.1.2 Genotyping of mESCs and mice

For mESCs, DNA was purified using the QIAGEN DNA Blood and Tissue kit (QIAGEN, 69504) according to the manufacturer's protocol.

Ear punch tissue samples from mice were lysed in 100µL lysis buffer (Tris-HCL 1M 2.5 mL, EDTA 0.5M 0.1 mL, Tween-20 0.25 mL made up to 50mL with water) and proteinase K treated for 1-2 hours at 65°C, followed by 95°C heat inactivation for 5 minutes.

DNA obtained as above was directly used as a template for PCR. Primers spanning either side of the insertion site were used in the following conditions: (98°C 30s, 39 cycles of 98°C 10s, 62°C 15s, 72°C 15s, then 72°C 5min). Primer sequences are as follows: 5'-gggagatttgtgttgtttatgtca-3' (forward) and 5'-ggcagaaactggctgttga-3' (reverse). The wild type allele is detected as a band of 186bp, and the edited allele is 233bp.

### 2.1.3 Peripheral blood analysis

Peripheral blood was collected via tail vein sampling or via cardiac puncture at the time mice were euthanised, into EDTA-treated (Sarstedt, 20.1288) or heparin-containing (Sarstedt, 20.1292) Microvette® tubes. Blood counts were obtained using a Sysmex KX-21N 3-part differential automated haematology analyser prior to downstream analysis.

For analysis of peripheral white blood cell engraftment or following transplantation with virally transduced donor cells, samples were resuspended in 9mL of lysis buffer containing ammonium chloride (150mM) and potassium bicarbonate (10mM) for 3 minutes at room temperature. 1mL of 10X phosphate-buffered saline

(PBS) was added to stop the lysis reaction and samples were centrifuged (500g, 5 minutes, 4°C), the supernatant removed and intact cells resuspended in PBS/5% fetal calf serum (PBS/FCS). Following the lysis reaction, cells were blocked and stained for immunophenotypic analysis as described in Chapter 2.1.5. Data were analysed and graphs generated using Prism GraphPad for Mac (v8.0).

### 2.1.4 Dissection and preparation of tissue

For all experiments where haematopoietic and leukaemic cell populations were used, the following steps were carried out to isolate tissue samples. Mice were euthanized using $CO_2$, and confirmation of death was achieved by cervical dislocation as per schedule 1 of the Home Office licensing requirements. Mice were placed under a sterile containment hood and doused with 70% ethanol solution. Tibiae, femurs ± cristae and vertebrae were dissected, crushed using a pestle and mortar, then flushed through a 50µm filter with 10mL of filtered PBS/FCS. Splenic tissue was dissected, weighed, imaged and crushed using a syringe plunger through a 50µm filter, then flushed through with PBS/FCS to create a single cell suspension. Cell counts were obtained using a Sysmex KX-21N 3-part differential automated haematology analyser prior to downstream analysis.

### 2.1.5 Immunophenotypic analysis and cell sorting

All samples were first incubated in purified anti-mouse CD16/32 to prevent non-specific antibody binding, with the exception of the staining of myeloerythroid progenitor populations where samples were blocked in fluorophore (PE-Cy7) conjugated anti-CD16/32. Incubations took place in the dark at 4°C for 20 minutes. Next, an equal volume of 2X concentrated antibody cocktail was added to the cell solution and incubated at 4°C for 20 minutes in the dark. Cells were then washed, resuspended in PBS/FCS and either DAPI or 7-aminoactinomycin D (7AAD) (Sigma-Aldrich, A9400) viability dyes were added to each sample prior to analysis/sorting. Antibody staining panels are listed in Tables 2.1, 2.2, 2.3 &

2.4. A complete list of antibodies used can be found in Chapter 7, Table 7.1 and used at previously titrated concentrations.

Stained cells were analysed on either the BD LSRFortessa/BD LSRFortessa X20, or sorted using the BD FACSAriaII/III. Single stained compensation beads and fluorescence minus one (FMO) controls for each fluorophore were used for each experiment. For cell sorting experiments, purity was determined prior to collection, proceeding only if above 95%. Flow cytometric analysis was carried out using FACSDiva v8.1 (BD Biosciences) and FlowJo (v10.6.1) software. Data were analysed and graphs generated using Prism GraphPad for Mac (v8.0).

**Table 2.1:** Immunophenotypic markers used to define HSC and multipotent progenitor populations as defined by Cabezas-Wallscheid *et al.* 2014.

| **Cell type** | **Cell surface marker** |
|---|---|
| HSC | Lin- Sca1+ c-Kit+ CD34- CD48- CD150+ CD135- |
| MPP1 | Lin- Sca1+ c-Kit+ CD34+ CD48- CD150+ CD135- |
| MPP2 | Lin- Sca1+ c-Kit+ CD34+ CD48+ CD150+ CD135- |
| MPP3 | Lin- Sca1+ c-Kit+ CD34+ CD48+ CD150- CD135- |
| MPP4 | Lin- Sca1+ c-Kit+ CD34+ CD48+ CD150- CD135+ |

**Table 2.2:** Immunophenotypic markers used to define myeloerythroid progenitors as defined by Pronk *et al.* 2007.

| **Cell type** | **Cell surface marker** |
|---|---|
| MkP | Lin- c-Kit+ CD150+ CD41+ |
| GMP | Lin- c-Kit+ CD41-CD150- CD16/32+ |
| Pre MegE | Lin- c-Kit+ CD41- CD16/32- CD150+ CD105- |
| Pre CFU-E | Lin- c-Kit+ CD41- CD16/32- CD105+ CD150+ |
| CFU-E & Pro Ery | Lin- c-Kit+ CD41- CD16/32- CD150- CD105+ |
| Pre GM | Lin- c-Kit+ CD41- CD16/32- CD150- CD150- |

### 2.1.6 Cell purification using magnetic-activated cell sorting

The Lin- BM fraction was purified using the Direct Lineage Cell Depletion Kit (Miltenyi Biotec, 130-110-470) according to manufacture's instructions. BM cells were incubated with microbeads conjugated with monoclonal antibodies against CD5, CD45R (B220), CD11b, Ly-6G/C (Gr-1), 7-4, and Ter-119 and then separated

**Table 2.3:** Immunophenotypic markers used to define lymphoid progenitors as defined by Buono *et al.* 2016.

| Cell type | Cell surface marker |
|---|---|
| NK/T progenitors | Lin– B220- CD49b- NK1.1- CD122+ |
| LMPP | Lin- B220- CD49b- NK1.1- CD122- Flt3+ CD127- c-Kit$^{hi}$ Sca1$^{hi}$ |
| ALP | Lin- B220- CD49b- NK1.1- CD122- Flt3+ CD127+ Ly6D- c-Kit$^{low}$ |
| BLP | Lin- B220- CD49b- NK1.1- CD122- Flt3+ CD127+ Ly6D+ c-Kit- |
| CLP | Lin- B220- CD49b- NK1.1- CD122- Flt3+ CD127+ Ly6D- c-Kit$^{low}$ Sca1$^{low}$ |

**Table 2.4:** Immunophenotypic markers used to define peripheral blood lineages

Immunophenotypic markers used to define peripheral blood lineages.

| Cell type | Cell surface marker |
|---|---|
| T cell | CD4+ and/or CD8+ |
| B cell | CD19+ |
| Myeloid | NK1.1- Mac1+ |

using LD columns (Miltenyi Biotec, 130-042-901) in the magnetic field of a suitable MACS separator, retaining the flow-through containing the Lin- fraction.

### 2.1.7 Retroviral transduction and serial replating

*MLL-AF9/mCherry* in a retroviral vector was cloned using in-house constructs, and viral particles were generated by co-transfection of 293T cells with the MLL-AF9/mCherry construct and Ecotropic packaging plasmids in the presence of *Trans*IT®-LT1 reagent (Mirus Bio LLC, MR2300). Viral supernatant was harvested 48, 72 and 96 hours following transfection, pooled, concentrated using Retro-X$^{TM}$ reagent (Takara, 631456) as per manufacturer's instructions and stored as a 100X suspension in 50-75µL aliquots at -80°C.

Following purification, 5 x $10^4$ to 5 x $10^5$ GMP or Lin- cells were incubated in Opti-MEM$^{TM}$ (Gibco$^{TM}$, 51985034), supplemented with 15% FCS (Gibco$^{TM}$), 10ng/µL mSCF, 5ng/µL mIL-3 and 5ng/µL mIL-6 (STEMCELL$^{TM}$ technologies), 1 x penicillin/streptomycin (Gibco$^{TM}$, 15140122) overnight. The following morning, 1

x $10^3$ cells/µL retroviral supernatant and 7µg/mL Polybrene (final concentration) was added to the cells. Samples were then spinoculated at 1500 G for 90 mins at 37°C. 48 hours after infection, mCherry expressing cells were sorted using the BD FACSAriaII/III. 1 x $10^3$ to 1 x $10^4$ cells were seeded in 1mL of MethoCult™ GF M3434 (STEMCELL™ technologies) supplemented with 1 x penicillin/streptomycin. Where possible, biological replicates were plated out in triplicate. Every 7 days, colonies were counted, solubilised with 10ml Iscove Modified Dulbecco Media (IMDM) (Gibco™, 21980032) supplemented with 10% FCS, and 1-5 x $10^3$ cells re-seeded 1mL MethoCult™ GF as above.

## 2.1.8 Functional assays

### 2.1.8.1 Competitive repopulation assays

Unfractionated BM cells from 8 week old CD45.2 *Myb*$^{MRE/MRE}$ or wild type (WT) mice were mixed in a 1:1 ratio with unfractionated competitor BM cells from CD45.1/2 WT mice and transplanted via tail vein injection into age-matched lethally irradiated (2 x 5Gy) CD45.1 WT recipient mice. Peripheral blood analysis was performed every four weeks following transplantation to monitor engraftment and end point BM analysis was performed at 18 weeks post-transplantation.

### 2.1.8.2 Transduction-transplantation assays

Purified bone marrow populations from 8 week old CD45.2 *Myb*$^{MRE/MRE}$ or WT mice were transduced with retrovirus and expanded via serial replating for three weeks as described above 2.1.7. Prior to transplantation, plated cells were solubilised and 1 x $10^6$ donor cells (CD45.2) and 2.5 x $10^5$ unfractionated BM helper cells (CD45.1) were then injected via the tail vein into lethally irradiated CD45.1 WT recipient mice. Peripheral blood analysis was performed four weeks following transplantation and fortnightly thereafter to monitor engraftment. Mice that fulfilled previously defined humane end-point criteria based on clinical and laboratory parameters were sacrificed under Schedule 1. At the time of sacrifice, peripheral blood, tibiae, femurs and splenic tissue were harvested for downstream analysis.

## 2.2 Cell lines and cell culture techniques

### 2.2.1 Cell line culture conditions

Mouse bone marrow cells were cultured in Opti-MEM™ (Gibco™, 51985034), 10%FCS, 100Units/mL Penicillin, 100µg/mL Streptomycin (Gibco™, 15140122), SCF (5ng/mL), IL3 (5ng/mL), IL6 (10ng/mL) (STEMCELL™ technologies), 20% FCS and GlutaMax™ (Gibco™, 35050061).

293T cells were purchased from ATCC and cultured in DMEM (Gibco™, A12860-01) with 10% FCS and GlutaMax™ (Gibco™, 35050061).

E14 (TOT2N variant) mESCs were a gift from the Klose lab and cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco™, A12860-01) supplemented with 10% FCS, 100Units/mL Penicillin, 100µg/mL Streptomycin (Gibco™, 15140122), LIF (produced in-house), MEM Non-Essential Amino Acids (NEAA, Gibco™, 11140050) and β-mercaptoethanol (Gibco™, 31350010).

THP1 cells were purchased from ATCC (www.lgcstandards-atcc.org) and cultured in RPMI-1640 (Gibco™, 21875034) with 10% FCS and GlutaMax™ (Gibco™, 35050061).

### 2.2.2 Generation of TetR protein expressing mESC lines

E14 (TOT2N variant) mESCs were grown to 60% confluency on 0.1% gelatin in a 10cm dish. Media was removed and the cells were washed using PBS. Cells were replenished with 9ml of fresh media. 28µg of TetR expression plasmid DNA and 80µl of Lipofectamine 2000 (Invitrogen™, Life Technologies, 11668-019) were mixed together with 1ml of Opti-MEM™ (Gibco™, 51985034). This mixture was incubated for 5 minutes at room temperature. Following incubation, the mixture was gently pipetted onto mESCs. Cells were incubated for 48 hours and were either harvested for downstream applications or selected to make a stable cell line. For stable selection, 1µg/ml of puromycin (Sigma-Aldrich, P8833) was added to the cells. After 24 hours, this selection media was removed, cells were washed with PBS and media and puromycin replaced. Puromycin selection was

continued for 7-10 days until colonies formed. Colonies were picked and resuspended into 150µl media (+1µg/ml puromycin) in a gelatinised 96-well dish. Cells were passaged and expanded to a 6-well dish. Western blot analysis was performed to verify the expression of the integrant.

### 2.2.3 Colony forming assays

24 hours following the second transfection with siRNA (see Chapter 2.3), THP-1 cells were plated at a density of 500 cells/mL in IMDM diluted MethoCult$^{TM}$ H4100 (STEMCELL$^{TM}$ Technologies) supplemented with 20% FCS and cultured at 37°C, 5% $CO_2$ for 14 days before counting. Colony-forming assays were run in triplicate with at least three biological repeats.

## 2.3 RNAi assays

In transgenic *Myb*$^{MRE/MRE}$ mESCs, knockdown assays using *miR-196-b* mimic were performed using the NEON Transfection System (Thermo Fisher Scientific) in a 10µL tip (1150 V, 20 ms, 2 pulses). 2µL of 20µM miRNA was added to 1 x $10^6$ mESC per transfection. Following transfection, cells were plated on 0.1% gelatin for 24 hours before harvesting for gene expression analysis.

In THP-1 cells, siRNA knockdown assays were performed using a rectangle pulse EPI 2500 electroporator (Fischer Heidelberg, Germany (Martinez *et al.* 2004)). 10µL of siRNA was added to 2 x $10^7$ cells resuspended in 800µL media, and electroporated at 330V for 0.1 seconds in a 2mm cuvette. Following electroporation, cells were incubated at room temperature for 10-15 minutes. The cells were then transferred to warm RPMI media (supplemented with 10% FCS and 1% GlutaMax$^{TM}$) at a concentration of 1x$10^6$/mL and incubated at 37°C, 5% $CO_2$. Where required, a second knockdown was conducted 48 hours following the first electroporation.

## 2.4 Protein assays

### 2.4.1 Western blotting

Whole cell lysates were obtained from 1 x $10^6$ cells. Cells were washed using PBS and resuspended in 100µl 2X NuPAGE™ LDS Sample Buffer (Invitrogen™, NP0008). Samples were heated at 95°C for 10 minutes, cooled and sonicated using a Diagenode Bioruptor® Standard at high frequency, 30s on/30s off for 5 minutes.

Alternatively, salt-soluble protein extraction was performed on 1 × $10^6$ cells. Cells were washed in PBS and resuspended in 50µl BC300 (20 mM Tris-HCl, pH 7.5; 20% glycerol; 300 mM KCl; 5 mM EDTA) + 0.5% NP-40 (IGEPAL® CA-630, Sigma-Aldrich, I8896)+ protease inhibitor cocktail (PIC) (Sigma-Aldrich, P8340), then incubated on ice for 30 minutes with intermittent agitation. Samples were then spun down at 20000G for 10 minutes at 4°C and the supernatant retained. An equal volume of 4X NuPAGE™ LDS Sample Buffer (NuPAGE™ Invitrogen™, NP0008) plus 0.1% β-mercaptoethanol (Sigma-Aldrich, M6250) was then added to the supernatant. Samples were heated at 95°C for 10 minutes, then cooled. All protein samples were stored at -80°C prior to western blotting.

For Western blot analysis, protein lysate/extract from the equivalent of 1 × $10^5$ cells was loaded onto pre-cast 4–12% Bis-Tris gels (NuPAGE™ Invitrogen™, NP0322) and run in 1X MES SDS Running buffer (NuPAGE™ Invitrogen™, NP0002) at 180V for approximately 1 hour. Proteins were transferred to a PVDF membrane (Immobilon®-P Transfer Membrane, EMD Millipore, IPVH00010) using a wet transfer setup. Gel and membrane were sandwiched between filter paper and 2x sponges and submerged in a tank filled with a transfer buffer (192mM glycine, 25mM Trizma® base, 10% methanol). 100V was supplied to the tank for 1 hour at 4°C.

Following transfer, membranes were incubated in Ponceau stain (Sigma-Aldrich, P7170) for 5 minutes, then washed in 100% methanol and air dried. The membrane was then re-hydrated and blocked in 5% milk for 1 hour, then incubated with antibody overnight at 4°C. The following day, the membrane was washed in Tris-buffered saline (TBS) with 0.05% TWEEN® 20 (Sigma-Aldrich, P1379) (0.05% TBS-

TWEEN) (3 x 5 minutes) then incubated in 5% milk with the appropriate secondary antibody for 1 hour. Subsequently, the membrane was washed with 0.05% TBS-TWEEN (3 x 5 minutes) and then probed using enhanced chemiluminescence (ECL) (ECL™ Prime Western Blotting Detection Reagent, GE Healthcare, RPN2232) according to the manufacturer's instructions. The membrane was developed using chemiluminescence film (Hyperfilm™ ECL™, GE Healthcare, 28906837) and Xograph Imaging.

A complete list of antibodies and concentrations used can be found in Table 2.5

### 2.4.2 Intracellular FACS

Up to a maximum of 1 x $10^6$ cells were stained with live/dead dye for 30 minutes in the fridge, washed twice in FACS buffer, then fixed in 200µL 2% paraformaldehyde for 10 minutes on ice. Samples were then diluted with 1mL FACS buffer and centrifuged for 5 minutes at 1200G. The supernatant was discarded and cells were permeabilised in 100µL 0.1% Triton-X for 10 minutes on ice. Samples were once again diluted with 1mL FACS buffer and centrifuged for 5 minutes at 1200G. Fluorophore conjugated antibody or isotype control was then added to the cells at pre-determined concentrations and incubated in the dark at room temperature for 30 minutes. Samples were then washed twice with FACS buffer then analysed on either the BD LSRFortessa/BD LSRFortessa X20. Flow cytometric analysis was carried out using FACSDiva v8.1 (BD Biosciences) and FlowJo (v10.6.1) software. Data were analysed and graphs generated using Prism GraphPad for Mac (v8.0).

## 2.5 Chromatin immunoprecipitation (ChIP) assays

Prior to fixation, cells were counted, pelleted and washed with PBS. For single fixation, up to 1 × $10^7$ cells were fixed in 1mL 1% formaldehyde (FA) (10 minutes at room temperature, with rotation). For double fixation, up to 1 × $10^7$ cells were fixed in 1mL 2mM disuccinimidyl glutarate (DSG) (Sigma-Aldrich, 80424) (30 minutes at room temperature, with rotation), then pelleted (10000G, 1 minute)

and the supernatant discarded. The pellet was then resuspended in 1mL 1% FA (30 minutes at room temperature, with rotation). For both fixation conditions, following the last fixation step, pellets were rinsed in PBS before being frozen on dry ice and stored at -80°C prior to downstream analysis.

Fixed pellets were thawed on ice and resuspended in 100µL SDS lysis buffer (50mM Tris-HCl pH8.1, 10mM EDTA, 1% SDS) plus 1:1000 PIC (Sigma-Aldrich, P8340) and syringe-passaged through a 27-gauge needle. This was transferred to a Covaris tube (microTUBE AFA Fiber Pre-Slit Snap-Cap 6x16mm, Covaris Inc, 520045) and sonicated using a Covaris S220 (Woburn, MA, USA) using the following parameters: duty cycle = 10%, intensity = 5 Cycles, burst = 200. Single fixed samples were sonicated for 5 minutes and double fixed samples for 10 minutes. After sonication, samples were transferred to Eppendorf tubes and spun at 20000G for 10 minutes at 4°C. The supernatant was retained and diluted 1:10 in ChIP dilution buffer (CDB) (16.7mM Tris-HCl pH8.1, 167mM NaCl, 1.2mM EDTA, 0.01% SDS, 1.1% Triton$^{TM}$ X-100 (Sigma-Aldrich, X100-100mL)). A mixture of Protein A and Protein G Dynabeads$^{TM}$ (Invitrogen$^{TM}$, 10001/4D) was added to the chromatin for pre-clearing (10 minutes at 4°C, with rotation). Samples were then centrifuged at 1000G for 2 minutes and the chromatin was split into 1mL aliquots with 50µL being retained for input material and the remainder for sonication efficiency testing. Antibodies were added to the 1mL aliquots and incubated with rotation at 4°C overnight. A complete list of antibodies and concentrations used can be found in Table 2.5

Antibody:chromatin complexes were isolated using Protein A/G Dynabeads (4 hours at 4°C, with rotation). Beads were then washed three times with a solution of 50 mM HEPES-KOH, pH 7.6, 500 mM LiCl, 1 mM EDTA, 1% NP-40, and 0.7% Na deoxycholate. Following a Tris-EDTA wash, samples were eluted (65°C heat block, 30 minutes), treated with 1µl RNase A (Thermo Fisher Scientific, EN0531) (37°C heatblock, 30 minutes), and 1µl proteinase K (Thermo Fisher Scientific, EO0491) (45°C heatblock, 30 minutes), and DNA crosslinks reversed overnight (65°C heatblock). Samples were then purified using the Zymo ChIP DNA Clean

& Concentrator Kit (Zymo Research, D5205) according to the manufacturer's instructions.

**Table 2.5:** Antibodies used in Western blot and ChIP experiments

| Antibody name | Company/catalog number | Host | Working concentration |
|---|---|---|---|
| BCL2 | Cell Signaling/4223 | Rabbit | 1:10000 WB |
| BRD4 | Bethyl/A301-985A | Rabbit | 1:500 ChIP |
| CBP | Santa Cruz Biotechnology/sc-369 | Rabbit | 1:500 ChIP |
| GAPDH | Bethyl/A300-641A | Rabbit | 1:10000 WB |
| H3K27ac | Diagenode/C15410196 | Rabbit | 1:500 ChIP |
| H3K4me1 | Diagenode/pAb-194-050 | Rabbit | 1:500 ChIP |
| H3K4me3 | Diagenode/pAb-003-050 | Rabbit | 1:500 ChIP |
| H3K79me3 | Diagenode/pAb-068-050 | Rabbit | 1:500 ChIP |
| H3K9ac | Gift from Mahadevan lab | Mouse | 1:500 ChIP |
| MED1 | Bethyl/A300-793A | Rabbit | 1:500 ChIP |
| MYB | Abcam/ab109127 | Rabbit | 1:500 WB, 1:200 ChIP |
| P300 | Bethyl/A300-358A | Rabbit | 1:250 ChIP |
| RNAPII | Santa Cruz Biotechnology/sc-899X | Rabbit | 1:500 ChIP |
| TBP | Bethyl/A301-229A | Rabbit | 1:500 ChIP |
| TetR (FS2) | Klose lab | Rabbit | 1:1000000 WB, 1:500 ChIP |
| Anti-Rabbit HRP (2°) | Sigma-Aldrich/A6667 | Goat | 1:10000 WB |

### 2.5.1 ChIP qPCR

All ChIP quantitative polymerase chain reaction (qPCR) experiments were carried out on chromatin from 1 x $10^7$ cells unless otherwise indicated. DNA was quantified by qPCR, with ChIP samples normalised using input chromatin.

### 2.5.2 ChIP-sequencing

For ChIP-seq experiments on histone modifications, 4 x $10^7$ cells were used and for non-histone modifications 1 x $10^8$ cells were used.

DNA libraries were generated using the NEBNext® Ultra™ DNA library preparation kit for Illumina® (New England Biolabs, E7370). Libraries were sequenced paired-end on the NextSeq® 500 (Illumina®) using a NextSeq® High Output v2 75 cycle sequencing kit and data analysed as described in Chapter 2.9.

### 2.5.3 Reference normalised ChIP-sequencing (ChIP-rx)

ChIP-rx was performed exactly as described for ChIP-seq experiments apart from the addition of Drosophila melanogaster S2 cells (kindly provided by the Fulga

lab, WIMM) during the lysis step of the protocol. S2 cells were fixed as described above for ChIP qPCR and ChIP seq applications and added to the human samples prepared for ChIP seq in a 1:4 ratio during the lysis step. The normal ChIP-seq protocol was followed from this point onward.

## 2.6 RNA assays

### 2.6.1 Gene expression analysis

Total RNA was extracted using the QIAGEN RNeasy Mini Kit (QIAGEN, 74106) from $1 \times 10^6$ cells as per the manufacturer's instructions. RNA was quantified using the NanoDrop™ (Thermo Fisher Scientific) and equivalent amounts of each sample (to a maximum of 4µg) were used as template for reverse transcription.

To generate cDNA, 8µL RNA was first mixed with 1µL dNTP Mix (Invitrogen™, 8427-013) and 1µL random hexamers (Roche, S06405) and placed at 65°C for 5 minutes before being snap-cooled on ice for 2 minutes. The reverse transcription reaction (5X First strand buffer (Invitrogen™, 02321) 4µl, 0.1M DTT (Invitrogen™, 00147) 2µl, RNaseOUT (40 units/µl Invitrogen™, 51535) 1µl, Superscript III reverse transcriptase (200 units/µl Invitrogen™, 56575) 1µl, nuclease free water 2µl) was run using a Bio-Rad thermal cycler (95°C 3min, 30 cycles of 98°C 20s, 65°C 15s, 72°C 1min, then 72°C 2min ). cDNA thus generated was quantified by qRT-PCR.

### 2.6.2 Nascent RNA-sequencing

$1 \times 10^8$ THP-1 cells were treated with 500µM 4-Thiouridine (4-SU) for 1 h. Cells were lysed with TRIzol (Thermo Fisher Scientific, 15596026) and total RNA was extracted with chloroform and precipitated with ethanol. Following DNase treatment with TURBO DNA-free™ Kit (Thermo Fisher Scientific, AM1907) as per the manufacturer's instructions, 4-SU-incorporated RNA was biotinylated by labelling with 1mg/ml Biotin-HPDP for 90 minutes at room temperature. Following chloroform extraction, labelled RNA was separated using magnetic streptavidin beads. Beads were washed using a magnetic µMACS stand before RNA was eluted in two rounds of elution with 100µL 100mM DTT. RNA was purified using a QIAGEN

RNeasy MinElute Cleanup Kit (QIAGEN, 74204). Libraries were prepared using the NEBNext® Ultra™ Directional RNA Library Preparation Kit for Illumina® (New England Biolabs, E7420). Libraries were sequenced paired-end on the NextSeq® 500 (Illumina®) using a NextSeq® High Output v2 150 cycle sequencing kit and data analysed as described in Chapter 2.9.

### 2.6.3 Poly(A) minus RNA-sequencing

Confluent 10cm plates of mESCs were washed with PBS and directly lysed with 5ml QIAzol lysis reagent (QIAGEN, 79306), and total RNA was precipitated with ethanol. Following DNase treatment with TURBO DNA-free™ Kit (Thermo Fisher Scientific, AM1907) as per the manufacturer's instructions, RNA was precipitated with ethanol overnight at -80°C. 4µg of total RNA was depleted of ribosomal RNA (rRNA) using the RiboMinus™ Eukaryote Kit for RNA-Seq (Thermo Fisher Scientific, 1083708) according to the manufacturers instructions. rRNA depleted RNA was subject to poly(A) depletion using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, E7490). The resultant enriched poly(A) minus RNA was precipitated with ethanol overnight at -80°C. Libraries were prepared using the NEBNext® Ultra™ II Directional RNA Library Preparation Kit for Illumina® (New England Biolabs, E7645S). Libraries were sequenced paired-end on the NextSeq® 500 (Illumina®) using a NextSeq® High Output v2 75 cycle sequencing kit and data analysed as described in Chapter 2.9.

## 2.7 ATAC-seq

The ATAC seq protocol was adapted from Buenrostro *et al.* 2013. At least three biological replicates were prepared and analysed per condition. For each replicate, 50000 cells were harvested and washed in PBS and resuspended gently in 50µl cold lysis buffer (10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.1% IGEPAL CA-630). Cells were spun down immediately at 500G for 10min at 4°C. The supernatant was discarded and the pellet set on ice, then resuspended in a transposase reaction mix (25µL 2xTD buffer, 2.5µL Tn5 transposase, and 22.5µL nuclease free water)

(Illumina®, FC-121-1030) and incubated at 37°C for 30 min. The reaction was stopped using a QIAGEN MinElute PCR Purification Kit (QIAGEN, 28004) as per the manufacturer's instructions and samples were eluted in 10µL 10mM Tris pH 8. DNA fragments were then amplified and indexed (transposed DNA 10µL, customised universal primer 2.5µL, customised index primer 2.5µL, NEBNext® 2X Mastermix 25µL (New England Biolabs), nuclease free water 10µL) in a PCR reaction (72°C 5min, 98°C 30s, 12 cycles of 98°C 10s, 63°C 30s, 72°C 1min). The amplified library was purified using QIAGEN MinElute PCR Purification Kit and eluted in 20µL nuclease free water. Library quality was assessed using a DNA 1000 high sensitivity Screen Tape on the Agilent 2200 TapeStation system and quantified by qRT-PCR (KAPA library quantification kit, KK4828). Libraries were sequenced paired-end on the NextSeq® 500 (Illumina®) using a NextSeq® High Output v2 75 cycle sequencing kit and data analysed as described in Chapter 2.9.

## 2.8 Capture-C

A total of $2 \times 10^7$ THP-1 cells were assayed per sample according to the protocol by Davies *et al.* 2016. Cells were fixed with 2% formaldehyde for 10 minutes with rotation. The fixed cells were then incubated in lysis buffer for 20 minutes, then washed, resuspended in PBS and frozen at -80°C. The cell suspension was then thawed, centrifuged and PBS removed. The pellet was resuspended in DpnII buffer and digestion reactions were set up with the DpnII enzyme. Controls were also set up with undigested chromatin. Digests were incubated overnight at 37°C, shaking. Additional DpnII was added twice throughout the digest. Following this, a ligation reaction was carried out, adding DNA ligase and ligation buffer to the reaction with overnight incubation. The material was then de-crosslinked (65°C, overnight) and the DNA was extracted using a phenol-chloroform extraction method.

DpnII-generated 3C libraries (10-12µg per sample) were sonicated to a fragment size of 200 bp and Illumina paired-end sequencing adaptors (New England Biolabs, E6040, E7335, and E7500) were added using Herculase II (Agilent, 600677) for the final PCR. Indexing was performed in duplicate to maintain library complexity, with

libraries pooled after indexing. Capture-C probes targeting promoters were designed as 70 or 120bp biotinylated DNA oligonucleotides using the online CapSequm tool (http://apps.molbiol.ox.ac.uk/CaptureC/cgi-bin/CapSequm.cgi; Davies *et al.* 2016) Enrichment was performed with two successive rounds of hybridisation, streptavidin bead pulldown (Invitrogen™, M270), bead washes (NimbleGen SeqCap EZ, Roche, 05634261001), and PCR amplification (KAPA, KK4824; NimbleGen SeqCap EZ accessory kit v2, Roche, 07145594001). Libraries were sequenced paired-end on the NextSeq® 500 (Illumina®) using a NextSeq® High Output v2 300 cycle sequencing kit and data analysed as described in Chapter 2.9.

## 2.9 Bioinformatic analysis

### 2.9.1 Sequence analysis

For ChIP-seq, ChIP-rx, and ATAC-seq, quality control of FASTQ reads, alignment, PCR duplicate filtering, blacklisted region filtering, and UCSC data hub generation was performed using an in-house pipeline (https://github.com/Hughes-Genome-Group/NGseqBasic/releases). Briefly, the quality of the FASTQ files was checked with fastQC, and then mapped using Bowtie against the human genome assembly (hg19). Unmapped reads were trimmed with trim-galore and then mapped again. Short unmapped reads from this step were combined using Flash and then mapped again. PCR duplicates were removed using samtools rmdup, and any reads mapping to Duke blacklisted regions (UCSC) were removed using bedtools. Directories of sequence tags (reads) were generated from the sam files using the Homer tool makeTagDirectory. The makeBigWig.pl command was used to generate bigwig files for visualisation in UCSC, normalising tag counts to tags per 10 million. For ChIP-rx, the normalisation factor was adjusted to take into account the ratio of mapped human (hg19) and Drosophila (dm3) reads in the bound and input samples. Peaks were called using the Homer tool findPeaks, with the input track provided for background correction, using the -style histone or -style factor options to call peaks in histone modification or TF/ATAC datasets, respectively. Metagene profiles were generated using the Homer tool annotate- Peaks.pl. Statistical analysis of

differences between ATAC peaks was conducted with Diffbind, using EdgeR. Peaks were considered different between conditions if they had an adjusted p value (false discovery rate (FDR)) of <0.05. Motif analysis was conducted using the Homer tool findMotifsGenome.pl using the parameters -size given and -mask.

### 2.9.2 Locus Overlap Analysis

The R/Bioconductor package LOLA (Locus Overlap Analysis) was used for enrichment analysis of published ChIP-seq datasets that overlap with each genomic region set as the reference database (Sheffield *et al.* 2016, http://lola.computational-epigenetics.org). A single shared base pair was sufficient for regions to count as overlapping. Fisher's exact test with false discovery rate correction (Storey *et al.* 2020) was used to assess the significance of overlap in each pairwise comparison.

### 2.9.3 RNA-seq and gene expression analysis

Following quality control analysis with the fastQC package (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), reads were aligned using STAR (Dobin *et al.* 2013) against the human genome assembly (hg19). Reads that were identified as PCR duplicates using Samtools (H. Li *et al.* 2009) were discarded. Gene expression levels were quantified as read counts using the featureCounts function from the Subread package (Liao *et al.* 2014) with default parameters. The read counts were used for the identification of global differential gene expression between specified populations using the edgeR package (Robinson *et al.* 2010). RPKM (reads per kilobase of transcript, per million) values were also generated using the edgeR package. Genes were considered differentially expressed between populations if they had an adjusted p value (FDR) of <0.05.

### 2.9.4 Capture-C analysis

For Capture-C, data analysis was performed using an in-house pipeline (https://github.com/Hughes-Genome-Group/CCseqBasicF/releases). Capture-C interactions between captured promoters and enhancers were quantified for statistical

analysis. Peaks outside of the bounds of Capture-C interaction domains (visually determined using UCSC genome browser) and those on trans chromosomes were removed from the analysis. Peaks within 10 kb of the Capture-C probe hybridisation site were also removed. Holm–Bonferroni-adjusted p values for each peak were calculated by comparing all of the normalised read counts for each DpnII fragment and all replicates within a peak using a paired Mann–Whitney test for the two treatment conditions (control vs. siRNA).

# 3

# Myb repression during leukaemogenesis results in an attenuated leukaemia phenotype

## Contents


**3.1 Introduction . . . . . 44**
3.1.1 MicroRNA mediated regulation of gene expression . . . 44
3.1.2 Engineering an endogenous miRNA mediated *Myb* repression circuit . . . . . 46
**3.2 Aims . . . . . 48**
**3.3 Results . . . . . 48**
3.3.1 Generation of an endogenous *Myb* repression circuit in mice 48
3.3.2 $Myb^{MRE/MRE}$ mice exhibit no major deficits in steady state haematopoiesis . . . . . 51
3.3.3 $Myb^{MRE/MRE}$ mice have preserved bone marrow reconstitution ability . . . . . 56
3.3.4 *In vitro* assay optimisation . . . . . 59
3.3.5 $Myb^{MRE/MRE}$ cells show a decreased trend in transformation potential *in vitro* . . . . . 61
3.3.6 *In vivo* experimental approach to test the requirement for Myb in leukaemic transformation . . . . . 65
3.3.7 *MLL-AF9* transduced $Myb^{MRE/MRE}$ and WT cells are immunophenotypically similar prior to transplantation . 66
3.3.8 MLL-AF9 transduced $Myb^{MRE/MRE}$ cells give rise to an attenuated leukaemia phenotype *in vivo* . . . . . 68
**3.4 Discussion . . . . . 71**


*In this section, the* $Myb^{MRE/MRE}$ *mouse line was generated in collaboration with the Fulga (MRC Weatherall Institute of Molecular Medicine, Oxford) and Davies labs (Wellcome Trust Centre for Human Genetics, Oxford). sgRNA and ssODN design and optimisation was performed by Yale Michaels (Fulga Lab), and*

*CRISPR-Cas9 editing experimental work was performed by the Davies lab. FACS staining panels were designed by Alex Thomas and Simona Valletta (Nerlov lab, MRC Weatherall Institute of Molecular Medicine, Oxford), both of whom assisted with initial immunophenotypic analysis.*

## 3.1 Introduction

Leukaemogenesis is a process which ultimately leads to the transformation of a normal haematopoietic cell into a leukaemic stem cell capable of propagating disease. In the context of *MLL*r leukaemia, this process is initiated by powerful MLL-FP oncoproteins, which drive leukaemia progression by increasing the expression of their target genes via both transcriptional elongation and epigenetic mechanisms, thus disrupting normal transcriptional networks.

As mentioned in Chapter 1.3.3, MYB is a direct target of MLL-FPs and has been implicated as an important gene in mediating *MLL*r leukaemogenesis (Hess *et al.* 2006). However, the requirement for MYB early on in the process of leukaemia initiation and transformation is still unclear, in part because systems to ensure that robust MYB repression only occurs at the onset of leukaemia initiation have not been available. To address this, we exploited the inherent properties of endogenous microRNAs (miRNAs) to modulate gene expression (see Chapter 3.1.1, by designing a conditional miRNA mediated *Myb* repression circuit, which is active only in the presence of the MLL-AF9 oncoprotein (see Chapter 3.1.2.

### 3.1.1 MicroRNA mediated regulation of gene expression

miRNAs are part of the RNA inhibition (RNAi) pathway. They are endogenous, small, non-coding RNAs, which play a critical role in the post-transcriptional regulation of gene expression (Bartel 2004). They function by base-pairing with short, well-defined sequences called miRNA response elements (MREs), which have complementarity with the miRNA and are typically located within the 3' untranslated regions (UTRs) of transcribed mRNA (Bartel 2004; Brennecke *et al.* 2005; Bartel 2009).

miRNAs have been shown to be capable of repressing gene expression in all mammalian cell types tested and many miRNAs show evolutionary conservation across metazoan species (Landgraf *et al.* 2007; Friedman *et al.* 2009). Over half of the mammalian transcriptome contains predicted or validated MREs (Landgraf *et al.* 2007). Their functional importance in regulating gene expression is most apparent when studied in a disease context, where disruption of miRNA mediated gene regulation is observed. Multiple disease phenotypes are associated with global dysregulation of miRNA expression or perturbations affecting specific miRNAs or MREs (Bartel 2018).

miRNA biogenesis is a multistep process involving multiple protein complexes both within the nucleus (Lee *et al.* 2002; Lee *et al.* 2003) and later in the cytosol (Yi *et al.* 2003; Zhang *et al.* 2004), finally resulting in a single stranded miRNA (known as the guide strand) being loaded onto an Argonaut (AGO) protein (Bartel 2018) which is part of a multi-component complex known as the RNA-induced silencing complex (RISC).

Based on its interaction with AGO, the miRNA guide has been characterised with several functional domains - the anchor, seed, central, 3' supplementary and tail regions (Figure 3.1A). Of these, the seed sequence is the main determinant of binding specificity between a miRNA and its MRE, and the 3' supplementary region further facilitates this interaction (Wee *et al.* 2012; Chandradoss *et al.* 2015).

Following miRNA-mediated RISC (miRISC) binding to an MRE, target repression may occur via multiple mechanisms which can be broadly separated into two categories: 1) processes that lead to transcript degradation through direct RNA cleavage (also known as slicing) by AGO2 (Hutvagner *et al.* 2002), or destabilisation via de-adenylation (Rehwinkel *et al.* 2005; Behm-Ansmant *et al.* 2006b; Bartel 2018, and 2) processes that lead to translational repression (Behm-Ansmant *et al.* 2006a; Chu *et al.* 2006; Meijer *et al.* 2013) (Figure 3.1B).

Overall, the strength of miRNA-mediated repression is thought to be positively correlated with the following parameters: 1) increased miRNA:MRE complementarity (Brennecke *et al.* 2005; Wee *et al.* 2012; Chandradoss *et al.* 2015;

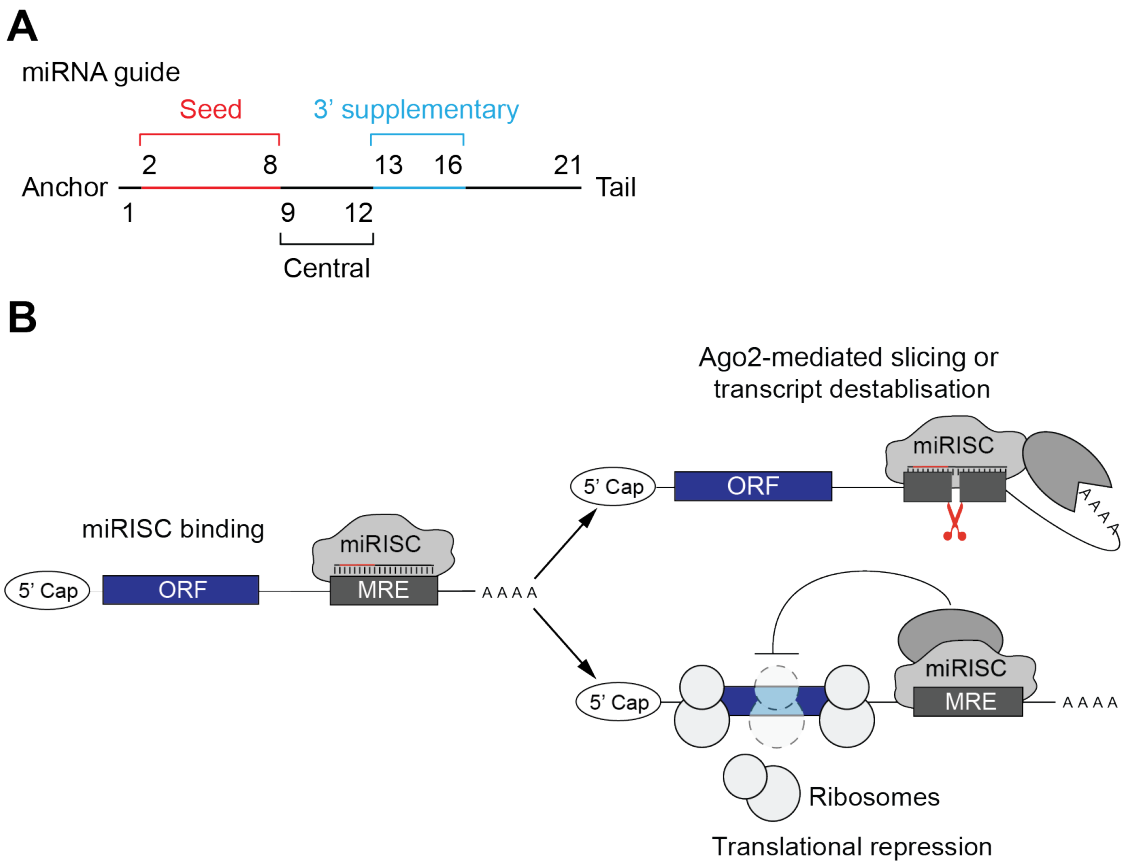



**Figure 3.1: MicroRNA mediated gene regulation.** (A) Functional domains of a miRNA guide. Numbers indicate ribonucleotide position from the 5' end. (B) Schematic representation of the two major pathways underlying miRNA-mediated repression. (Adapted from Michaels *et al.* 2019 with permission.)

Vainberg Slutskin *et al.* 2018; Michaels *et al.* 2019), 2) increased number of MREs per 3'UTR, and 3) increased stoichiometric ratio between miRNA molecules and its target MRE pool (Arvey *et al.* 2010; Bosson *et al.* 2014).

Therefore, it follows that effective repression of a particular mRNA species may be achieved by manipulating the above parameters.

### 3.1.2 Engineering an endogenous miRNA mediated *Myb* repression circuit

As MYB is upregulated in MLL-AF9 leukaemia, to investigate if high levels of MYB is required in early leukaemogenesis, it was necessary to ensure that MYB repression is temporally coupled with the onset of leukaemia initiation. Therefore, using MLL-AF9 as a leukaemia model, where the fusion protein drives leukaemogenesis,

MYB repression should occur only at the onset of MLL-AF9 expression. To this end, in collaboration with the Fulga lab (MRC Weatherall Institute of Molecular Medicine, Oxford), we designed a miRNA mediated *Myb* repression circuit in a transgenic mouse model, where repression of *Myb* mRNA occurs only in the presence of MLL-AF9 (Figure 3.2).

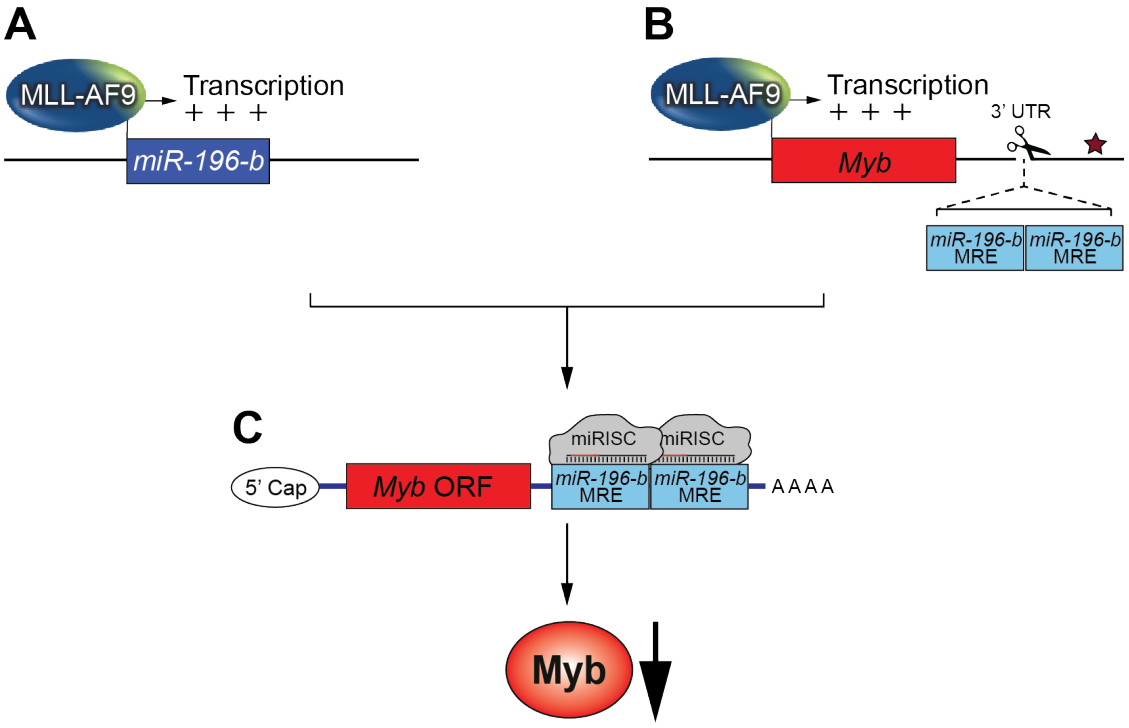



**Figure 3.2: Schematic of engineering an endogenous *Myb* repression circuit.** MLL-AF9 results in upregulation of (A) miRNA species e.g. *miR-196-b* and (B) *Myb*. (B) CRISPR-Cas9 insertion of 2 x MREs complementary to *miR-196-b* into the 3' UTR sequence of the *Myb* gene. The star at the 3' end represents the polyadenylation site. (C) *miR-196-b* miRISC mediated repression of Myb expression.

Our model exploits the fact that in addition to upregulating the expression of protein coding genes, MLL-AF9 also causes widespread dysregulation of miRNA expression, leading to upregulation of particular miRNA species e.g. *miR-196-b* (Fleischmann *et al.* 2016; S. E. Meyer *et al.* 2018) (Figure 3.2A). Therefore, the introduction of 2 sequential MREs with perfectly complementarity to *miR-196-b* into the 3' UTR of *Myb* (Figure 3.2B), should effectively create a conditional repression circuit for *Myb*, which is active only when MLL-AF9, and hence miR-196-b, is expressed (Figure 3.2C). Theoretically, this avoids the temporal ambiguity

of other methods (as described in Chapter 1.3.3) to suppress Myb expression during leukaemia initiation.

A potential flaw in the circuit design, is that *miR-196-b* is transcribed from the *HoxA* cluster of genes (Debernardi *et al.* 2007; Jongen-Lavrencic *et al.* 2008; Schotte *et al.* 2010), and in haematopoietic cells, the expression of *HoxA* genes appears to largely mirror the pattern observed for *Myb* expression. *HoxA* genes are preferentially expressed at high levels in immature stem/progenitor compartments and their expression is downregulated during differentiation and maturation (Lawrence *et al.* 1996; Sauvageau *et al.* 1994; Moretti *et al.* 1994; Thorsteinsdottir *et al.* 2002; Pineault *et al.* 2002). There is therefore a possibility that *miR-196-b* may be expressed to a sufficient level in stem/progenitor cells as to cause a sufficient knockdown of Myb expression that may impair definitive haematopoiesis.

## 3.2 Aims

1. To establish an experimental system that will effect *Myb* repression during leukaemia transformation.

2. To determine the effect of reduced *Myb* expression on leukaemia transformation *in vitro*.

3. To determine the effect of reduced *Myb* expression on leukaemia transformation in a mouse transplantation model.

## 3.3 Results

### 3.3.1 Generation of an endogenous *Myb* repression circuit in mice

Prior to engineering the *Myb* repression circuit in mice, it was important to confirm the upregulation of each circuit component by MLL-AF9. Therefore, lineage depleted (Lin-; using antibodies against CD5, CD45R (B220), CD11b, Gr-1 (Ly-6G/C), 7-4, and Ter-119) mouse bone marrow (BM) cells were retrovirally transduced with human *MLL-AF9* constitutively expressed from the Murine Stem

Cell Virus (MSCV) promoter and an mCherry reporter constitutively expressed from the 3-phosphoglycerate kinase (PGK) promoter (Figures 3.3A & 7.1 ). In keeping with in-house and published data (Zuber *et al.* 2011; Bousquet *et al.* 2013; Pattabiraman *et al.* 2014), RNA qRT-PCR analysis two weeks post-transduction confirmed increased *Myb* mRNA expression (Figure 3.3B) and Nanostring (data from Fulga lab) demonstrated a significant upregulation of *miR-196-b* expression (Figure 3.3C) in the presence of MLL-AF9.

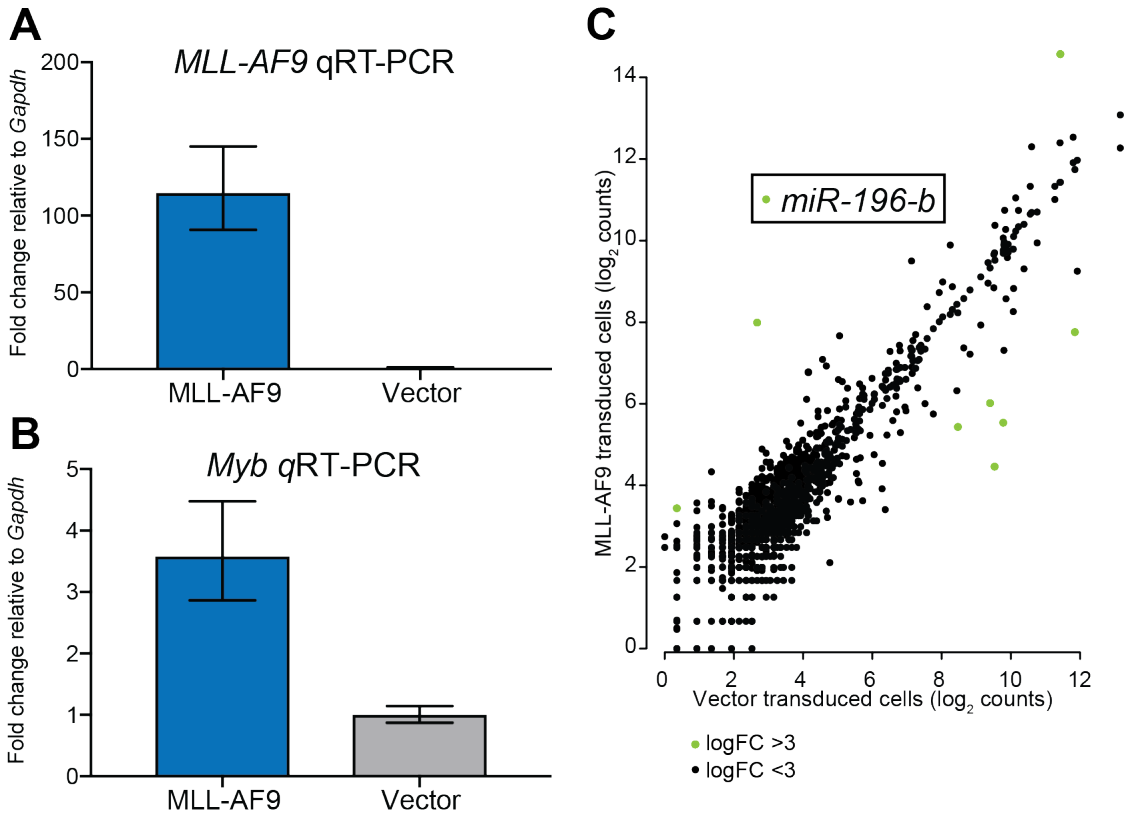



**Figure 3.3: *MLL-AF9* expression in mouse Lin- bone marrow cells results in upregulation of *Myb* and *miR-196-b*.** (A) qRT-PCR analysis of *MLL-AF9* and (B) *Myb* expression two weeks following transduction with *MLL-AF9*, normalised as fold change relative to the housekeeping gene *Gapdh*. Error bars represent the standard deviation from three biological replicates. (C) Nanostring $\log_2$ counts of cells transduced with virus containing either *MLL-AF9* (y-axis) or vector control (x-axis). Data points represent single miRNAs. miRNAs with logFC>3 are highlighted in green (data from Fulga lab). Data analysed in R from 3 biological replicates.

In view of the above results confirming *miR-196-b* upregulation in the presence of MLL-AF9, a gene editing strategy to insert two *miR-196-b* MREs with perfect

complementarity into the *Myb* 3' UTR was designed by Yale Michaels (Fulga lab) and experimental work undertaken by the Davies lab.

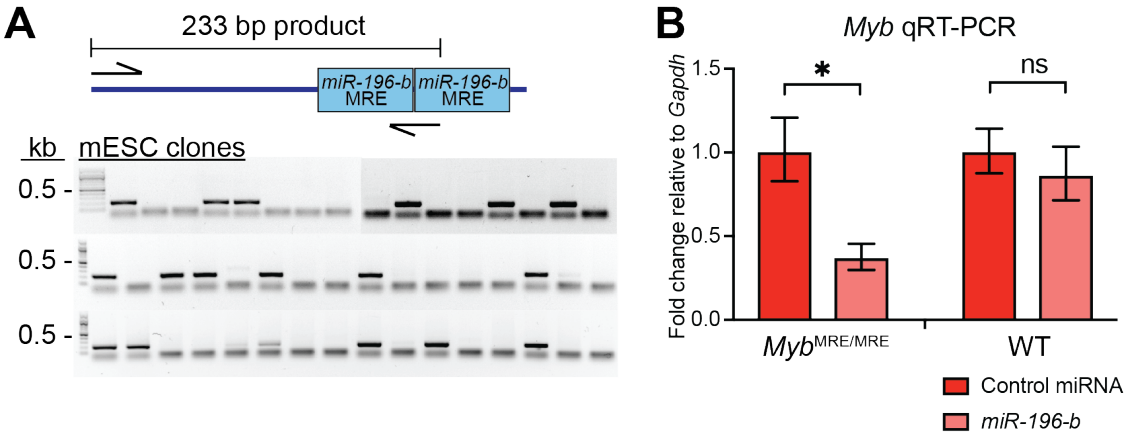



**Figure 3.4: Engineered *miR-196-b* MREs at the *Myb* 3' UTR are functional in mESCs.** (A) PCR screening of mESC clones for insertion of MREs. PCR product run on 1% agarose gel with a 2-log DNA ladder to facilitate identification of PCR product (233 bp) specific for positive clones. Wild type Myb does not produce a PCR product. (B) qRT-PCR analysis of *Myb* expression in engineered mESCs with homozygous insertion of the *miR-196-b* MREs into *Myb* ($Myb^{MRE/MRE}$) compared to WT cells, 24 hours following transfection with *miR-196-b* or control non-targeting mimic. Error bars represent the standard deviation from three biological replicates. Significance tested by unpaired t-test, ns = not significant, * = $p < 0.05$.

As an initial experiment, CRISPR-Cas9-mediated editing of *Myb* was performed in mESCs, resulting in surprisingly high targeting efficiency, with >40% of the clones screened being found to be positive for the insertion (Figure 3.4A). Next, functional validation of the MREs by transfecting ($Myb^{MRE/MRE}$) with *miR-196-b* mimic, resulted in robust *Myb* knockdown compared with a non-targeting miRNA control (Figure 3.4B). This suggests that *miR-196-b* was able to repress *Myb* in this system.

As editing frequency was relatively high in mESCs, we proceeded with C57BL/6J zygote pronuclear injections. Microinjected zygotes were cultured overnight then implanted in foster mothers. On screening the 14 pups obtained, 1 male pup was found to be homozygous, and 2 female pups were found to be heterozygous for the transgene. These mice were crossed to generate a $Myb^{MRE/MRE}$ transgenic mouse line.

### 3.3.2 *Myb*$^{MRE/MRE}$ mice exhibit no major deficits in steady state haematopoiesis

Given the known expression profile of *miR-196-b* (see Chapter 3.1.2), there was risk that in *Myb*$^{MRE/MRE}$ mice, *miR-196-b* may be expressed to a sufficient level in stem/progenitor cells as to effect a sufficient reduction in Myb expression to impair definitive haematopoiesis. Significantly, *Myb*$^{MRE/MRE}$ mice appear phenotypically normal and are viable and fertile, producing normal litter sizes compared to wild type (WT) mice. Furthermore, there appeared to be no gross deficit in their haematopoiesis, suggesting absent or minimal circuit activity at baseline. Nonetheless, for the purposes of downstream experimental design, it was important to ensure that the bone marrow composition and function of *Myb*$^{MRE/MRE}$ mice was comparable to that of WT mice.

To this end, detailed immunophenotypic analyses of the bone marrow and peripheral blood from 8 week old *Myb*$^{MRE/MRE}$ mice were performed and compared with *Myb*$^{+/+}$ wild type litter mates (WTLM) and age-matched WT controls.

Using the gating strategy as described in Cabezas-Wallscheid *et al.* 2014 (Figure 3.5A), immunophenotypic analysis of HSC and multipotent progenitor (MPP) populations was performed. Of interest, this showed a statistically significant expansion of the MPP1 and MPP2 compartments compared to controls (Figure 3.5C and D). While the HSC compartment demonstrated a slight upward trend in the *Myb*$^{MRE/MRE}$ mice, this was not statistically significant (Figure 3.5B) and MPP3 and MPP4 compartments were also not significantly different (Figure 3.5E and F).

Analysis of myeloerythroid progenitors according to Pronk *et al.* 2007 (Figure 3.6A) conducted at the same time, revealed no statistically significant differences in the size of the megakaryocyte progenitor (MkP), granulocyte monocyte progenitor (GMP), pre–granulocyte-macrophage progenitor (Pre GM), pre-megakaryocyte erythroid (Pre MegE), pre-erythroid colony forming unit (Pre CFU-E) or erythroid colony forming unit (CFU-E) & proerythroblast (Pro Ery) populations in *Myb*$^{MRE/MRE}$ mice compared to control mice (Figure 3.6B-G).

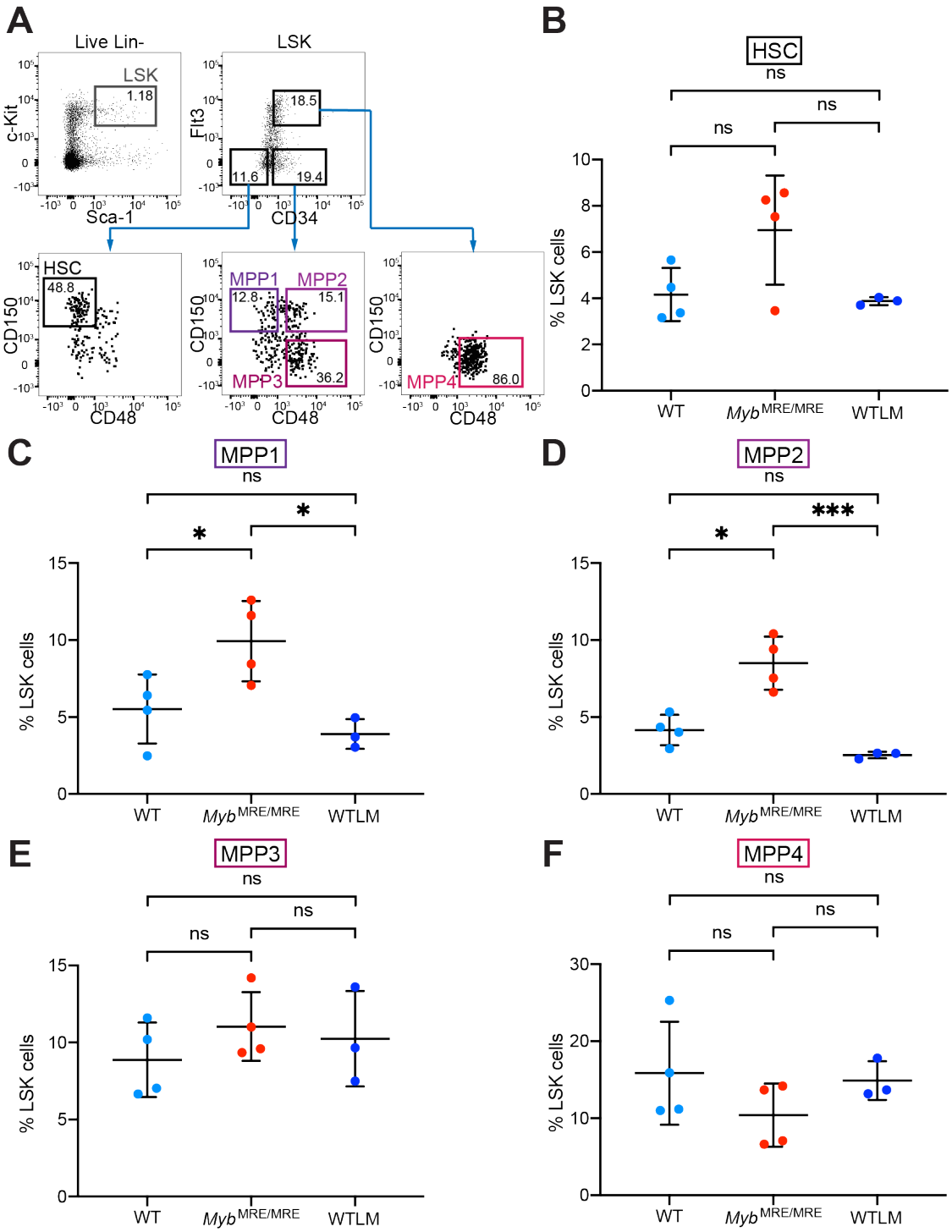



**Figure 3.5: Analysis of HSC and multipotent progenitor populations.** (A) Gating strategy used to define HSC and multipotent progenitor (MPP) populations as per Cabezas-Wallscheid *et al.* 2014. Endogenous HSC (B), MPP1 (C), MPP2 (D), MPP3 (E) and MPP4 (F) frequency as a % of Lin- Sca-1+ c-Kit+ (LSK) cells in WT mice (light blue), $Myb^{MRE/MRE}$ mice (red) and WTLM (royal blue). Bars show mean ± SD, n = 4 from two experiments. Significance tested by one-way ANOVA, ns = not significant, * = p <0.05, ** = p <0.01.

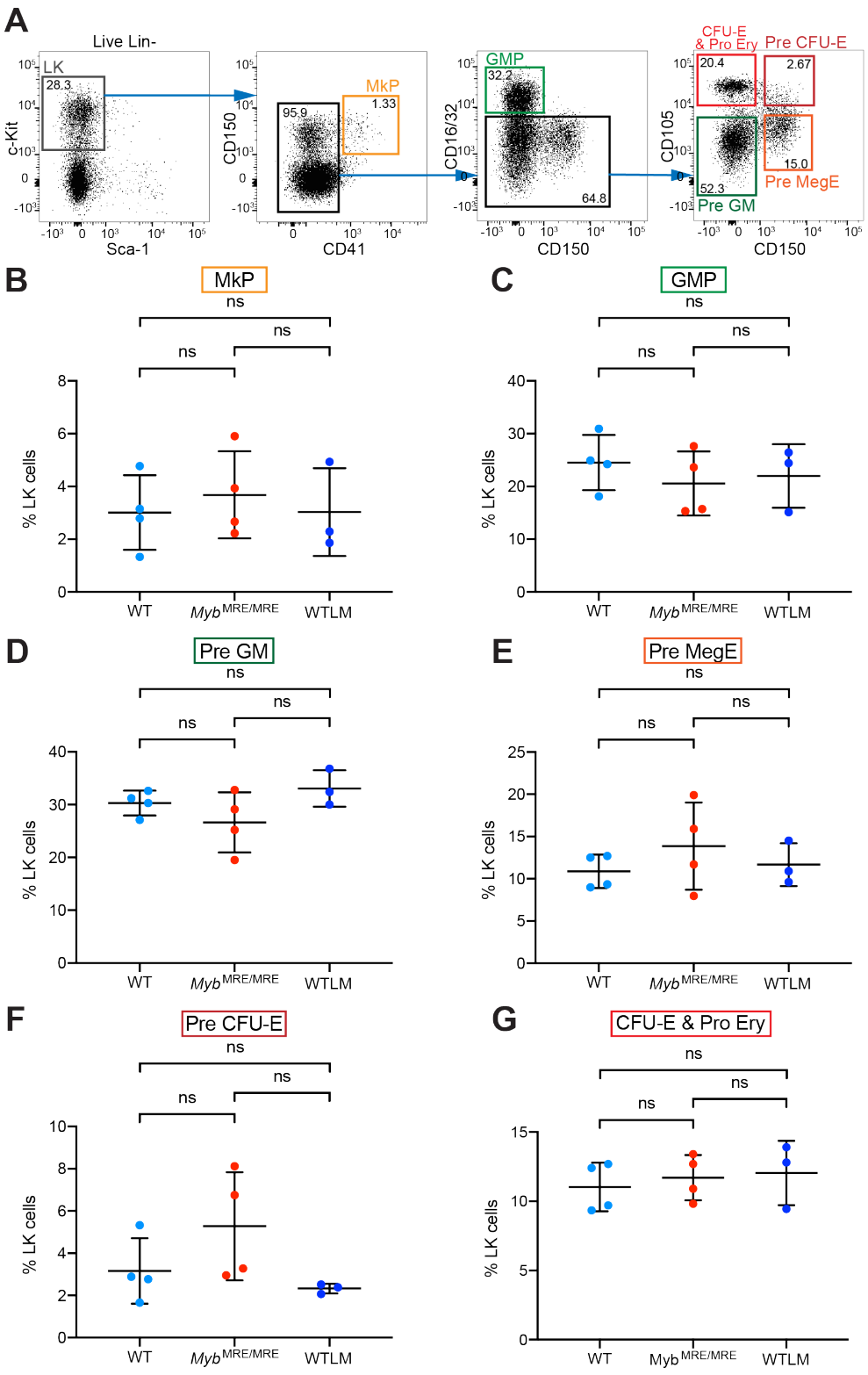



**Figure 3.6: Analysis of myeloerythroid progenitor populations.** (A) Gating strategy used to define myeloerythroid progenitor populations as per Pronk *et al.* 2007. Endogenous MkP (B), GMP (C), Pre GM (D), Pre MegE (E), Pre CFU-E (F) and CFU-E (G) & Pro Ery frequency as a % of Lin- c-Kit+ (LK) cells in WT mice (light blue), $Myb^{MRE/MRE}$ mice (red) and WTLM (royal blue). Bars show mean ± SD, n = 4 from two experiments. Significance tested by one-way ANOVA, ns = not significant.

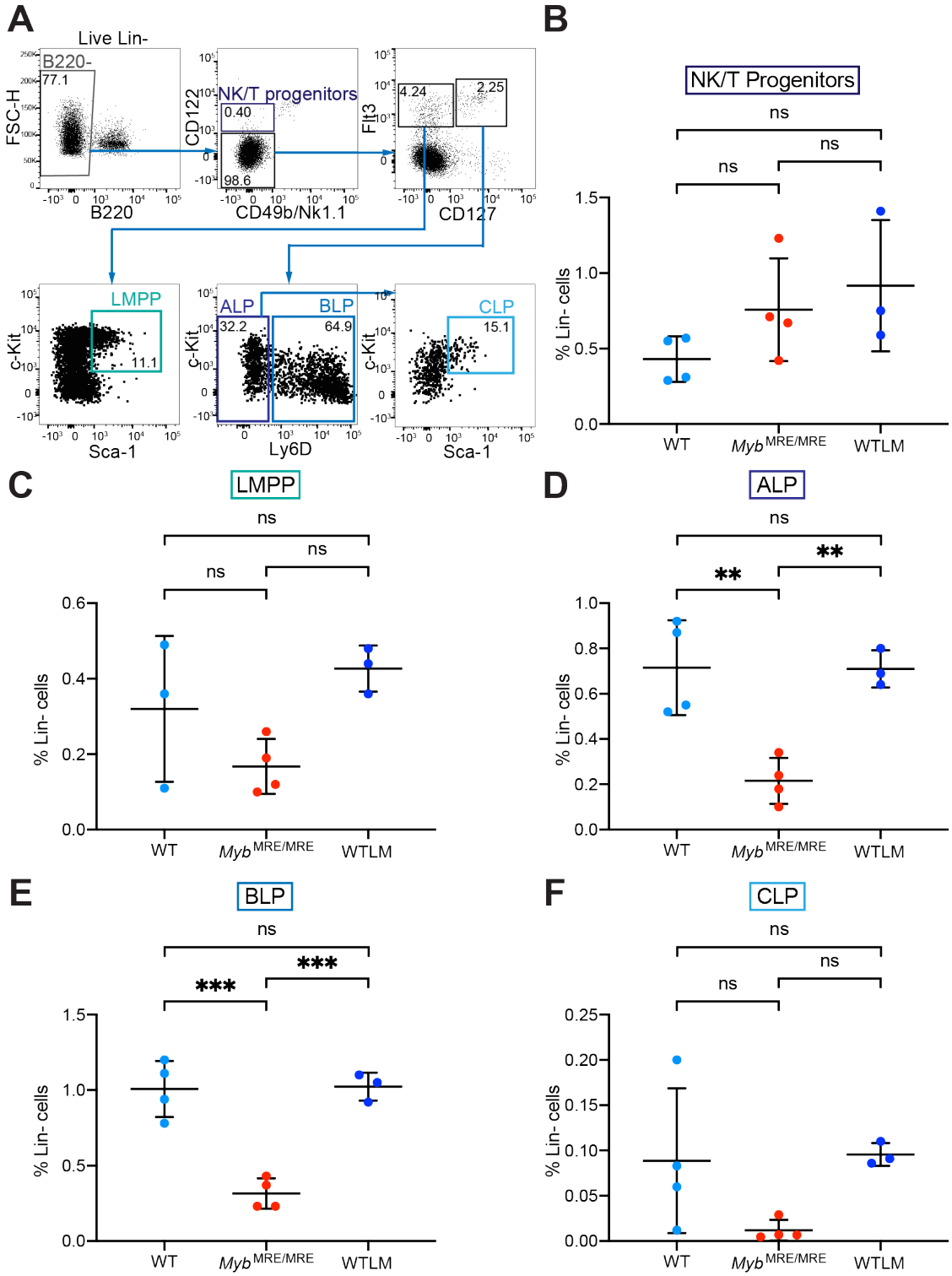



**Figure 3.7: Analysis of lymphoid progenitor populations.** (A) Gating strategy used to define lymphoid progenitor populations as per Buono *et al.* 2016. Endogenous NK/T progenitors (B), LMPP (C), ALP (D), BLP (E) and CLP (F) frequency as a % of Lin- cells in WT mice (light blue), $Myb^{MRE/MRE}$ mice (red) and WTLM (royal blue). Bars show mean ± SD, n = 4 from two experiments. Significance tested by one-way ANOVA, ns = not significant, **p = <0.01, ***p = <0.001.

In contrast, immunophenotypic analysis of lymphoid progenitor populations as defined by Buono *et al.* 2016 (Figure 3.7A), demonstrated a statistically significant reduction in all lymphoid progenitor (ALP) and B-lymphoid progenitor (BLP) (Figure 3.7D & E) compartments of *Myb*$^{MRE/MRE}$ compared to controls. Differences in natural killer T cell (NK/T) progenitor (Figure 3.7B), lymphoid-primed multipotent progenitor (LMPP) (Figure 3.7C) and common lymphoid progenitor (CLP) (Figure 3.7F) populations did not reach statistical significance.

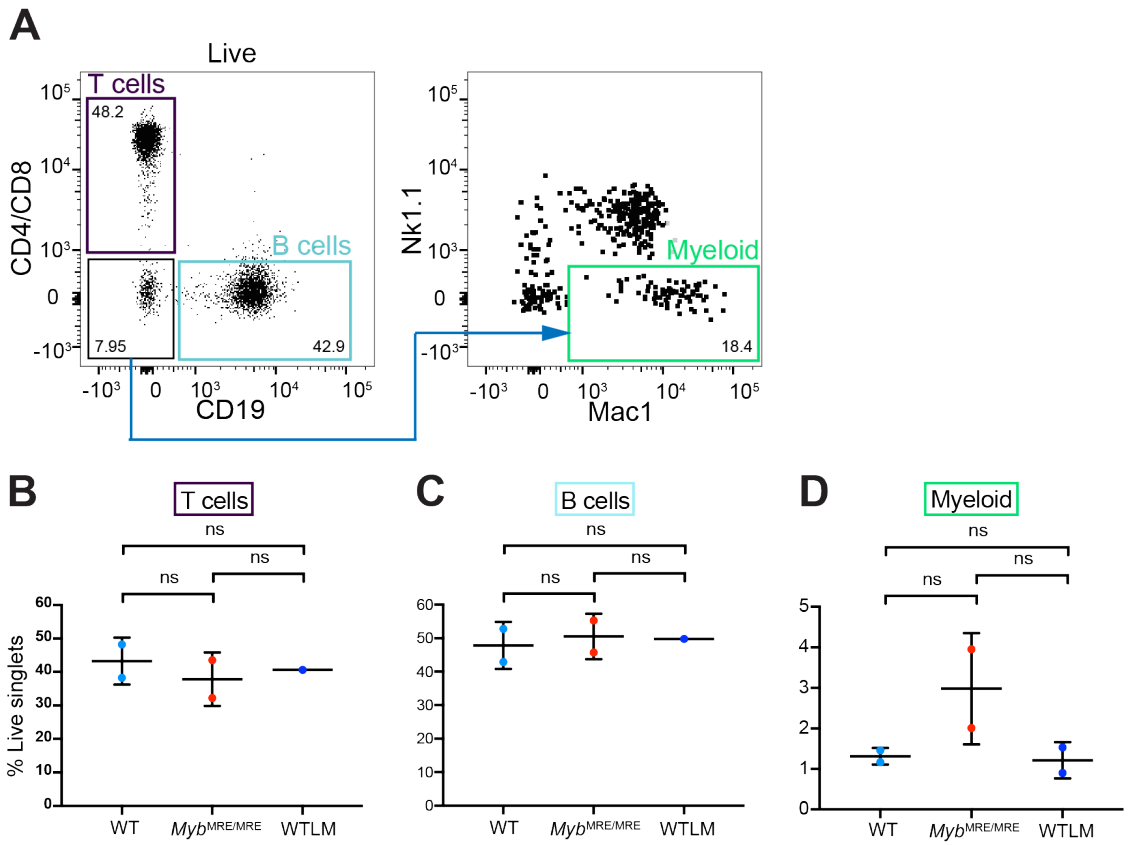



**Figure 3.8: Analysis of peripheral blood lineages.** (A) Gating strategy used to define peripheral blood lineage populations. Endogenous T cells (CD4+ and/or CD8+) (B), B cells (CD19+) (C) and Myeloid (Mac1+, Nk1.1-) (D) frequency as a % of live singlets in WT mice (light blue), *Myb*$^{MRE/MRE}$ mice (red) and WTLM (royal blue). Bars represent mean ± SD, n = 2 from a single experiment. Significance tested by one-way ANOVA, ns = not significant.

Interestingly, the differences observed in MPP1, MPP2, ALP and BLP compartments in *Myb*$^{MRE/MRE}$ mice did not result in significant abnormalities in peripheral blood cell populations (Figure 3.8). Importantly, *Myb*$^{MRE/MRE}$ mice had normal T

and B cell numbers (Figure 3.8B & C). While there was a slight upward trend in the peripheral myeloid population (Figure 3.8D), this was not statistically significant.

These data suggest that there is a possible trend towards myeloid bias in $Myb^{\mathrm{MRE/MRE}}$ mice. As discussed in Chapter 1.3.2, it is well documented that lymphopoiesis is particularly vulnerable to loss of Myb, and that lower levels of Myb expression result in a myeloid differentiation bias. Therefore, at face value, the data observed in $Myb^{\mathrm{MRE/MRE}}$ mice could possibly be explained by low level Myb KD due to activation of the circuit in some progenitor populations. However, as the sample number was relatively small (n = 4 per genotype analysed over two experiments), these observations could be attributed to biological variation and certainly more biological replicates are needed make any confident conclusions. Overall however, the data suggest that there were no major deficits in steady state haematopoiesis observed in $Myb^{\mathrm{MRE/MRE}}$ mice.

As the bone marrow populations of WTLMs were not statistically different compared to WT controls, the transgenic mice were maintained as a homozygous line in keeping with the 3R principles, and all subsequent experiments were conducted using WT mice as controls.

### 3.3.3 $Myb^{\mathrm{MRE/MRE}}$ mice have preserved bone marrow reconstitution ability

Next, to investigate if $Myb^{\mathrm{MRE/MRE}}$ mice have normal BM function, a competitive repopulation assay was performed (Figure 3.9). BM from each donor mouse was transplanted into three recipient mice, and three donors were used for each genotype. Peripheral blood analysis was performed every four weeks following transplantation to monitor engraftment (Figure 3.10) and end-point BM analysis was performed at 18 weeks post-transplantation (Figure 3.11).

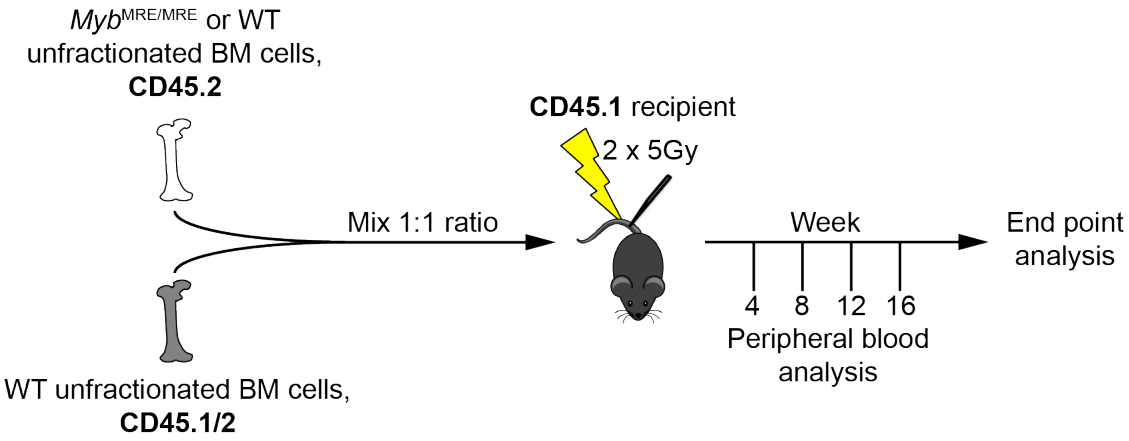



**Figure 3.9: Experimental design for competitive repopulation assay.** Unfractionated BM cells from 8 week old CD45.2 $Myb^{MRE/MRE}$ or WT mice were mixed in a 1:1 ratio with unfractionated BM cells from CD45.1/2 WT mice and transplanted via tail vein injection into age-matched lethally irradiated (2 x 5Gy) CD45.1 recipient mice. Engraftment was monitored by regular peripheral blood analysis and end-point BM analysis was performed at 18 weeks post-transplantation.

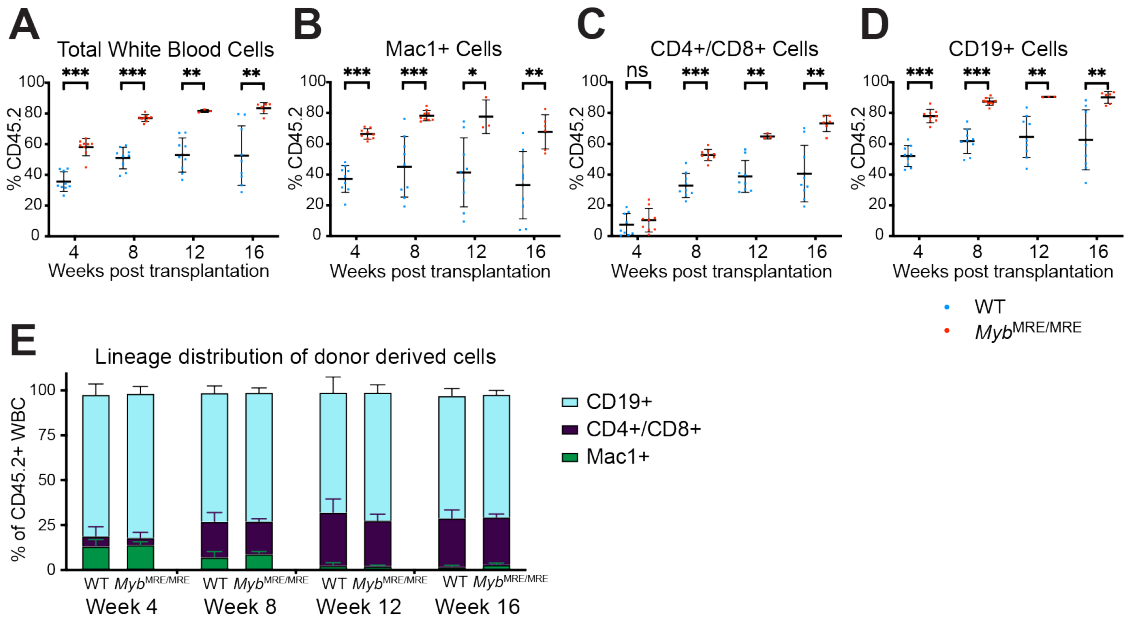



**Figure 3.10: $Myb^{MRE/MRE}$ mice have increased PB reconstitution ability.** Immunophenotypic analysis of CD45.2 peripheral blood chimerism of total white blood cells (A), myeloid (Mac1+) (B), T (CD4+ and/or CD8+) (C) and B (CD19+) (D) cells in CD45.1 recipient mice, following transplantation of WT (light blue) or $Myb^{MRE/MRE}$ (red) unfractionated bone marrow with CD45.1/2 competitor BM cells. BM from each donor mouse was transplanted into three recipient mice. (E) Stacked bar graph of % contribution of each lineage to the total peripheral blood compartment. Myeloid represented in green, T cells in purple and B cells in light blue. Bars represent mean ± SD, n = 9 from a single experiment. Significance tested by unpaired t-test, ns = not significant, ** = p <0.01, *** = p <0.001.

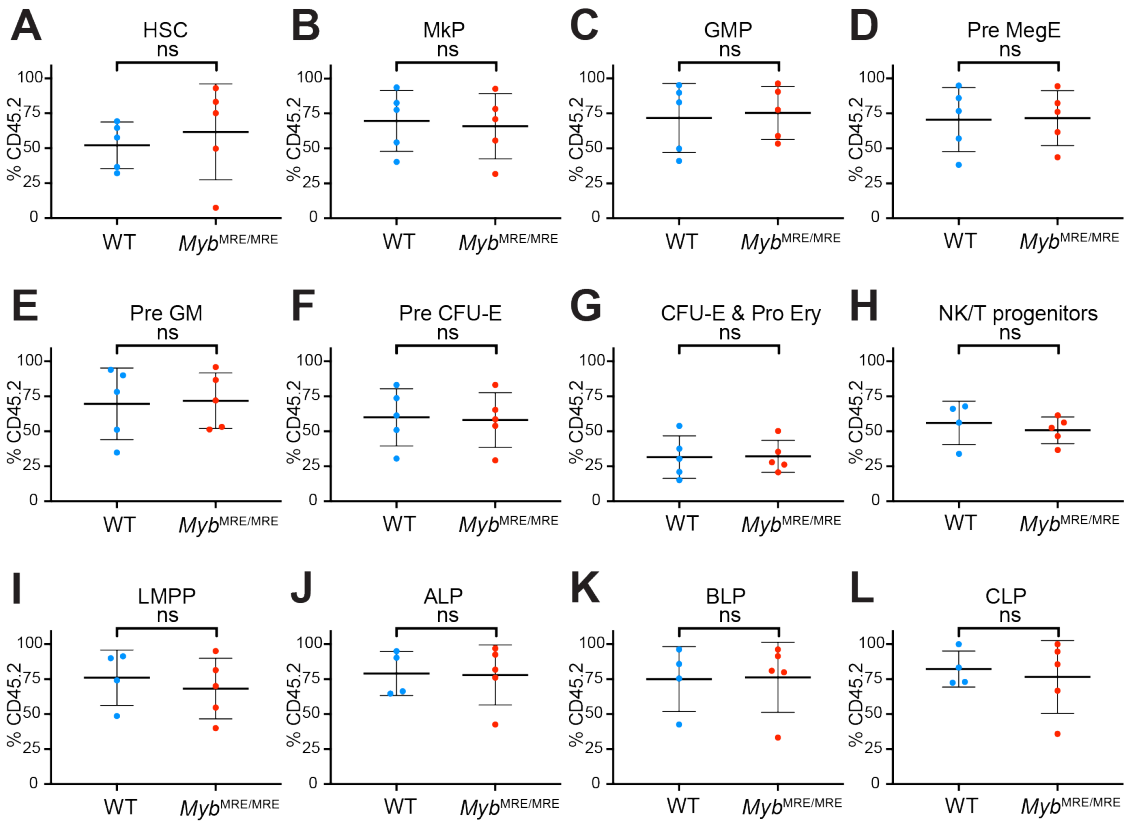



**Figure 3.11: *Myb*$^{MRE/MRE}$ mice have preserved BM reconstitution ability.** Immunophenotypic analysis of CD45.2 chimerism of (A) HSC, (B) MkP, (C) GMP, (D) Pre MegE, (E) Pre GM, (F) Pre CFU-E, (G) CFU-E & Pro Ery, (H) NK/T, (I) LMPP, (J) ALP, (K) BLP and (L) CLP populations in the bone marrow compartment of CD45.1 recipient mice, following transplantation of WT (light blue) or *Myb*$^{MRE/MRE}$ (red) unfractionated bone marrow with CD45.1/2 competitor BM cells. BM from each donor mouse was transplanted into three recipient mice. Error bars show mean ± SD, n = 4-5 from a single experiment. Significance tested by unpaired t-test, ns = not significant.

Interestingly, while no significant differences in the PB populations of *Myb*$^{MRE/MRE}$ mice were observed during steady state haematopoiesis (Figure 3.8B-D), in a competitive transplantation setting, *Myb*$^{MRE/MRE}$ bone marrow demonstrated increased peripheral blood reconstitution ability compared to WT controls (Figure 3.10). The total white blood cell (WBC) contribution of *Myb*$^{MRE/MRE}$ BM was significantly increased (Figure 3.10A) and this was reflected in the increased contributions to each lineage, including Mac1+ (Figure 3.10B), CD4+ and/or CD8+ (Figure 3.10C) and CD19+ (Figure 3.10D) populations. Importantly, the proportional contribution from *Myb*$^{MRE/MRE}$ BM to each cell lineage remained the same compared to WT (Figure 3.10E). This indicates that while the PB

reconstitution potential of $Myb^{MRE/MRE}$ BM cells is increased, this is not due to biased production of any particular lineage, but possibly due to overall expansion of stem/progenitor compartments or increased downstream proliferation and/or differentiation.

Somewhat unexpectedly, given the results observed in the peripheral blood, end-point bone marrow analysis revealed no significant differences in reconstitution of stem and progenitor compartments between $Myb^{MRE/MRE}$ and WT donors (Figure 3.11), despite increased contribution from $Myb^{MRE/MRE}$ to the peripheral blood (Figure 3.10A-D). Taken together, this suggests that there may be increased proliferation and/or differentiation downstream of the stem/progenitor compartments in $Myb^{MRE/MRE}$ cells, resulting in the increase in PB reconstitution potential observed.

### 3.3.4 *In vitro* assay optimisation

As discussed in Chapter 1.2.3, retroviral transduction is a well-established model to study MLL-AF9 leukaemogenesis (Krivtsov *et al.* 2006; Somervaille *et al.* 2009; Krivtsov *et al.* 2013). While GMPs are the most widely used cell to target for transduction, their numbers are severely limiting when compared to the Lin- population as a whole, particularly if using individual mice as biological replicates. Therefore to identify an ideal target cell population in which to test the effect of the engineered *Myb* repression circuit on leukaemia initiation, GMPs (from a pool of 10 mice) or Lin- cells (from a pool of 3 mice) were purified from WT mice. Following 12 hours in culture, cells were transduced with control or *MLL-AF9* retrovirus with an mCherry reporter.

Importantly, based on mCherry expression, there was no significant difference in the transduction efficiency between the two cell populations (Figure 3.12A). An equivalent number of mCherry+ cells from each experimental condition were then plated onto methylcellulose containing recombinant mouse stem cell factor (SCF), mouse interleukin 3 (IL-3), mouse interleukin 6 (IL-6) and human erythropoietin (EPO). Following 7 days in culture, the resultant colonies were visualised and counted. In leukaemogenic colony forming assays, diffuse colonies are generally composed

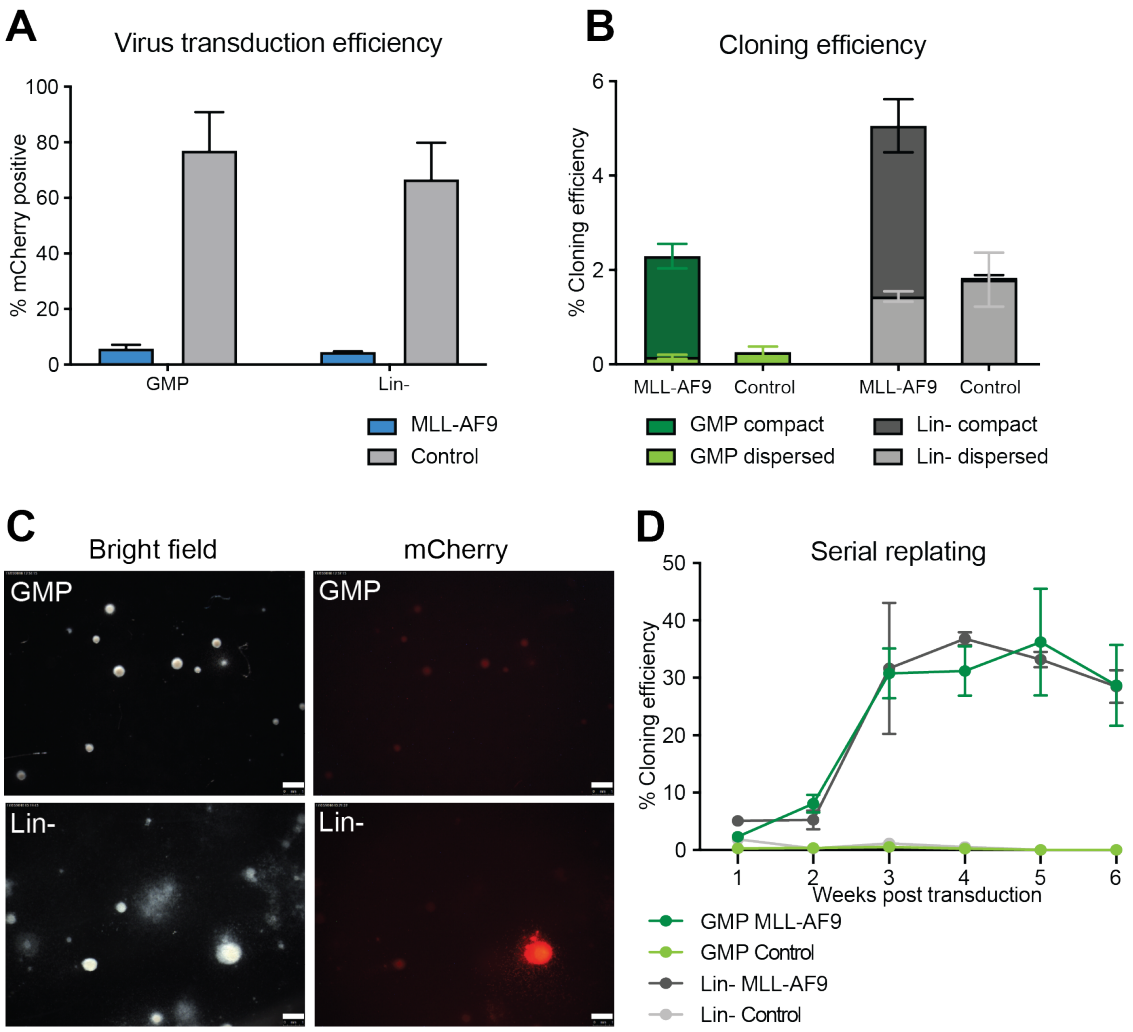



**Figure 3.12: Mouse GMP and Lin- cells have comparable transformation efficiency *in vitro* when transduced with *MLL-AF9*.** (A) Transduction efficiency in GMP and Lin- cells using *MLL-AF9* retrovirus (blue bars) compared with control (grey bars), based on mCherry expression by flow cytometry. (B) Stacked bar graph of % contribution of compact (darker shade) vs dispersed (lighter shade) colonies to % colonies formed by GMP (green bars) or Lin- cells (grey bars) transduced with *MLL-AF9* or control retrovirus. Cloning efficiency as a percentage of $10x^3$ cells plated. (C) Colony morphology of transduced GMP and Lin- cells under bright field and fluorescent microscopy. White bar represents 1mm. (D) Serial replating of GMP and Lin- cells transduced with *MLL-AF9* or control retrovirus. Cloning efficiency as a percentage of $10x^3$ cells plated. Error bars represent ± SD of 3 independent experiments.

of differentiating cells and do not go on to form further colonies, while at least a proportion of compact colonies will have self-renewal and proliferation potential (Hess *et al.* 2006; Krivtsov *et al.* 2006; Krivtsov *et al.* 2013). Interestingly, while the overall cloning efficiency was relatively reduced in MLL-AF9 transduced GMPs

compared with transduced Lin- cells, there were similar numbers of "compact" colonies produced by both cell types, with Lin- cells additionally producing a larger number of "diffuse" colonies (Figure 3.12B and C), suggesting that both cell populations possess similar leukaemia initiating potential.

After counting, colonies were solubilised and a similar number of cells from each transformed cell population was replated. This was repeated weekly for up to 6 weeks and no significant differences in compact colony numbers were observed over this period (Figure 3.12D). Additionally, from week 3 onward, there was an absence of diffuse colonies observed in cultures of both cell types (data not shown) and a large increase in cloning efficiency (Figure 3.12D), suggesting terminal differentiation and depletion of all non-transformed cells.

Although there were minor differences in haematopoietic compartments observed in $Myb^{\mathrm{MRE/MRE}}$ mice compared with WT mice, these differences were not observed in any of the cell populations thought to be susceptible to transformation by MLL-AF9 i.e. HSCs and GMPs (Collins *et al.* 2000; Krivtsov *et al.* 2006; W. Chen *et al.* 2008; Krivtsov *et al.* 2013; Stavropoulou *et al.* 2016). Therefore, taking into account that there were no significant differences in cloning efficiency between WT GMPs and Lin- cells and given the relative abundance of Lin- cells per mouse, downstream experiments were conducted with Lin- cells isolated from individual mice as biological replicates to allow us to capture individual biological variation.

### 3.3.5 $Myb^{\mathrm{MRE/MRE}}$ cells show a decreased trend in transformation potential *in vitro*

To test the effect of *Myb* repression on leukaemic transformation, an *in vitro* experimental approach was designed (Figure 3.13). Lin- cells were purified from the bone marrow of $Myb^{\mathrm{MRE/MRE}}$ or WT mice. Following 12 hours in culture, cells were transduced with *MLL-AF9/mCherry* retrovirus. 48 hours post transduction, mCherry+ cells were seeded directly into methylcellulose using fluorescence-activated cell sorting (FACS). Colonies were counted after 7 days in culture, solubilised and replated. Serial replating was carried out for a total of 10 weeks.

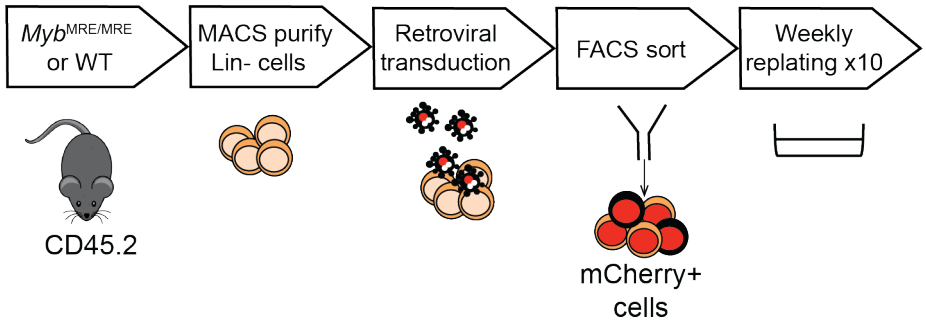



**Figure 3.13: Experimental design for *in vitro* assay.** Lin- cells were purified from the bone marrow of *Myb*$^{MRE/MRE}$ or WT mice using magnetic-activated cell sorting (MACS) technology and transduced with *MLL-AF9/mCherry* retrovirus. 48 hours post-transduction, mCherry+ cells were FACS sorted directly into supplemented methylcellulose. After 7 days in culture, colonies were counted, then solubilised and cells replated in methylcellulose. Serial replating was carried out for a total of 10 weeks.

Interestingly, there was a noticeable downward trend in the clonogenic potential of *MLL-AF9* transduced *Myb*$^{MRE/MRE}$ (MRE:MA9) Lin- cells compared to transduced WT (WT:MA9) Lin- cells (Figure 3.14), suggesting that the ability of *miR-196-b* to target *Myb* may have partially perturbed leukaemogenesis in these cells. However, this reduction was not statistically significant. This may be explained by the large variation observed between biological samples, likely due to the stochastic nature of effecting transformation in a mixed cell population purified from different mice. Alternatively, it is also possible that in MRE:MA9 cells, *Myb* repression was not sufficiently robust or indeed, that the repression circuit was not functional. Therefore, the next step was to validate the efficacy of the repression circuit in MRE:MA9 cells.

Given the low transduction efficiency of approximately 5% (Figure 3.12A and data not shown), and that individual mice were used as biological replicates, assays to confirm *Myb* KD were limited by the number of cells available at each replating time point. This was particularly challenging in the first three weeks post transduction where a large drop-off in cell numbers was observed. Therefore, multiple approaches were used to validate *Myb* KD following MLL-AF9 transduction of *Myb*$^{MRE/MRE}$ cells.

Firstly, *Myb* qRT-PCR was performed on samples harvested at weeks 2 and 5. At week 2, similarly low levels of *Myb* expression was observed in both MRE:MA9 and WT:MA9 cells (Figure 3.15A), possibly reflecting the low clonogenic potential

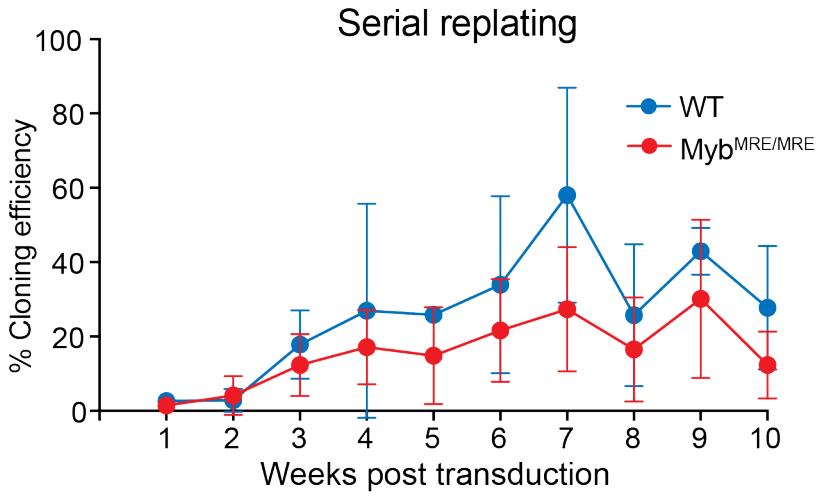



**Figure 3.14: No significant difference in the clonogenic potential of *MLL-AF9* transduced *Myb*$^{MRE/MRE}$ and WT Lin- cells.** Serial replating of MLL-AF9 transduced *Myb*$^{MRE/MRE}$ and WT Lin- cells. Cloning efficiency is plotted as a percentage of cells plated. WT:MA9 colonies are represented in blue and MRE:MA9 in red. Data points and error bars represent the mean ± SD, n = 3-4 from two independent experiments.

of cells at this time point (Figure 3.14). However, *Myb* mRNA levels were seen to diverge by week 5, with a statistically significant increase in *Myb* expression in WT:MA9 cells compared to MRE:MA9 cells at this time point (Figure 3.15A). This difference was also observed at the protein level at week 5. Western blot analysis (on separate samples from those used in Figure 3.15A) showed a loss of Myb protein expression in MRE:MA9 cells compared to WT:MA9 cells (Figure 3.15B). However, this result was not consistently reproducible between biological replicates and time points, and the cell numbers required for a robust Western blot result precluded examination of samples at time points prior to week 4 of the experiment. Therefore, Myb intracellular FACS (IFACS) was used as an alternative method to compare protein expression between samples. While this was only performed on samples from week 1, there is a significant decrease in the median fluorescence intensity of Myb in MRE:MA9 cells compared to WT:MA9 cells (Figure 3.15C), in contrast to the similar mRNA levels observed at week 2 (Figure 3.15A). The discrepancy observed between *Myb* mRNA and protein levels could be due to the miRNA circuit effecting translational repression rather than mRNA degradation. qRT-PCR and ICFACs will need to be repeated at multiple time points in the experiment to

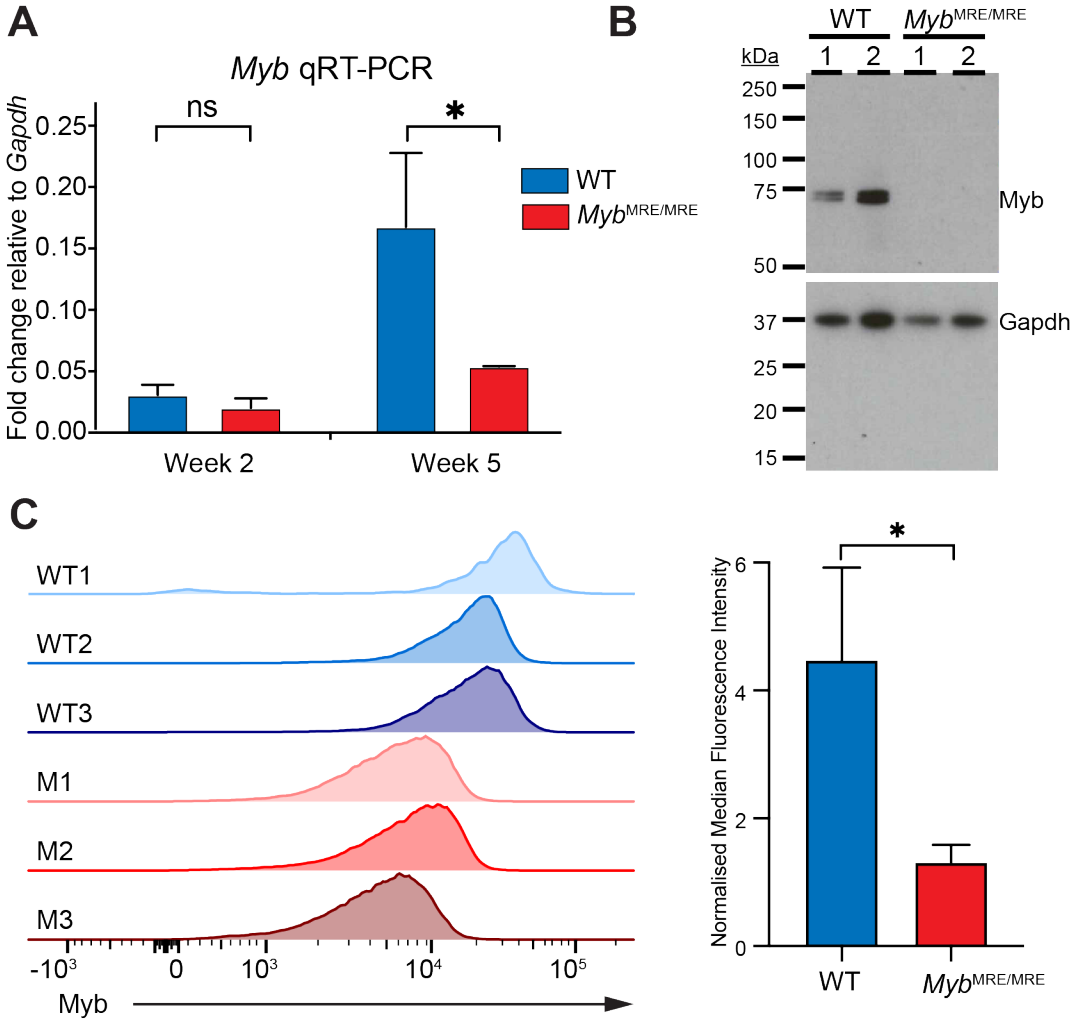



**Figure 3.15: *MLL-AF9* transduced *Myb*$^{MRE/MRE}$ cells have lower Myb expression compared to WT controls.** (A) *Myb* qRT-PCR in MRE:MA9 (red) and WT:MA9 (blue) cells at weeks 2 and 5 following transduction with MLL-AF9, normalised to housekeeping gene *Gapdh*. (B) Western blot analysis at week 5 following transduction using antibodies against Myb and Gapdh (as a loading control). Two biological replicates are shown. (C) Intracellular FACS analysis of Myb expression. Histogram and bar graph of median fluorescence intensity of WT:MA9 (WT1-3, blue) and MRE:MA9 (M1-3, red) at week 1 following transduction with *MLL-AF9*. Mean ± SD from three biological replicates. Significance tested by unpaired t-test, * = p <0.05

fully understand the dynamics of *Myb* mRNA and protein repression in our system during leukaemic transformation by MLL-AF9.

Although not conclusive, the above experiment provides several interesting observations. Firstly, it is possible to engineer a conditional miRNA circuit that utilises an endogenous miRNA (*miR-196-b*) to effect a change in target gene expression (*Myb*) under specific conditions (MLL-AF9 expression). Secondly,

while apparent reduction in Myb is sufficient to produce a downward trend in the clonogenic potential of MRE:MA9, these cells still acquire abnormal long-term colony forming ability. This suggests that a reduction in Myb expression during leukaemic transformation is insufficient to fully abrogate increased clonogenicity mediated by MLL-AF9, indicating that there is either a low threshold requirement for Myb in MLL-AF9 leukaemogenesis, or that MLL-AF9 is able to utilise alternative pathways to drive an aberrant self-renewal and proliferation programme *in vitro*.

### 3.3.6 *In vivo* experimental approach to test the requirement for Myb in leukaemic transformation

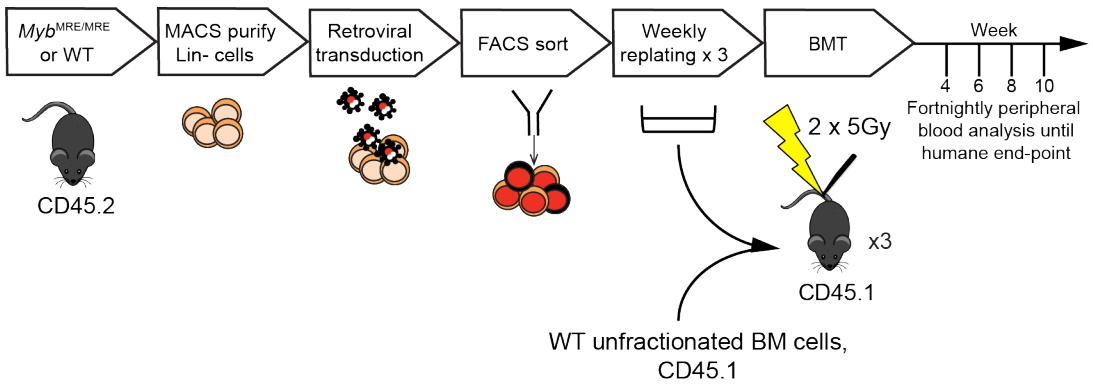



**Figure 3.16: Experimental design for *in vivo* assay.** BM cells from CD45.2 $Myb^{MRE/MRE}$ or WT mice were purified, transduced with *MLL-AF9* retrovirus and plated in methylcellulose as per the *in vitro* assay (Figure 3.13). After 3 weeks of serial replating, 1 x $10^6$ donor cells (CD45.2) and 2.5 x $10^5$ unfractionated BM helper cells (CD45.1) were injected into the tail vein of each lethally irradiated CD45.1 recipient. To monitor engraftment, peripheral blood analysis was performed four weeks following transplantation and fortnightly thereafter until mice were sacrificed.

As it appeared that MRE:MA9 cells were still able to form long-term colonies in culture despite reduced Myb expression, it was important to ascertain if these cells were also able to give rise to leukaemia *in vivo*. To test the effect of *Myb* repression on leukaemic transformation *in vivo*, both MRE:mA9 and WT:MA9 (CD45.2) cells were expanded by serial replating for three weeks before co-transplantation with CD45.1 helper cells via tail vein injection into lethally irradiated (2 x 5Gy) CD45.1 recipient mice. Cells from each donor mouse were transplanted into three recipient mice and four donors were used per genotype. Each recipient was injected with 1 x

$10^6$ donor cells (CD45.2) and 2.5 x $10^5$ unfractionated BM helper cells (CD45.1). To monitor engraftment, peripheral blood analysis was performed four weeks following transplantation and fortnightly thereafter until the mice were sacrificed (Figure 3.16).

### 3.3.7 *MLL-AF9* transduced *Myb*$^{MRE/MRE}$ and WT cells are immunophenotypically similar prior to transplantation

Given that *MLL-AF9* was transduced into a mixed target cell population (Lin- cells), it was important to identify any differences in potential leukaemic stem cell (LSC) populations between MRE:MA9 and WT:MA9 cells prior to transplantation. Therefore, when cells were resuspended for replating on a weekly basis, a proportion were subjected to immunophenotypic analysis as per the gating strategy in Figure 3.17.

The staining panel was designed based on data from Krivtsov *et al.* 2013, which showed that following expansion in methylcellulose, *MLL-AF9* transduced HSCs and GMPs gave rise to distinct cell populations with demonstrable LSC potential in limiting dilution transplantation assays. These populations were defined as follows: 1) Lin+, 2) LSK (Lin- c-Kit+ Sca1+), 3) Lin- c-Kit-, 4) HSC-like (Lin- c-Kit+ Sca1+ CD34-) and 5) GMP-like (Lin- c-Kit+ CD16/32+ CD34+).

Importantly, immunophenotypic analysis of MRE:MA9 and WT:MA9 cells revealed no statistically significant differences by unpaired t-test in these putative LSC populations (Figure 3.17B-G). While this does not necessarily prove that there were no functional differences between the two experimental groups, there was at least no gross bias for a particular LSC population as defined by immunophenotypic analysis prior to transplantation.

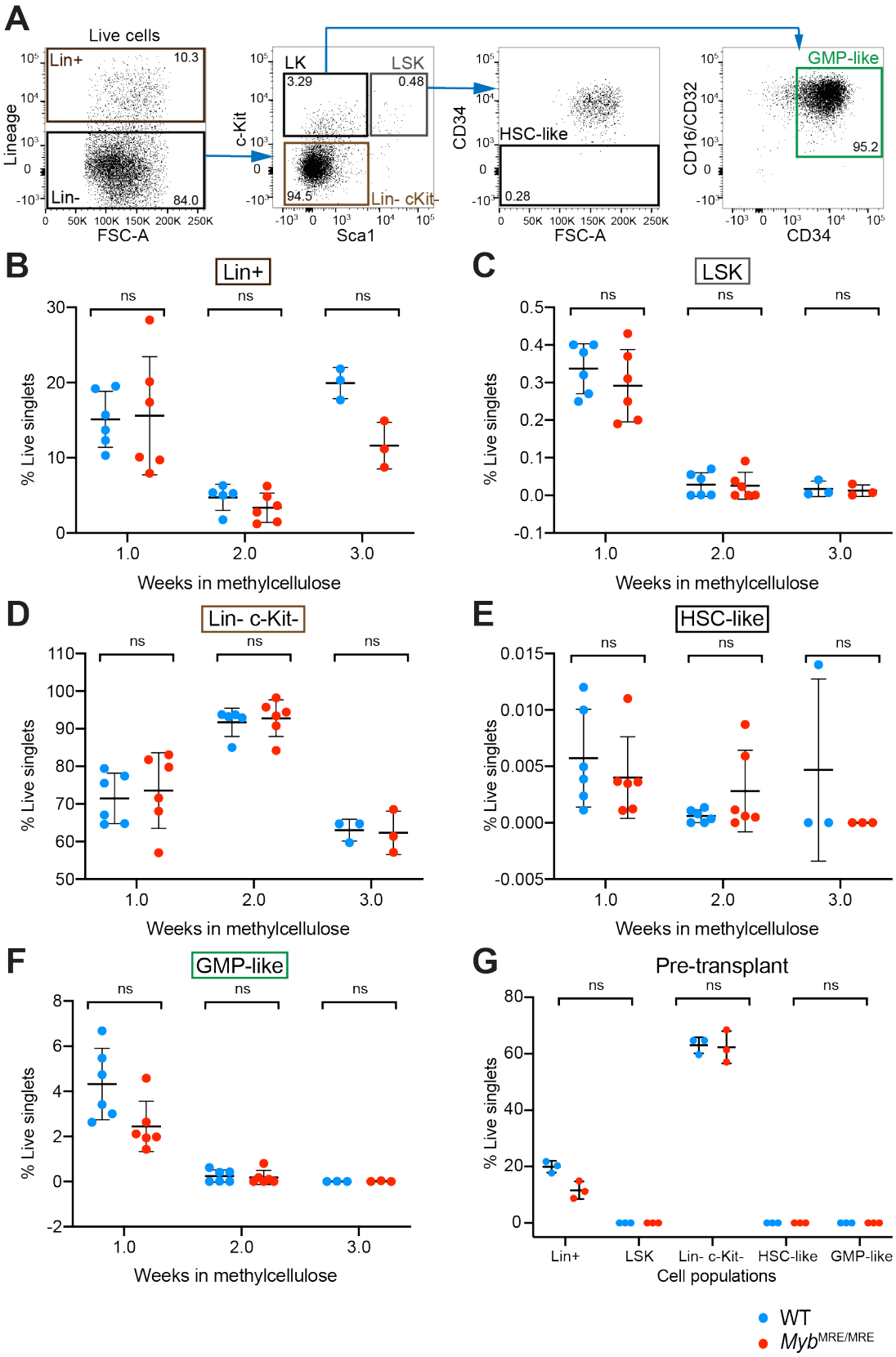



**Figure 3.17: MLL-AF9 transduced *Myb*$^{MRE/MRE}$ and WT cells are immunophenotypically similar prior to transplantation.**

### 3.3.8 MLL-AF9 transduced *Myb*$^{MRE/MRE}$ cells give rise to an attenuated leukaemia phenotype *in vivo*

Following three weeks of expansion in methylcellulose, an equivalent number of donor cells (CD45.2) were transplanted via tail-vein injection into lethally irradiated (2 x 5Gy) recipient mice (CD45.1). Recipients were weighed daily and monitored according to clinical (hunched posture, ruffled fur, labored breathing, weight loss, loss of mobility and non-responsive to stimuli) and laboratory (>40% PB donor chimerism, symptomatic anaemia and thrombocytopenia) criteria for symptoms and signs consistent with leukaemia. If any end-point criteria were fulfilled, mice were humanely sacrificed under Schedule 1 as per UK Home Office standards and regulations. At the time of sacrifice, peripheral blood, tibiae, femurs and splenic tissue were harvested for analysis and to confirm that leukaemia was the most likely cause of morbidity.

Since both MRE:MA9 and WT:MA9 cells gave rise to long-term colonies *in vitro*, it was perhaps not surprising that transduced cells from both genotypes were capable of engraftment, where the threshold for engraftment was defined as >0.1% donor chimerism in the peripheral blood (Figure 3.18A). Consistent with published viral transduction MLL-AF9 models (Krivtsov *et al.* 2013; Krivtsov *et al.* 2006; Somervaille *et al.* 2009; Krivtsov *et al.* 2013), donor cells contributed exclusively to the myeloid lineage (Mac1+) in peripheral blood ((Figure 3.18B-D).

However, the most striking finding was that MRE:MA9 recipient mice had a considerable survival advantage over WT:MA9 recipient mice ($p = 0.0075$) (Figure 3.18M), with a median survival time of 105 days and 67 days respectively. Significantly, fortnightly peripheral blood chimerism analysis revealed that MRE:MA9 cells had a markedly increased latency of engraftment compared with WT:MA9

---

**Figure 3.17:** continued. (A) Gating strategy used to define putative LSC populations. Frequency of Lin+ (B), LSK (C), Lin- c-Kit- (D), HSC-like (E) and GMP-like (F) populations in WT:MA9 (blue) and MRE:MA9 (red) cells at weeks 1-3. (G) Frequency of all populations prior to transplantation (week 3) in WT:MA9 (blue) and MRE:MA9 (red) cells. Bars represent the mean ± SD, n = 3-6 of two experiments.

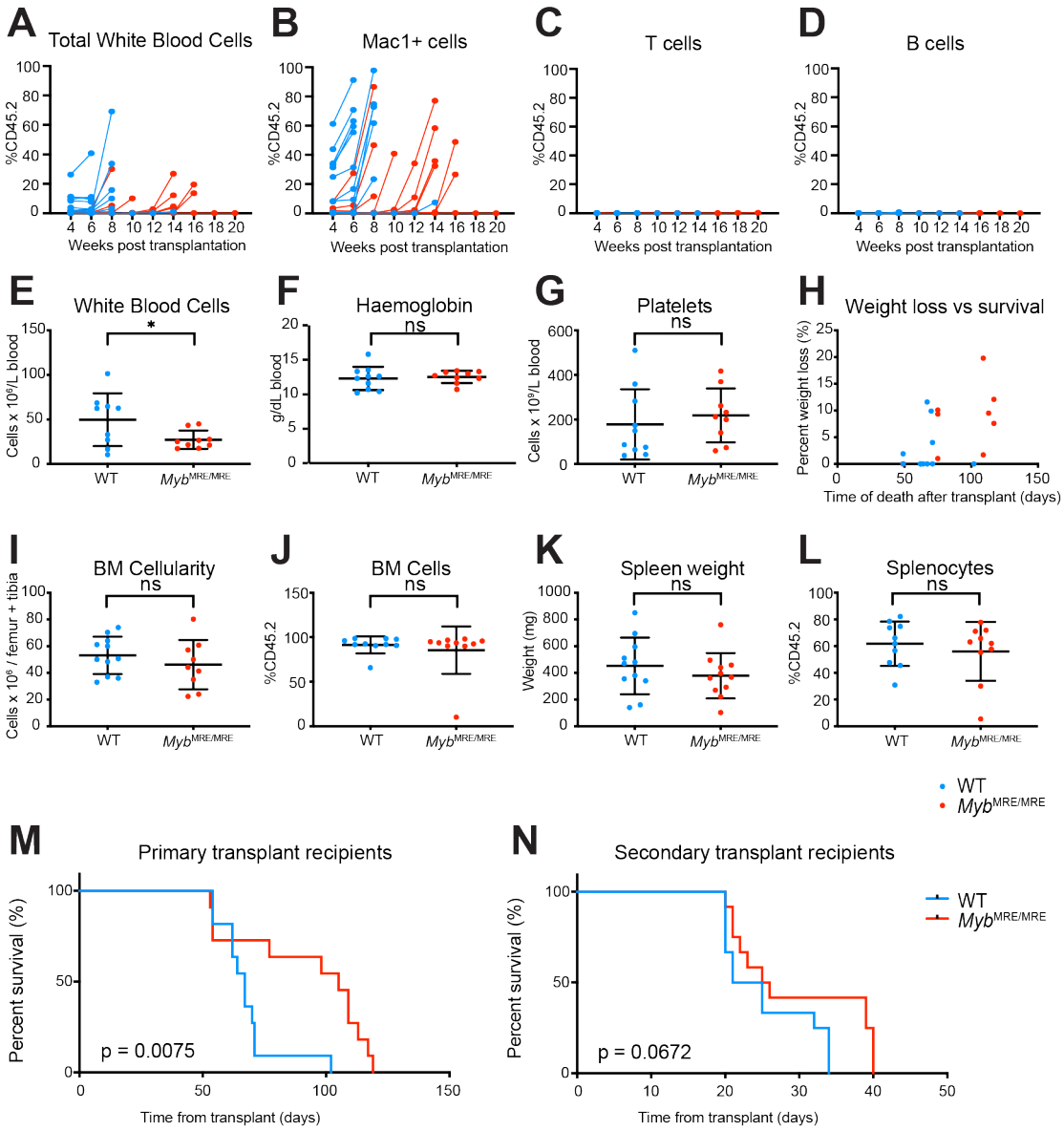



**Figure 3.18: Mice transplanted with MLL-AF9 transduced *Myb*$^{MRE/MRE}$ cells have a survival advantage.** Immunophenotypic analysis of level of CD45.2 chimerism of (A) total white blood cells, (B) myeloid (Mac1+), (C) T (CD4+ and/or CD8+) and (D) B (CD19+) cells in peripheral blood of CD45.1 recipient mice post-transplantation of MRE:MA9 or WT:MA9 cells with CD45.1 support BM cells. Peripheral blood count parameters, (E) white blood cell count, (F) haemoglobin, and (G) platelets count at end-point. (H) Maximum percent weight loss at the time of death post-transplant, compared to pre-transplant weight. (I) Bone marrow cellularity and (J) level of CD45.2 chimerism at end point. (K) Spleen weight and (L) level of donor cell infiltration at end point. Each dot represents a data from individual recipient mice with recipients of WT:MA9 cells in blue and MRE:MA9 cells in red. Bars represent the mean ± SD, n = 11 from 2 experiments. Significance tested by unpaired t-test, ns = not significant, * = p <0.05. Survival curves for (M) primary and (N) secondary transplantation experiments. The x-axis represents time of death after transplantation. Significance tested by log-rank test.

cells (10 weeks ± 4.382 vs 5.3 weeks ± 3.133, p = 0.0086) (Figure 3.18A and B) and overall, MRE:MA9 recipient mice had a lower leukaemic burden in the peripheral blood compared to WT:MA9 recipients at end-point analysis (Figure 3.18A and E).

At the time of sacrifice, in keeping with the clinical features of leukaemia, all recipient mice appeared to be mildly anaemic (average Hb 12.5g/dL and 11.4g/dL in MRE:MA9 and WT:MA9 mice respectively; normal range 13.9-15.9 g/dL) (Figure 3.18F) and thrombocytopenic (average platelet count 218 x $10^9$/L and 178 x $10^9$/L in MRE:MA9 and WT:MA9 mice respectively; normal range 700-1300 x $10^9$/L) (Figure 3.18G), with splenomegaly (Figure 3.18K) demonstrating on average ∼60% infiltration by CD45.2 positive cells (Figure 3.18L) and near 100% donor chimerism in the bone marrow (Figure 3.18J). Of interest, MRE:MA9 recipient mice were observed to experience a marked amount of weight loss compared to WT:MA9 recipient mice (Figure 3.18H). Indeed, weight loss was the most frequent parameter to trigger humane end-point criteria in MRE:MA9 recipient mice. Ironically, this fits with the much more aggressive disease phenotype displayed by WT:MA9 cells, as WT:MA9 recipient mice did not survive long enough to experience weight loss.

The survival advantage of MRE:MA9 cells was only restricted to primary transplant recipients. In secondary transplant recipients, there was a convergence of survival curves (p = 0.0672) and median survival was reduced to 23 days in WT:MA9 recipients and 25.5 days in MRE:MA9 recipients (Figure 3.18N).

Taken together, these results suggest that in keeping with the slight reduction in cloning efficiency observed *in vitro*, reduction of *Myb* expression during leukaemia transformation reduces leukaemia fitness *in vivo*. However, as demonstrated by data from secondary transplantation, this effect is not permanent, and under selection pressure, MLL-AF9 transformed $Myb^{\mathrm{MRE/MRE}}$ cells appeared to be able to recover from the initial effects of lower *Myb* expression and evolve an aggressive and sustained leukaemia phenotype.

## 3.4 Discussion

In summary, we have successfully engineered a conditional *Myb* repression circuit mediated by an endogenous miRNA in a transgenic mouse model, allowing us to study the effect of *Myb* repression on leukaemia initiation (Figures 3.2 & 3.4). We demonstrated that the expression of *MLL-AF9* in transgenic mouse BM cells was coupled with significantly lower levels of Myb expression compared with WT controls, albeit only at a few relatively early time points (Figure 3.15), indicating that the repression circuit was functional. This reduction in Myb expression was correlated with a trend towards decreased clonogenic potential *in vitro* (Figure 3.14), as well as a convincing reduction in leukaemia fitness *in vivo*, as demonstrated by the survival advantage seen in primary transplant recipient mice (Figure 3.18), suggesting that at least early on in leukaemic transformation, a threshold level of Myb is required to propagate an aggressive leukaemia phenotype.

Based on the experimental data generated, there appears to be several caveats to these conclusions and further work needs to be done to address these issues. Firstly, we assume that the disruption to leukaemogenesis in $Myb^{\mathrm{MRE/MRE}}$ cells is due to *Myb* repression alone and not because of significant *miR-196-b* sponging by the engineered *miR-196-b* MREs and de-repression of *miR-196-b* targets. While this is unlikely due to the massive upregulation of *miR-196-b* ($\log_2 8$), it would nonetheless be helpful if the absolute transcript number of mature *miR-196-b*, *Myb*, and potential *miR-196-b* targets could be obtained perhaps via Nanostring, digital droplet PCR or other suitable technique. If the levels of *miR-196-b* are found to be significantly reduced in transduced $Myb^{\mathrm{MRE/MRE}}$ cells compared with WT cells, regulation of *miR-196-b* expression by a leukaemia specific promoter to correct for differences in *miR-196-b* expression between the two genotypes would allow us to address the role of Myb in transformation more accurately. In addition, while not completely applicable for reasons discussed in Chapter 1.3.3, a useful control that would support a Myb-specific requirement for leukaemogenesis would be to knockdown Myb in WT cells at the time of transduction with MLL-AF9, for example with shRNA against *Myb*.

Secondly, while the activity of the circuit was designed to be conditional upon MLL-AF9 expression, analysis of bone marrow composition and function have demonstrated an expansion of the MPP1 and MPP2 compartments (Figure 3.5), and a reduction in the ALP and BLP compartments of $Myb^{\mathrm{MRE/MRE}}$ mice (Figure 3.7), albeit from a small number of samples. This raises the possibility that the system may not be absolutely stringent, particularly as the *HoxA* cluster within which *miR-196-b* is located may be actively transcribed in certain stem/progenitor cell types at a time when Myb expression is simultaneously required. It is possible that these observations could simply be attributed to biological variation due to the relatively small sample number examined (n = 4), therefore more replicates will be needed for us to be confident in the robustness of the circuit. To this end, it would also be useful to perform additional complementary experiments with individual stem/progenitor compartments isolated from WT and $Myb^{\mathrm{MRE/MRE}}$ mice, including *in vitro* colony culture to assess functional output, as well as gene expression analysis of *Myb, miR-196-b* and known *Myb* targets using Fluidigm qRT-PCR.

Nevertheless, given that there are no gross haematopoietic deficiencies in $Myb^{\mathrm{MRE/MRE}}$ mice, both at steady state and under stressed competitive transplantation conditions (Figures 3.10 & 3.11), the extent of *Myb* repression, if in effect, is likely to be at a low level and affecting only specific cell types, therefore by no means invalidating the observations from our transformation experiments so far.

As mentioned above, we observed a significant survival advantage in primary transplant recipients of MRE:MA9 cells, which was lost upon secondary transplantation (Figure 3.18), suggesting that the cells have acquired resistance to the miRNA circuit. Two immediate questions arise from these observations which this chapter has not been able to address, but would be interesting to explore. Firstly, what are the dynamics of Myb repression and downstream effects during leukaemia transformation that result in an initial milder phenotype? Secondly, how do transgenic cells evolve a more aggressive phenotype upon serial transplantation?

We have taken steps towards answering these questions by sending paired samples for RNA-sequencing, including samples from the post-transduction weekly *in vitro*

expansion, pre-transplantation cells and end-point bone marrow from primary transplants. Analysis of this data will allow us to examine and compare changes in the transcriptome over time in both WT:MA9 and MRE:MA9 cells. We will also be able to track the dynamics of *miR-196-b* levels and *Myb* repression, and its effect on known *Myb* target genes. In addition, the experimental approach may be further refined. For example, instead of heterogenous Lin- cells, a defined target cell population such as GMPs could be used in repeat experiments. This will control for the stochastic nature of viral transformation in a mixed population and facilitate our interpretation of experimental data.

Lastly, there are several possible mechanisms through which *MLL-AF9* transformed transgenic cells may escape the effects of Myb repression. This could involve inactivation of the circuit mechanism, for example by preventing recognition of the *Myb* mRNA by *miR-196-b*. This could be achieved via alternative polyadenylation, where the 3'UTR of the *Myb* transcript is cleaved and polyadenylated at an earlier site upstream of the engineered MREs, excluding them from the mature mRNA. Alternative polyadenylation is known to be a feature of cancer cells (Mayr *et al.* 2009; Erson-Bensan *et al.* 2016; Masamha *et al.* 2018). Alternatively, the MREs or *miR-196-b* itself could have acquired loss of function mutations, but this is less likely to be a common mechanism for all samples because the rapid and consistent recovery argues against the appearance of specific DNA mutations, given the relative genetic stability of MLL-AF9 leukaemias. The final and most likely explanation is that under continuous selection pressure from the circuit, the transformed $Myb^{\mathrm{MRE/MRE}}$ cells adapt by utilising alternative, Myb-independent transcription pathways to mediate leukaemogenesis, possibly via epigenetic rather than mutational mechanisms, and are thus able to evolve an aggressive phenotype. The RNA-seq analysis may be able to answer some of these questions, particularly if there is an enrichment of *Myb* isoforms with truncated 3' UTRs, downregulation of *miR-196-b* expression or global transcriptomic differences between MRE:MA9 and WT:MA9 samples, and we eagerly anticipate these results.

# 4

# MYB is sufficient for establishing an enhancer-like element

## Contents


**4.1 Introduction . . . . . 74**
4.1.1 MYB structure, regulation and function . . . . . 74
4.1.2 Understanding the role of MYB in establishing enhancer function . . . . . 79
4.1.3 The mouse embryonic stem cell TetR system . . . . . 80
**4.2 Aims . . . . . 82**
**4.3 Results . . . . . 83**
4.3.1 Myb binding is sufficient for the recruitment of functional CBP and P300 . . . . . 83
4.3.2 The Myb transactivation domain is sufficient for stable binding of its known interaction partners . . . . . 86
4.3.3 The Myb transactivation domain is sufficient for the recruitment of transcription activation complexes and RNAPII . . . . . 90
4.3.4 Myb is sufficient for establishing an enhancer-like element at chromatin . . . . . 91
4.3.5 A single TetR binding site is sufficient for Myb dependent enhancer-like activity . . . . . 95
**4.4 Discussion . . . . . 98**


## 4.1 Introduction

### 4.1.1 MYB structure, regulation and function

*MYB* encodes a nuclear protein with three conserved functional domains – an N-terminal DNA binding (DB) domain, a central transactivation (TA) domain and a C-terminal negative regulatory (NR) domain (Klempnauer *et al.* 1982; Boyle *et al.*

1984; Sakura *et al.* 1989; T. Nomura *et al.* 1993; Lipsick 1996) (Figure 4.1).

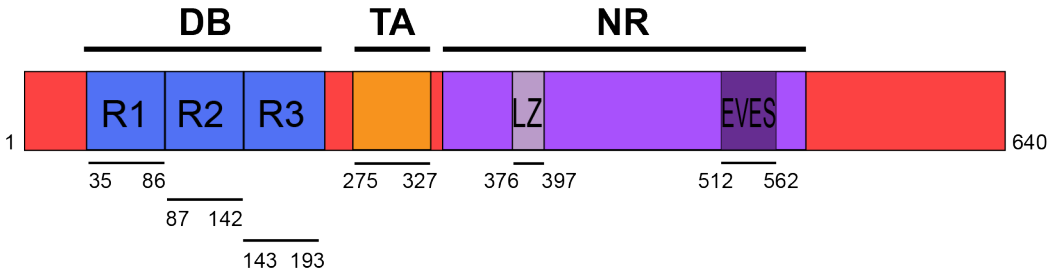



**Figure 4.1: Schematic of MYB protein structure and conserved domains.** Diagram of MYB protein structure, with the N-terminus on the left and the C-terminus on the right. Numbers in black represent amino acid residues corresponding to human MYB. The DNA binding (DB) domain with tandem repeats R1-3 is shaded in blue, the transactivation (TA) domain in yellow and the negative regulatory (NR) domain containing the leucine zipper (LZ) and EVES motif in purple.

The DB domain consists of three imperfect tandem repeats of 50-52 amino acids, termed R1, R2 and R3 (Sakura *et al.* 1989; Weston *et al.* 1989; Saikumar *et al.* 1990). Each repeat consists of conserved tryptophan triplets separated by 18-19 amino acids (Frampton *et al.* 1991) which cluster within a hydrophobic core and maintain the helix-turn-helix structure of the DB domain (Saikumar *et al.* 1990). Deletion analyses have demonstrated that R2 and R3 are absolutely required for DNA binding, while R1 is dispensable, but has a role in stabilising the interaction with DNA (Saikumar *et al.* 1990; Howe *et al.* 1991; Tanikawa *et al.* 1993). Complementary to this, R2 and R3 have been shown by nuclear magnetic resonance (NMR) to interact in a co-operative manner to bind the major groove of DNA (Ogata *et al.* 1992; Jamin *et al.* 1993; Ogata *et al.* 1994; Tahirov *et al.* 2001).

Deletion and linker insertion analyses have located the TA domain downstream of the DB domain (Sakura *et al.* 1989; Weston *et al.* 1989; Kalkbrenner *et al.* 1990). This domain corresponds to amino acids 275-327 in human MYB (Kalkbrenner *et al.* 1990), and amino acids 241-325 in mouse Myb (Sakura *et al.* 1989). The transactivation potential of MYB is largely mediated by the interaction of its TA domain with co-activating proteins, the most well described being CBP and P300 (Dai *et al.* 1996; Sandberg *et al.* 2005; Pattabiraman *et al.* 2014).

Single amino acid substitutions within the MYB TA domain which disrupt its interaction with the KIX domains of CBP or P300 have clearly demonstrated a loss of MYB transactivating and transforming ability (Parker *et al.* 1999; Zor *et al.* 2004; Sandberg *et al.* 2005; Pattabiraman *et al.* 2009; Papathanasiou *et al.* 2010; Pattabiraman *et al.* 2014). Mouse strains homyzygous for such mutations e.g. M303V and E308G, are observed to have multi-lineage haematopoietic abnormalities in their bone marrow with resultant in anaemia and marked thrombocytosis (Sandberg *et al.* 2005; Papathanasiou *et al.* 2010), reminiscent of $Myb^{kd/kd}$ mice (Emambokus *et al.* 2003). Importantly, Myb M303V is unable to transform FDB-1 cells (a murine haematopoietic cell line) (Sandberg *et al.* 2005), and c-Kit+ cells from $Myb^{E308G/E308G}$ mice cannot be transformed by the myeloid leukaemia oncogenes AML1-ETO, AML1-ETO9a, MLL-ENL, and MLL-AF9, whether *in vitro* or when transplanted into irradiated recipient mice (Papathanasiou *et al.* 2010).

More recently, the MYB TA domain has been shown to interact with the TAF12/TAF4 heterodimer, subunits of TFIID, which also includes the TATA-box binding protein (TBP) and 11 other TBP-associated factors (TAFs) (Xu *et al.* 2018). Importantly, this provides a mechanism through which MYB is able to directly stabilise transcription activating complexes at gene regulatory elements.

The NR domain of MYB is so termed as deletion of this region increases the transactivating and transforming potential of MYB (Sakura *et al.* 1989; Weston *et al.* 1989; Grasser *et al.* 1991; Hu *et al.* 1991; Gonda *et al.* 1996). Two structures within the NR domain of MYB, the leucine zipper (LZ) structure and the EVES motif, are thought to play a part in attenuating its activity. Firstly, the LZ, located at amino acids 376-397 and containing a 10-amino acid sub-region known as the "FAETL" motif (Fu *et al.* 1996), is thought to mediate homodimerisation of MYB. MYB dimers - possibly due to obscuration of the DB domain - are unable to bind DNA (T. Nomura *et al.* 1993). This is corroborated by the fact that the DNA binding capability of MYB has been observed to increase significantly when the NR domain including the LZ structure is deleted (Ramsay *et al.* 1991; Tanaka *et al.* 1997). Furthermore, the LZ has been shown to interact with proteins such as P67

**Table 4.1:** MYB post-translational modifications and effect on function

| Amino acid residue | MYB domain | Effect on MYB function | Modifying protein/ conjugate | Reference |
|---|---|---|---|---|
| *Phosphorylation* | | | | |
| Ser-11, Ser-12 | DB | Decreased DNA binding | CKII | Lüscher *et al.* 1990; Oelgeschläger *et al.* 1995 |
| Ser-8 | DB | Partial decreased DNA binding | PKA | Lüscher *et al.* 1990 |
| Ser-528 | NR | Decreased transactivation | $P42^{MAPK}$ | Aziz *et al.* 1995; Aziz *et al.* 1993; Dash *et al.* 1996; Vorbrueggen *et al.* 1996; Miglarese *et al.* 1996 |
| Ser-116 | DB | Decreased DNA binding & co-operativity with CEBP | PKA | K. B. Andersson *et al.* 2003 |
| Thr-486 | NR | Increased proteolytic turnover | $P38^{MAPK}$ | Bies *et al.* 2013 |
| *Acetylation* | | | | |
| Lys-480, Lys-485 | NR | Increased DNA binding & transactivation | P300 | Tomita *et al.* 2000 |
| Lys-438, Lys-441 | NR | Increased transactivation & interaction with CBP | CBP | Sano *et al.* 2001 |
| *SUMOylation* | | | | |
| Lys-499, Lys-523 | NR | Increased degredation/proteolytic turnover | SUMO1/2/3 | Bies *et al.* 2013; Bies *et al.* 2002; Sramko *et al.* 2006 |
| *Ubiquitination* | | | | |
| Widespread | - | Increased proteolytic turnover | FBW7 | Kitagawa *et al.* 2010 |

DB = DNA binding; NR = negative regulatory.

that have an inhibitory effect on MYB function (Kanei-Ishii *et al.* 1992; Favier *et al.* 1994; Tavner *et al.* 1998). The EVES motif, containing a highly conserved string of amino acids (in single-letter amino acid code), is located at amino acids 512-562 of the MYB NR domain. Phosphorylation of Ser-528 within the EVES motif facilitates its interaction with the DB domain (Aziz *et al.* 1993; Aziz *et al.* 1995; Dash *et al.* 1996; Tanaka *et al.* 1997). This intramolecular interaction is

thought to compete with other activating proteins e.g. P100 (which interacts with MYB via its own EVES motif) from binding with the DB domain, thus negatively regulating MYB activity (Dash *et al.* 1996; Leverson *et al.* 1998a).

**Table 4.2:** MYB direct protein interactions.

| **Partner** | **Effect on MYB** | **Myb domain** | **References** |
|---|---|---|---|
| CEBPB | Co-operates | DB | Mink *et al.* 1996 |
| P100 (via Pim-1) | Co-activator | DB | Dash *et al.* 1996; Leverson *et al.* 1998a |
| HSF3 | Unknown | DB | Kanei-Ishii *et al.* 1997 |
| MAF | Inhibitor | DB | Hedge *et al.* 1998 |
| CYP40 | Inhibitor | DB | Leverson *et al.* 1998b |
| CEBPE | Co-operates | DB | Verbeek *et al.* 1999 |
| RARA | Inhibitor | DB | Pfitzner *et al.* 1998 |
| TAF12/TAF4 | Co-activator | TA | Xu *et al.* 2018 |
| p300 | Co-activator | TA | Kasper *et al.* 2002; Sandberg *et al.* 2005; Papathanasiou *et al.* 2010; Pattabiraman *et al.* 2014 |
| CBP | Co-activator | TA | Dai *et al.* 1996; Kiewitz *et al.* 1997; Oelgeschläger *et al.* 1996; Papathanasiou *et al.* 2010; Pattabiraman *et al.* 2014 |
| Menin | Associate with MLL | TA | Jin *et al.* 2010 |
| MYB DB domain | Inhibitor | NR (EVES) | Dash *et al.* 1996 |
| MYB LZ | Inhibitor | NR (LZ) | T. Nomura *et al.* 1993 |
| p160/p67 | Inhibitor | NR (LZ) | Tavner *et al.* 1998 |

DB = DNA binding; TA = transactivation; NR = negative regulatory; LZ = leucine zipper; EVES = EVES motif.

The activity of MYB is tightly controlled by multiple mechanisms. Besides the effects mediated by the NR domain as mentioned above, MYB activity is further regulated via post-translational modifications to its DB and NR domains as summarised in Table 4.1. Furthermore, MYB has been found to interact directly with a large number of proteins (Table 4.2). These interacting proteins may be categorised into three distinct classes: 1) MYB regulatory proteins, 2) co-operating transcription factors that directly interact with MYB and 3) transcriptional co-activators. Overall, with the exception of co-operating TFs that interact with the

DB domain, the DB and NR domains of MYB generally associate with proteins that reduce MYB activity, while as its name suggests, the TA domain largely interacts with co-activating proteins.

### 4.1.2 Understanding the role of MYB in establishing enhancer function

As discussed in Chapter 1.1, gene regulation is a multi-faceted process in which TFs play a key role. Multiple examples of TFs acting alone or in combination have demonstrated their role in orchestrating distinct aspects of gene regulation - beginning with initiating and maintaining chromatin accessibility, to the assembly of activating complexes at gene regulatory elements, association of enhancers with their cognate promoters, and culminating in the final output of active transcription by RNAPII - in a temporally and spatially specific manner.

As highlighted in Chapter 1.3.3, recent evidence suggests that one possible function of MYB is to establish enhancer activity. However, as the control of transcription is such an intricate and interconnected process, dissecting the mechanistic role of MYB is not without its challenges.

Firstly, despite providing an appropriate cellular context, attempting to study the function of MYB at native gene regulatory elements would be complicated by all components required for active transcription at these loci. It would be further confounded by the fact that active enhancers and promoters are frequently co-bound by multiple TFs (Buecker *et al.* 2012; Schoenfelder *et al.* 2015).

Historically, reporter assays where an endogenous or synthetic gene regulatory element is tethered a short distance upstream of a reporter gene e.g. luciferase, and transfected into cells and tissues of interest have been used as a relative measure of transcriptional activity following TF binding. While these experiments provide tissue specific information on TF expression and transactivation ability, they lack the resolution and scope to analyse the intervening molecular chain of events from TF binding leading up to active transcription. Theoretically, chromatin immunoprecipitation (ChIP) could be performed to test for the presence of proteins

stabilised by MYB binding. However, if a native regulatory sequence were to be used, recruitment of co-factors, chromatin proteins and active transcription complexes due to the activity of other TFs binding at that same element cannot be ruled out. On the other hand, even if the element were to contain only a single MYB binding sequence, because reporter constructs are transfected into cells of interest, they will frequently be episomal and not fully chromatinised and thus lack physiological relevance (Inoue *et al.* 2017). In the event that the construct is stably integrated, this tends to occur at random (and usually accessible) genomic loci. In both situations, this precludes robust examination of changes in chromatin accessibility and recruitment of protein complexes that can be confidently attributed to the TF of interest. Furthermore, as the regulatory element is seldom more than a few hundred bases upstream of the reporter gene, it would be impossible to study TF function in the context of its binding at distal regulatory elements e.g. enhancers that are kilobases away.

Additionally, in order to fully understand the molecular role of MYB, it would be important to determine if MYB binding alone was sufficient to recruit co-activating complexes to chromatin. Various *in vitro* techniques have been used to analyse protein:protein interactions, including pull-down, co-immunoprecipitation and affinity purification assays. Although these are useful techniques to initially identify and test physical interactions between proteins *in vitro*, the interactions which are identified lack cellular context. Thus, it would be difficult be confident that these are indeed *bona fide* interactions *in vivo* and not simply artefacts of the techniques used.

Therefore, to establish if MYB is sufficient for establishing enhancer activity, it would be important to study its function in relative isolation of transcription and free from the activity of other TFs, whilst still maintaining physiological relevance.

### 4.1.3 The mouse embryonic stem cell TetR system

One system that allows the study of unique protein interactions in a cellular context is the mouse embryonic stem cell (mESC) Tetracycline repressor (TetR) system developed by Blackledge *et al.* 2014. This novel system utilises the interaction

between the bacterial TetR and the Tetracycline operator (TetO) to target proteins or protein domains to a chromatin environment free from the complexities and regulatory influences of native gene regulatory elements. This was achieved by inserting into chromosome 8 of mESCs, a sequence containing a TetO array (comprising fourteen TetO DNA-binding sites, each separated by ∼20bp) embedded within a 170kb stretch of DNA corresponding to a gene desert region on human chromosome 7 (Figure 4.2A). Thus, the TetO array is effectively flanked by long stretches of inert chromatin that lacks identifiable genes, CpG islands, chromatin modifications (including H3K4me3, H3K27me3, H3K9ac, H2AK119ub, H3K36me3, H4K20me2 and H3K9me3) or the presence of RNAPII. Thereafter, fusing a protein of interest to the TetR DNA-binding domain (TetR-fusion) would permit its *de novo* recruitment to the chromatinised TetO array (Figure 4.2B) and the direct consequences of fusion protein occupancy could be examined by ChIP qPCR.

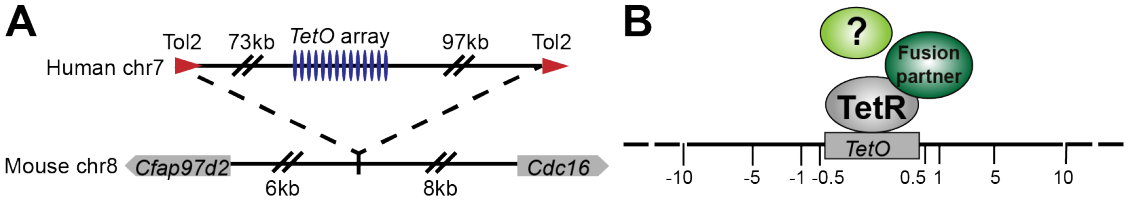



**Figure 4.2: The TetR system.** (A) The *TetO* array at its integration site on mouse chromosome 8. (B) Targeting of factors to the *TetO* via the TetR DNA-binding domain. Numbers represent qPCR primer positions (kb) with respect to TetO array. (*Adapted from Blackledge* et al. *2014*)

Another important advantage of the TetR system, is that the interaction between TetR and *TetO* is disrupted upon doxycycline treatment. This serves as an important control to demonstrate the dependence of a recruited protein on the presence of the TetR-fusion. Additionally, it can be used to understand the binding dynamics of recruited proteins following loss of the original effector protein. A cell line expressing TetR alone is used as a control to ensure detected proteins are not recruited by TetR itself.

As ChIP qPCR is used as a detection tool, results are interpreted in a binary manner, where an amplified signal relative to control (TetR alone or in the presence of doxycycline) is used to indicate the recruitment of a protein by the TetR-fusion. This

signal for a particular protein at the TetO array may be evaluated in comparison to its relative enrichment at endogenous loci. However, as ChIP qPCR is an antibody dependent technique, one limitation of this system is that it does not allow quantitative comparisons of enrichment between different proteins. Furthermore, it is worth bearing in mind that an absence of enrichment for a specific protein by ChIP qPCR may not necessarily mean it is not recruited by the TetR-fusion, but due to technical limitations (antibody affinity or crosslinking), its presence may be below the limits of detection for this assay. Similarly, interactions between TetR and its fusion partner, and/or misfolded domains may occlude legitimate interactions with other proteins, resulting in false-negative results.

## 4.2 Aims

In this chapter, the mESC TetR system was used to establish the following aims:

1. To determine if Myb is sufficient for the recruitment of its known interaction partners
2. To determine if Myb is sufficient for the recruitment of transcription activation complexes
3. To determine if Myb is sufficient for establishing an enhancer-like element at chromatin

## 4.3 Results

### 4.3.1 Myb binding is sufficient for the recruitment of functional CBP and P300

Firstly, full length mouse Myb was fused to the bacterial TetR DNA binding domain to generate a TetR-Myb fusion protein that was stably expressed in mESCs containing a *TetO* array (from now on referred to as *TetO*) (Figure 4.3A & B). As human MYB has 90.4% amino acid identity compared with mouse Myb (UniProt Consortium 2019), and all proteins assayed for in this system have >90% amino acid conservation between both species (see Table 4.3), interactions observed in this system are assumed to be applicable to both organisms.

**Table 4.3:** Conservation of human MYB vs mouse Myb interacting proteins.

| Protein | % amino acid identity |
|---|---|
| MYB | 90.4 |
| CBP | 95.3 |
| P300 | 93.7 |
| BRD4 | 93.1 |
| MED1 | 94.3 |
| TBP | 91.2 |
| TAF12 | 96.3 |

% amino acid identity of human MYB vs mouse Myb, MYB interacting proteins and transcription activation complex proteins (UniProt Consortium 2019)

To confirm DNA binding at *TetO* by the TetR-Myb fusion protein, ChIP qPCR for TetR (using an antibody recognising an N-terminal FS2 tag) and Myb was performed in the TetR-Myb mESC line. Signal was detected at *TetO* by both TetR and Myb ChIP qPCR in the absence of doxycycline, but signal dropped to near background 0.5 kb away from *TetO*, demonstrating sequence-specific binding (Figures 4.3D & F). After 6 hours of doxycycline treatment, which disrupts the affinity of the TetR DB domain for the TetO sequence, TetR-Myb binding at this locus was abolished (Figures 4.3D & F), thus validating the TetR-Myb system. In this and all subsequent experiments, a cell expressing TetR alone was used

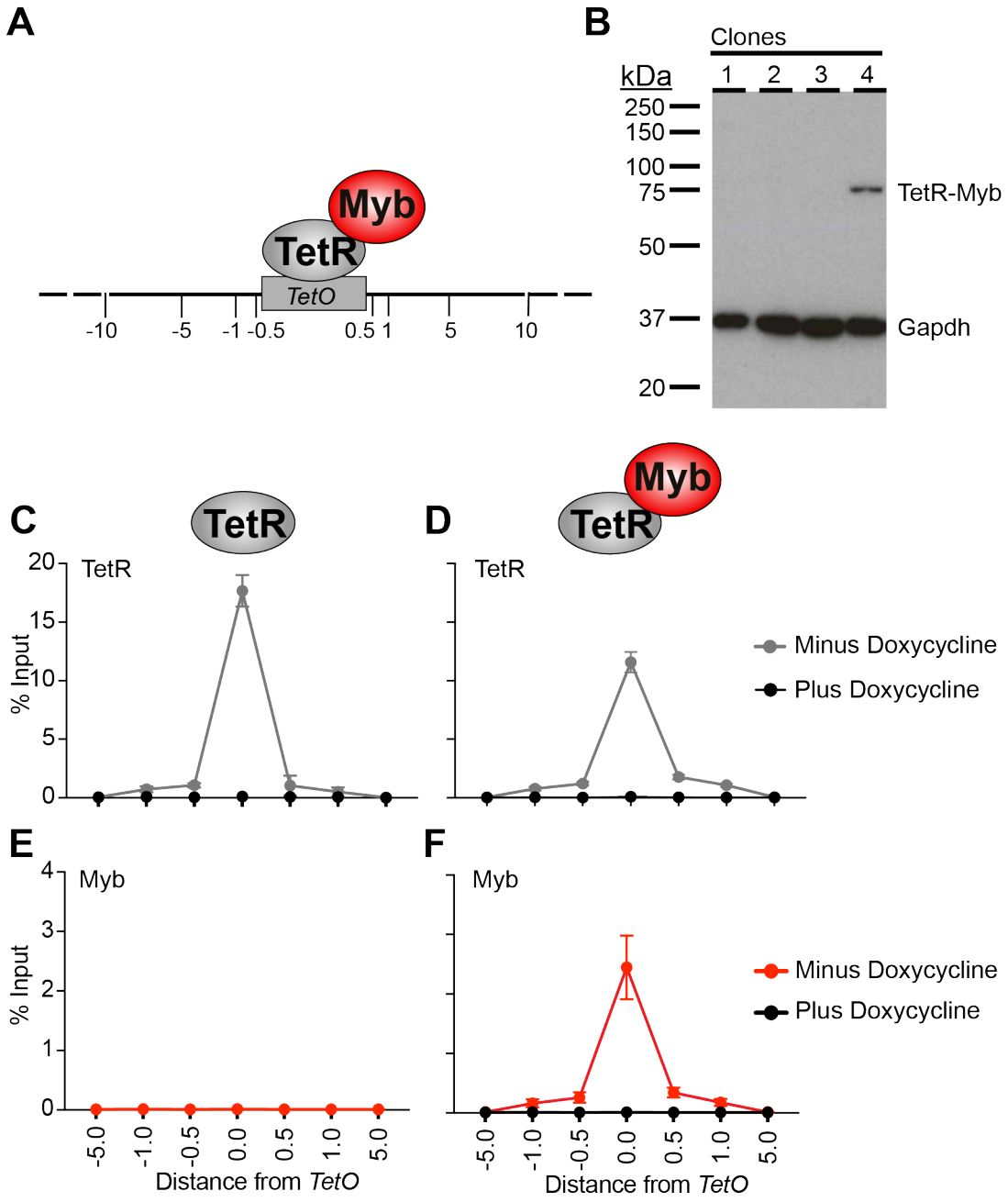



**Figure 4.3: Validation of the TetR-Myb system.** (A) Schematic of TetR-Myb bound at *TetO*. Numbers represent qPCR primer positions (kb) with respect to *TetO*. B) Western blot analysis using antibody against TetR (N-terminal FS2 tag), demonstrating the expression of TetR-Myb in a clonal cell line (lane 4) and negative controls (lanes 1-3). ChIP qPCR in cell lines stably expressing either TetR or TetR-Myb, using antibodies against (C-D) TetR (N-terminal FS2 tag) (grey lines) and (E-F) Myb (red lines) in the absence of doxycycline, and following 6 hours of doxycycline treatment (black lines). Y-axis indicates primer pairs used, based on their distance from *TetO*. Error bars represent the standard deviation from three biological replicates.

as an important negative control. In TetR expressing cells, TetR but not Myb is detected at *TetO* (Figures 4.3C & E).

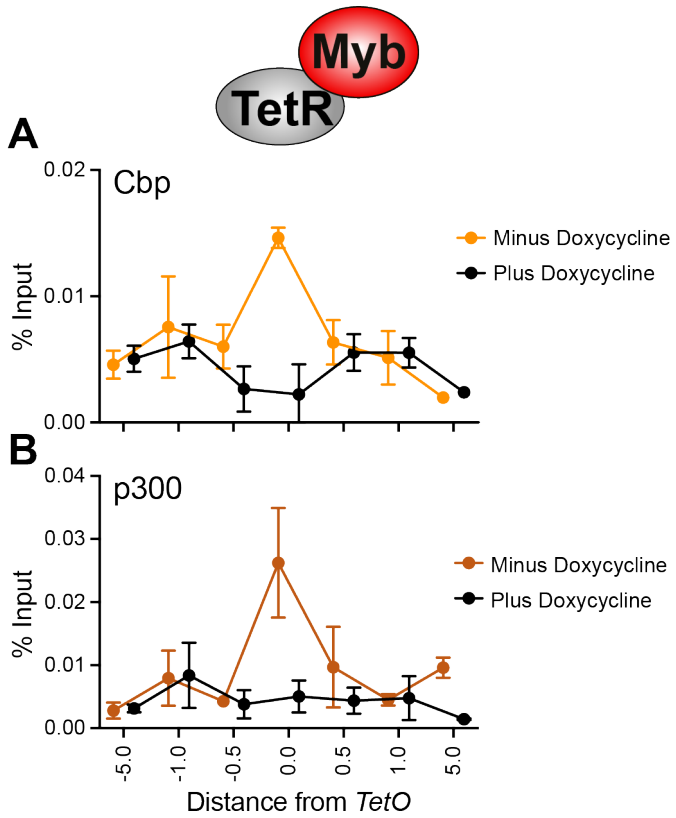



**Figure 4.4: Cbp and p300 recruitment is dependent on TetR-Myb.** ChIP qPCR in cell lines stably expressing either TetR or TetR-Myb, using antibodies against (A) Cbp (yellow line) and (B) p300 (brown line) in the absence of doxycycline, and following 6 hours of doxycycline treatment (black lines). Error bars represent the standard deviation from three biological replicates.

Next, to test if TetR-Myb was sufficient for stable binding of Cbp and p300, with which it is known to interact, ChIP qPCR was performed for the two proteins. Both proteins were enriched for at *TetO*, dependent on the binding of TetR-Myb (Figures 4.4A & B). In addition, H3K27ac was also detected at *TetO*, indicating that Cbp/p300 was functional in this context (Figure 4.5B). Consistent with a direct interaction with TetR-Myb, both Cbp/p300 and H3K27ac were localised at *TetO* and enrichment levels dropped rapidly away from this locus.

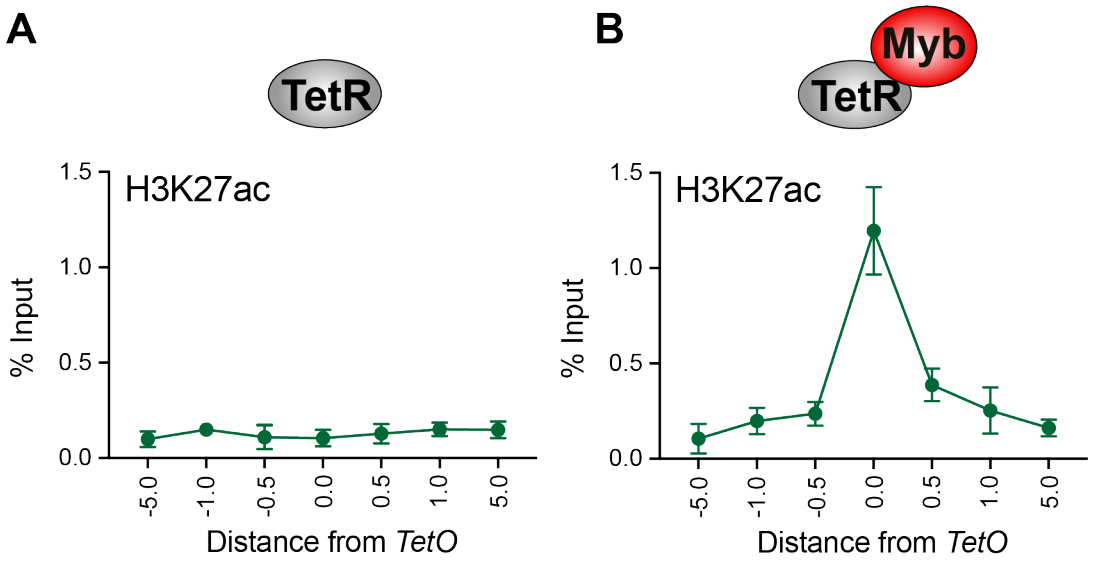



**Figure 4.5: H3K27ac is dependent on TetR-Myb.** ChIP qPCR in cell lines stably expressing either (A) TetR or (B) TetR-Myb, using antibodies against H3K27ac (green lines) in the absence of doxycycline. Error bars represent the standard deviation from three biological replicates.

### 4.3.2 The Myb transactivation domain is sufficient for stable binding of its known interaction partners

As mentioned above (Chapter 4.1.1), the interaction between the Myb TA domain with Cbp and p300 is well documented (Oelgeschläger *et al.* 1996; Kiewitz *et al.* 1997; Kasper *et al.* 2002; Sandberg *et al.* 2005). Therefore, to test if the Myb TA domain in isolation is capable of recruiting its interaction partners *in vivo*, amino acids 241-325 of mouse Myb was fused to the C-terminus of TetR, to produce a TetR-Myb TA domain fusion protein (TetR-TA). Several stable TetR-TA cell lines were generated (Figure 4.6B). As before, these were validated using ChIP qPCR for TetR in the presence and absence of doxycycline. Data from a representative TetR-TA clone is shown in Figure 4.6D, and this cell line was used in all downstream experiments.

Perhaps unsurprisingly, given the results observed with TetR-Myb, both Cbp and p300 were also detectable at *TetO* by ChIP qPCR in TetR-TA cells. However, relative to recruitment via full length Myb, both Cbp and p300 were enriched at *TetO* by at least one order of magnitude greater than at endogenous loci (Figure 4.7A & B). While this technique is not directly quantitative, this nonetheless represents

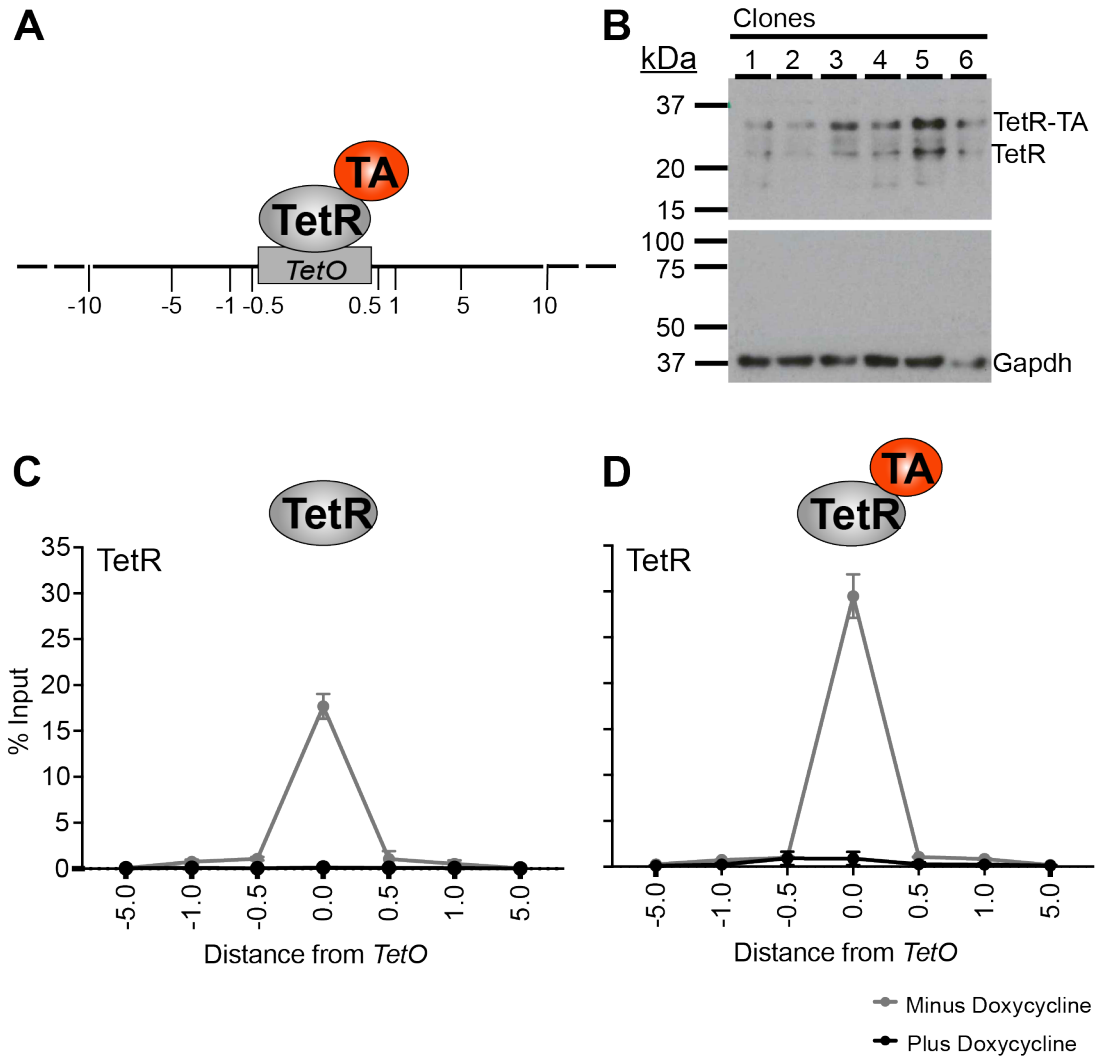



**Figure 4.6: Validation of the TetR-TA system.** (A) Schematic of TetR-TA bound at *TetO*. Numbers represent qPCR primer positions (kb) with respect to the *TetO* locus. (B) Western blot analysis using antibodies against TetR (N-terminal FS2 tag), demonstrating the expression of TetR-TA in a clonal cell lines (lanes 1-6). ChIP qPCR in cell lines stably expressing either (C) TetR or (D) TetR-TA, using antibodies against TetR in the absence of doxycycline (grey lines), and following 6 hours of doxycycline treatment (black lines). Error bars represent the standard deviation from three biological replicates.

a massive relative increase in the amount of bound protein detected. Furthermore, detection of H3K27ac by ChIP qPCR was noticeably increased in parallel with Cbp and p300. H3K27ac was detected not only at *TetO*, as seen in the TetR-Myb line, but was further enriched at the regions spanning the *TetO* locus, peaking at +/- 1kb but spreading up to +/-10kb away (Figure 4.7C). This pattern was also observed for H3K9ac, another histone residue which can be acetylated by Cbp and p300

(Figure 4.7D). This dip in signal at the TetO has been observed with other histone modifications e.g. H3K27me3 and H2AK119ub (Blackledge *et al.* 2014), and may be explained by relative nucleosome depletion at the *TetO* due to the high frequency of TetR-TA binding. It is not clear why this was not observed for the TetR-Myb fusion but may attributed to the decreased activity of full length Myb relative to the TA domain in isolation. Taken together, the differences observed between TetR-Myb and TetR-TA binding at the TetO, could be explained by the absence of both the DB and NR domains of full length Myb in the latter. As described above (Chapter 4.1.1), these domains have been shown to mediate both intramolecular and intermolecular interactions that result in decreasing the transactivation activity of Myb, and therefore potentially reducing the binding affinity for Cbp/p300.

Since we have demonstrated that the Myb TA domain alone is sufficient to recruit CBP/P300 and catalyse H3K27ac, we wanted to extend this result and determine if it is sufficient to recruit additional proteins involved in either promoter or enhancer function. Interestingly, recent work in human cell lines has suggested that MYB interacts with TAF12 (Xu *et al.* 2018), which is part of the TFIID general TF complex that includes TBP makes up part of the pre-initiation complex of RNAPII. Hence, to determine if the Myb TA domain is sufficient to recruit the pre-initiation complex, a canonical member of which is Tbp, ChIP qPCR was performed for Tbp in the TetR-TA cell line. Enrichment of Tbp was observed at high levels at *TetO* and lower levels were present up to 10kb on either side of this locus (Figure 4.7E). Importantly, the above observations demonstrate that, in the absence of its negative regulatory domains, the Myb TA domain is sufficient for stable chromatin association of its interaction partners.

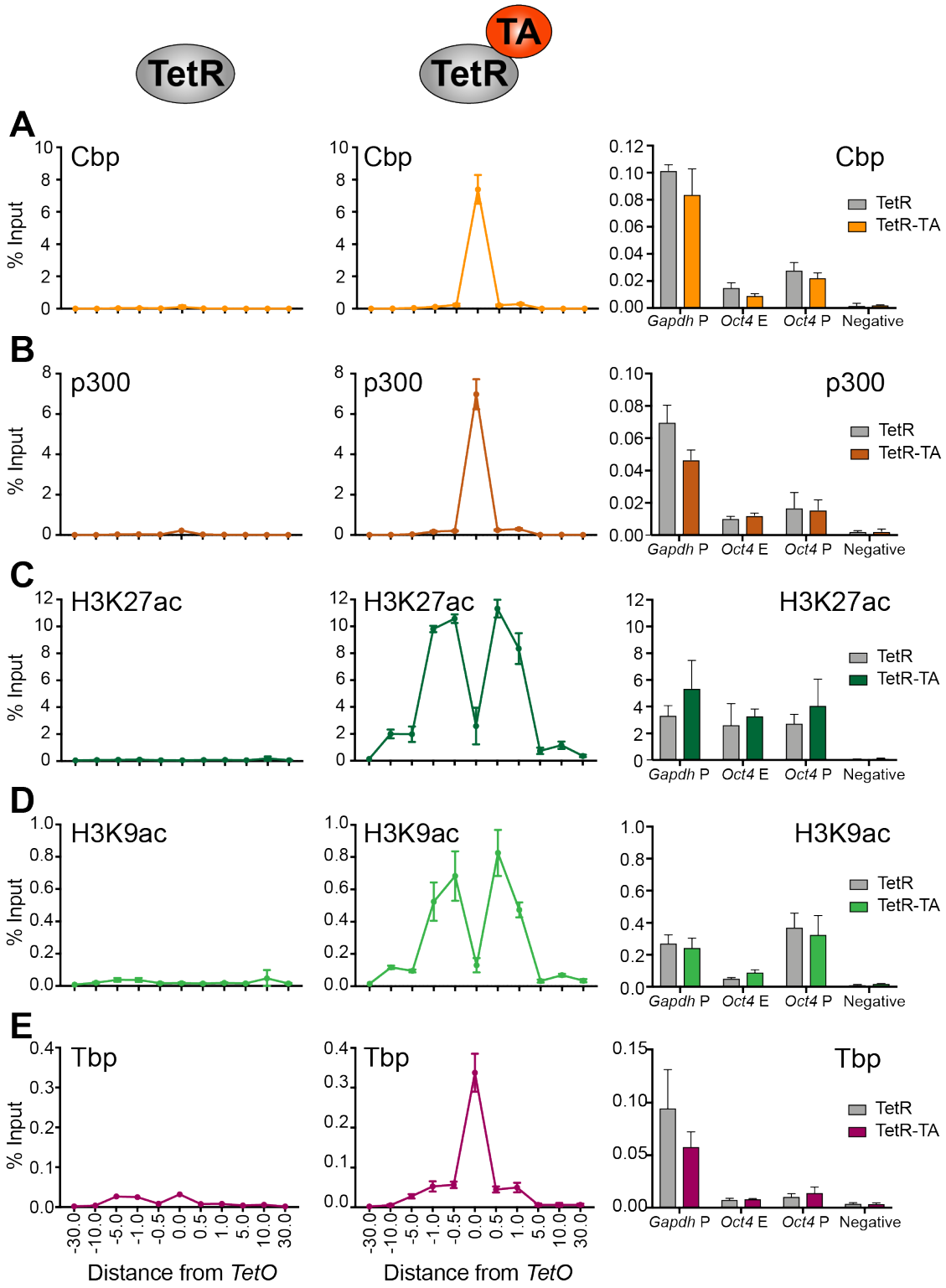



**Figure 4.7: The Myb transactivation domain is sufficient for stable binding of its known interaction partners.** Panels contain ChIP qPCR graphs of cell lines stably expressing either TetR (left column) or TetR-TA (middle column), using antibodies against (A) Cbp (yellow), (B) p300 (brown), (C) H3K27ac (dark green), (D) H3K9ac (light green) (E) and Tbp (mauve). qPCR for each ChIP at endogenous loci (right column) is shown for comparison. Error bars represent the standard deviation from three biological replicates.

### 4.3.3 The Myb transactivation domain is sufficient for the recruitment of transcription activation complexes and RNAPII

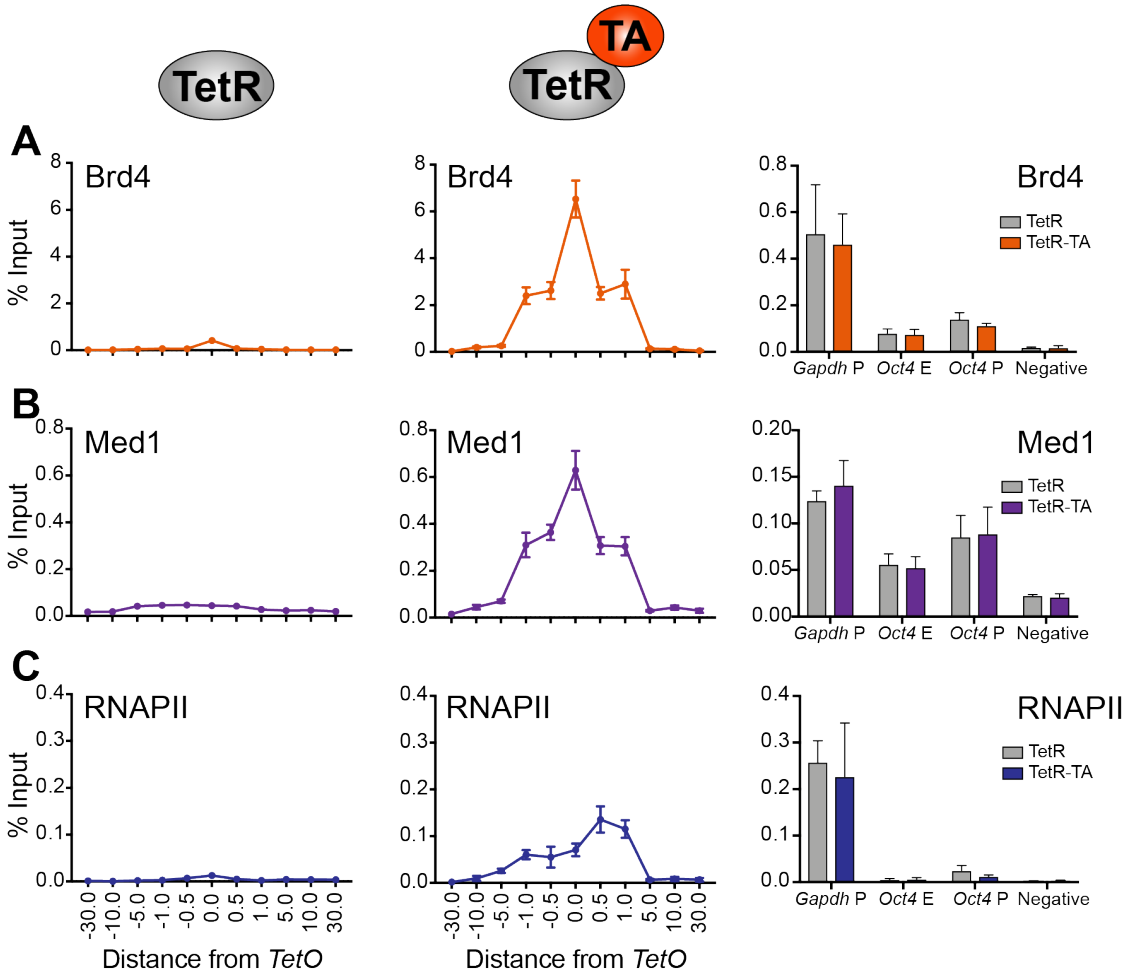



**Figure 4.8: The Myb transactivation domain is sufficient for the recruitment of Brd4, Med1 and RNAPII.** Panels contain ChIP qPCR graphs of cell lines stably expressing either TetR (left column) or TetR-TA (middle column), using antibodies against (A) Brd4 (orange), (B) Med1 (purple) and (C) RNAPII subunit RPB1 (blue). qPCR for each ChIP at endogenous loci (right column) is shown for comparison. Error bars represent the standard deviation from three biological replicates.

Building on the results from Chapter 4.3.2, we wanted to test if other chromatin proteins and co-activators not known to directly interact with Myb or its TA domain were also present. Specifically given that H3K27ac, H3K9ac and Tbp were all detectable at *TetO* via TetR-TA, it would be interesting to test for the presence of Brd4 (known to recognise and bind to acetylated lysines), Mediator

(known to interact with Tbp, Brd4 and RNAPII), and RNAPII (recruited by the pre-initiation complex, a member of which is Tbp). Furthermore, both Brd4 and Mediator are known components of phase separated activation complexes thought to mediate enhancer-promoter interactions.

Significantly, ChIP qPCR for both Brd4 and Med1 demonstrated enrichment at, and up to +/-5kb from *TetO* (Figure 4.8A & B). Even more striking, was the detection of RNAPII spreading up to +/-10kb from *TetO* (Figure 4.8C). Surprisingly, while the other proteins and histone modifications analysed were symmetrically distributed about the TetO, RNAPII showed a bias towards binding on the plus-side of the *TetO*. Collectively, these ChIP qPCR results suggest that the Myb TA domain alone was sufficient for stabilisation of transcription activation complexes, likely via its interacting proteins and deposited histone modifications.

### 4.3.4 Myb is sufficient for establishing an enhancer-like element at chromatin

As mentioned in Chapter 1.1.3, enhancer features include 1) TF binding, 2) the presence of activating complexes, 3) histone modifications such as H3K27ac and H3K4me1, 4) chromatin accessibility, 5) enhancer-promoter interactions, and 6) the production of eRNAs.

So far, the data presented here have demonstrated that the Myb TA domain alone is sufficient to recruit activating complexes and RNAPII to a chromatinised locus. To further explore this, we wanted to determine if the presence of the Myb TA domain alone had any impact on the surrounding chromatin landscape or on the initiation of transcription from the *TetO* locus. Thus, Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) and poly(A) minus RNA sequencing (poly(A) minus RNA-seq); to enrich for possible eRNA production) were performed to test if anchoring the Myb TA domain at the *TetO* locus would result in increased chromatin accessibility and active transcription. These techniques were conducted in a cell line stably expressing TetR-TA, and as controls, a cell line expressing TetR only and cells with no TetR or TetR-fusion protein (NP).

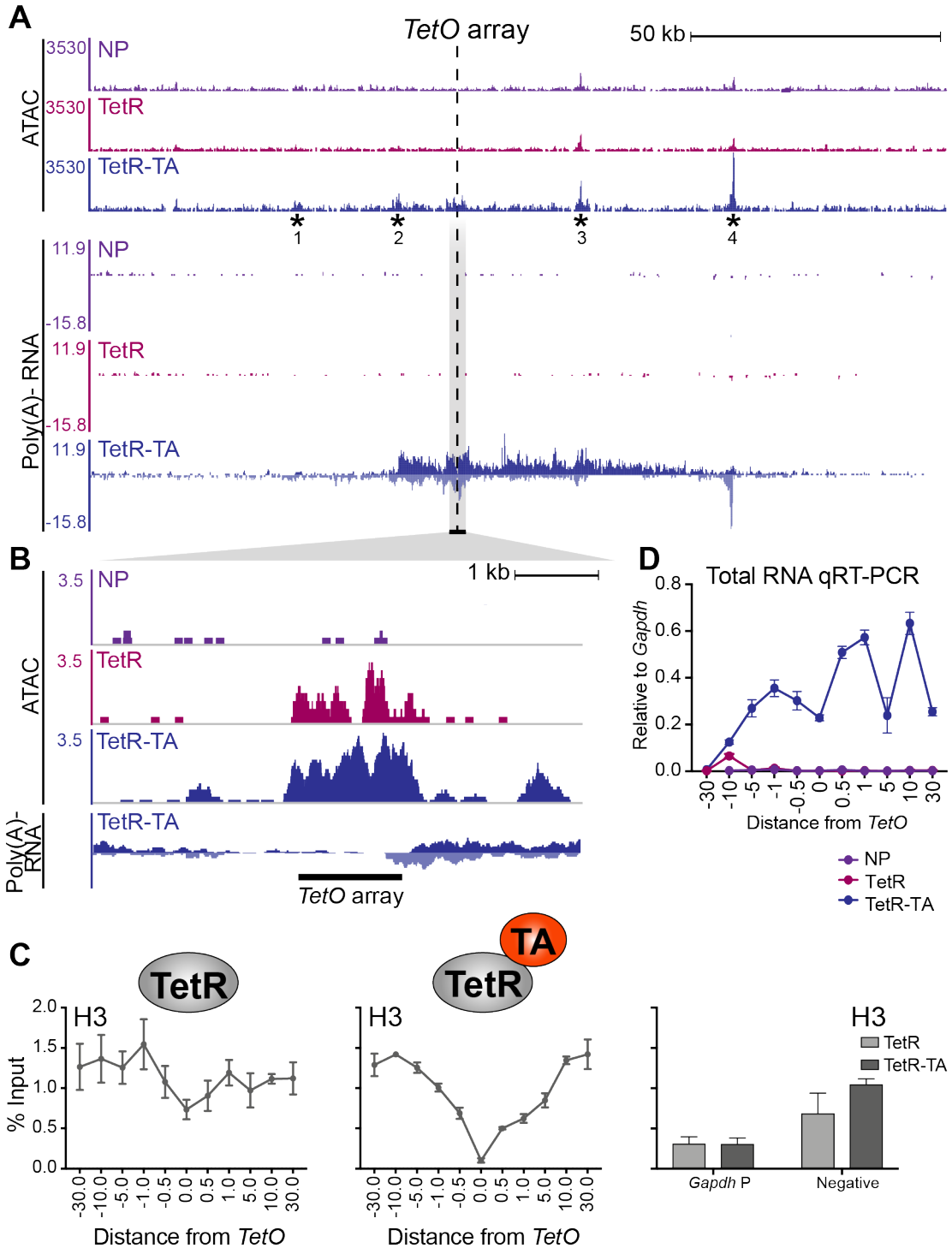



**Figure 4.9: Myb binding results in increased chromatin accessibility and active transcription.** (A) Representative tracks of ATAC-seq and strand specific poly(A) minus RNA-seq at the *TetO* locus in NP, TetR and TetR-TA cells. Regions of increased chromatin accessibility are marked with * and numbered 1-4. (B) Windowed view of ATAC-seq in all three cell lines and poly(A) minus RNA-seq in TetR-TA cells at the *TetO* locus.

Unsurprisingly, in NP cells, there was a lack of ATAC signal at the *TetO* locus given that there was no protein available to bind to the array. In contrast, chromatin accessibility was observed at *TetO* in both TetR and TetR-TA cells (Figure 4.9B), indicating that TetR binding was sufficient to generate accessible chromatin. In the TetR-TA line, however, the ATAC-seq signal at *TetO* was significantly increased, indicating greater accessibility. Complementary to these observations, ChIP qPCR demonstrated decreased histone H3 signal at *TetO* in TetR-TA compared with TetR cells (Figure 4.9C), supporting the idea that the region is less chromatinised (and therefore more accessible) in TetR-TA cells.

In TetR cells, no other differences in accessibility compared to NP cells were observed within ∼100kb of the *TetO* locus. Strikingly however, in TetR-TA cells, in addition to the changes at *TetO* (Figure 4.9B & C), accessibility was also increased at several distal sites (marked with *) (Figure 4.9A). Interestingly, some of these were already be accessible in both TetR and NP cells, however, the accessibility was markedly increased in TetR-TA cells (*3 and *4). Perhaps more significantly, there were other accessible sites distal to *TetO* that appeared to be unique to TetR-TA cells (*1 and *2).

Remarkably, poly(A) minus RNA-seq revealed the presence of active transcription centred around the *TetO* locus in TetR-TA but not TetR or NP cells (Figure 4.9A, B & D). Transcription from *TetO* appeared to be bidirectional, but contrary to canonical eRNA transcription, was observed to persist for more than 90kb upstream and 40kb downstream of *TetO* (Figure 4.9A, B and D in part). Additionally, there was evidence of transcript accumulation at certain loci both up and downstream of the *TetO* locus, with many of these RNA peaks also corresponding to the increased ATAC peaks described above (Figure 4.9A), suggesting that these may be sites of transcription initiation. At non-ATAC peaks, transcript accumulation is reminiscent

---

**Figure 4.9:** continued. (C) Panel containing ChIP qPCR graphs of cell lines stably expressing either TetR or TetR-TA, using antibodies against H3 (dark grey). Error bars represent the standard deviation from three biological replicates. (D) Expression analysis by qRT-PCR at the *TetO* locus in all three cell lines.

of the pattern of mRNA splicing observed at coding genes, however read analysis from all TetR-TA samples revealed only sporadic splicing events, with no enrichment for any particular "isoform" or splicing preference that would explain the transcript pattern observed (data not shown). Other explanations for this phenomenon would include increased RNAPII initiation at these regions or RNAPII pausing immediately downstream of these loci and should be investigated further. With this in mind, there does appear to be two possible separate transcription start sites (TSSs) approximately 10kb downstream and 55kb upstream of *TetO* coinciding with the amplified ATAC peaks *2 and *4, and demonstrating unidirectional transcription on the plus and minus strand respectively.

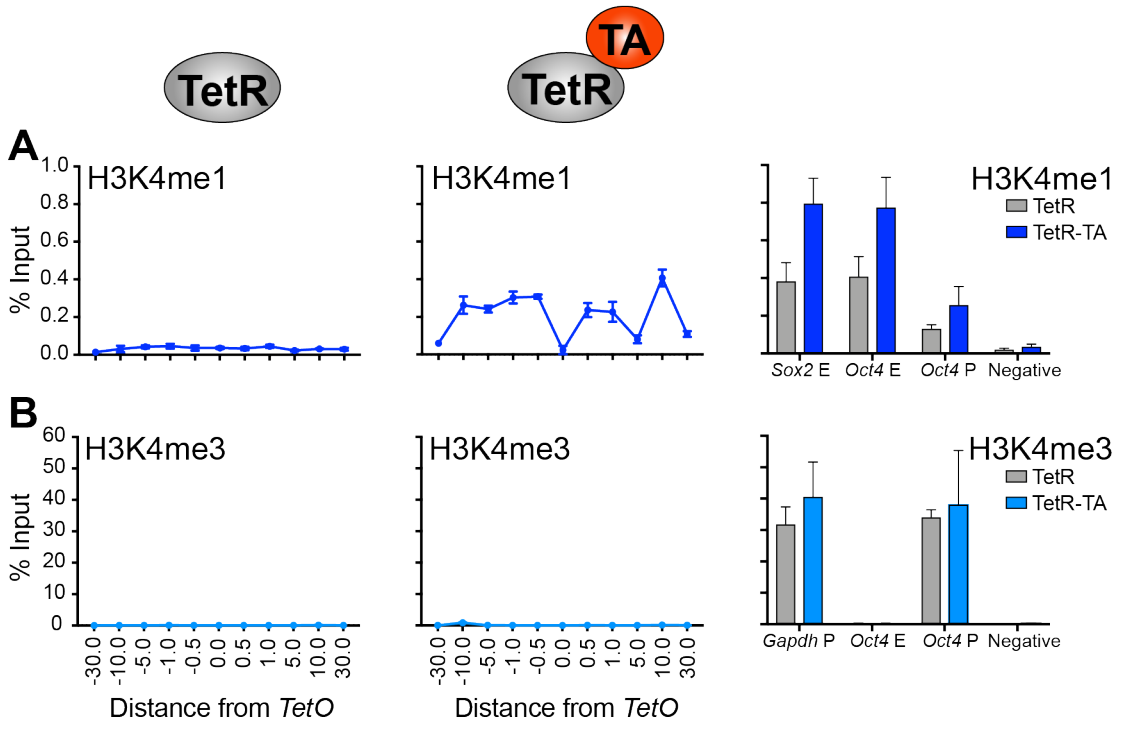



**Figure 4.10: The Myb transactivation domain is sufficient for the recruitment of H3K4me1 but not H3K4me3.** Panels contain ChIP qPCR graphs of cell lines stably expressing either TetR (left column) or TetR-TA (middle column), using antibody against (A) H3K4me1 (light blue) and (B) H3K4me3 (medium blue). qPCR for each ChIP at endogenous loci (right column) is shown for comparison. Error bars represent the standard deviation from three biological replicates.

Although it has not been specifically verified that TetR-TA does not also bind at these regions of increased chromatin accessibility and transcription, this

seems unlikely given that there are no endogenous *TetO* sequences in these regions. Excluding this possibility, the data therefore suggest that the Myb TA domain alone is capable of 1) initiating and enhancing chromatin remodelling, and 2) recruiting and stabilising active RNAPII, both locally and at distal loci.

Thus far, anchoring TetR-TA to the *TetO* locus has resulted in chromatin characteristics that are common to both enhancers and promoters. Therefore, ChIP qPCR for histone modifications H3K4me1 and H3K4me3 which selectively mark enhancers and promoters respectively was performed. In addition to recruitment of H3K27ac, Figure 4.10 demonstrates enrichment of H3K4me1 (Figure 4.10A) relative to H3K4me3 (Figure 4.10B) by ChIP qPCR in the region of *TetO*. Taken together, the data strongly suggest that simply anchoring the Myb TA domain to DNA, is sufficient to establish an enhancer-like element.

### 4.3.5 A single TetR binding site is sufficient for Myb dependent enhancer-like activity

One possible criticism of the findings at the *TetO* locus, is that the array consists of 14 TetO sequences. Therefore, the observations made may be unique to the engineered locus due to artificially inflated levels of bound TetR-TA.

Fortuitously, Blackledge *et al.* 2014 also identified a single locus on mouse chromosome 1 with sequence homology to the bacterial TetO consensus sequence (Figure 4.12A). This endogenous *TetO* (*ETetO*) is distant from surrounding genes and therefore provides a "natural" site at which to analyse the effects of TetR-TA binding.

TetR-TA binding to *ETetO* is able to recapitulate (albeit at reduced levels) Cbp, H3K27ac, H3K4me1, Brd4 and RNAPII recruitment (Figure 4.11B-G), which is not observed with TetR only binding. Importantly, as at *TetO*, the presence of TetR-TA at *ETetO*, is associated with increased chromatin accessibility (Figure 4.12). Consistent with the results at *TetO*, accessibility was increased both at *ETetO* itself and at a second peak ∼500bp away. Additionally, active bidirectional transcription was observed, originating from the *ETetO* locus in the TetR-TA but

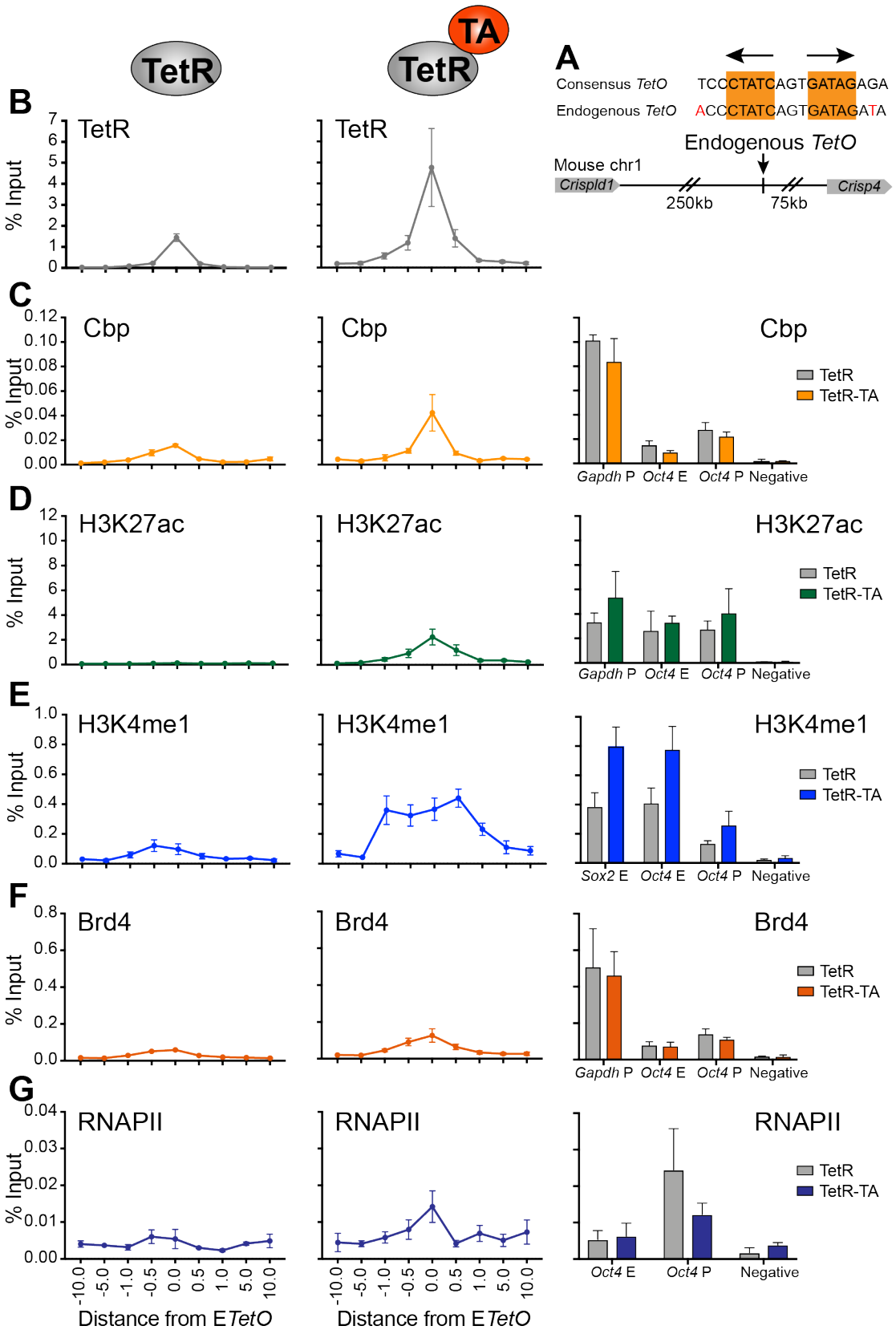



**Figure 4.11: Myb-TA recruitment of Cbp, H3K27ac, H3K4me1, Brd4 and RNAPII at a single endogenous TetO site in mouse ESCs.**

not TetR or NP cell lines (Figure 4.12), although at a reduced frequency relative to that observed at the *TetO* locus (Figure 4.9). This observation is somewhat unsurprising given that there are 14 TetO sequences at *TetO* compared with a single TetO sequence at *ETetO*.

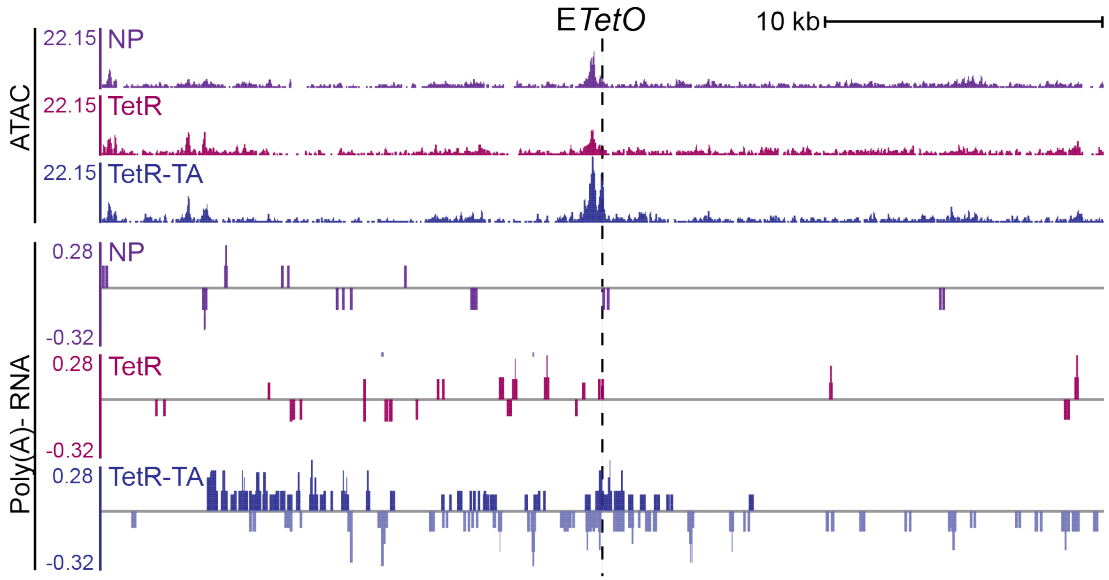



**Figure 4.12: Myb-TA binding at a single endogenous *TetO* site results in increased chromatin accessibility and active transcription.** Representative tracks of ATAC-seq and strand specific poly(A) minus RNA-seq at the *ETetO* in NP, TetR and TetR-TA cells.

Together, these findings demonstrate that Myb dependent enhancer-like activity can also be targeted to a single TetR binding site endogenous to the mouse genome. This argues that the effects observed in these experiments are not the result of copy number or site specific features inherent to the engineered *TetO* locus. Therefore,

---

**Figure 4.11:** continued. (A) A region on mouse chromosome 1, distant from surrounding genes, contains a sequence with homology to a single consensus bacterial TetO binding site. This provides a endogenous (*ETetO*) site in the genome at which to analyse TetR fusion protein binding and function. Panels contain ChIP qPCR graphs of cell lines stably expressing either TetR (left column) or TetR-TA (middle column), using antibodies against (B) TetR (N-terminal FS2 tag) (grey), (C) Cbp (yellow), (D) H3K27ac (dark green), (E) H3K4me1 (light blue), (F) Brd4 (orange) and (G) RNAPII subunit RPBI (dark blue). qPCR for each ChIP at endogenous loci (right column) is shown for comparison. Error bars represent the standard deviation from three biological replicates.

*de novo* recruitment of Myb is sufficient for the formation of an enhancer-like regulatory element.

## 4.4 Discussion

These results establish that Myb is sufficient to recruit Cbp/p300 and catalyse H3K27ac (Figure 4.4). Furthermore, we have uniquely demonstrated that this activity is mediated by the Myb TA domain, and more remarkably, that this domain alone is sufficient to establish a chromatin environment with characteristics of an enhancer-like element. This is evidenced by the following features: (1) co-activator occupancy (Cbp, p300, Tbp, Brd4, Med1) (Figures 4.7, 4.8 & 4.11), (2) an active enhancer chromatin signature (H3K27ac, H3K4me1) (Figure 4.7, 4.10 & 4.11), (3) increased chromatin accessibility both locally and distally (Figures 4.9 & 4.12), and (4) active transcription (Figures 4.9 & 4.12), which are all dependent on the presence of the Myb TA domain.

Several notable observations have arisen from these data. Firstly, an important finding is that even a single TetR-TA molecule anchored at *ETetO* is sufficient to generate the above mentioned enhancer characteristics (Figure 4.11 & 4.12), and where an increased concentration of TetR-TA is present, i.e. at *TetO*, exponentially enriched levels of co-activating proteins Cbp, p300, Brd4, as well as H3K27ac are observed, reminiscent of those found at super-enhancers (Sabari *et al.* 2018; Whyte *et al.* 2013; Hnisz *et al.* 2013) (Figure 4.7, 4.8). TetR-TA mediated recruitment of these co-activating complexes and activating histone modifications is associated with chromatin remodelling both locally, and, as in the case at *TetO*, also distally (Figure 4.9 & 4.12).

Interestingly, the R3 repeat of the Myb DB domain has been shown to be important in mediating its potential pioneer factor function, where the D152V mutation in R3 leads to decreased binding to closed chromatin due to a weakened histone interaction, resulting in reduced chromatin accessibility at these loci by ATAC-seq (Fuglerud *et al.* 2017). Mice harbouring the D152V mutation behave like $Myb^{kd/kd}$ mice and develop thrombocytosis, megakaryocytosis, anaemia and

lymphopenia (Emambokus *et al.* 2003; Carpinelli *et al.* 2004). Notably, this is a similar phenotype displayed by mice with Myb TA domain mutations such as L302A, M303V and E308G which disrupt its interaction with the KIX domains of CBP/p300 (Zor *et al.* 2004; Sandberg *et al.* 2005; Pattabiraman *et al.* 2009; Papathanasiou *et al.* 2010; Pattabiraman *et al.* 2014). Taking the above into consideration, and given that our data demonstrate that the Myb TA domain alone is capable of effecting increased chromatin accessibility, this suggests that Myb dependent chromatin remodelling is mediated by its TA domain activity and interactions with co-activators. This provides a molecular mechanism to explain Myb-mediated gene activation, where Myb is directed to gene regulatory elements by its DB domain, and appears to function by establishing chromatin accessibility and other aspects of enhancer (or promoter) activity via its TA domain.

Another significant observation is the pattern of transcription arising from the *TetO* locus, which persists for tens of kilobases both up and downstream, and shows evidence of transcript accumulation at distinct sites (Figure 4.9). Consequently, although this locus bears the hallmarks most characteristic of an active enhancer, the transcripts that have arisen are distinct from canonical eRNA transcripts which have typically been described as being short, abortive transcripts less than 2kb in length (see Chapter 1.1.3.2) (Figure 4.9). Interestingly, more recently, it has been suggested that there may be a degree of overlap between eRNAs and lncRNAs, where some lncRNAs may in fact be stably transcribed eRNAs with no significant biochemical function beyond their role in enhancer activity (Paralkar *et al.* 2016). However, the stability of these long eRNAs is thought have arisen through well described mechanisms such as the presence of splicing acceptor and donor sites and/or polyadenylation cassettes (Ørom *et al.* 2010). As we have not observed regular splicing events involving poly(A) minus transcripts arising from the *TetO* locus, this is unlikely to be the case here. Taking into account the accumulation of transcripts at several loci (Figure 4.9), an alternative explanation would include increased RNAPII initiation at these regions or RNAPII pausing immediately downstream of these sites and could be investigated by comparing the enrichment

of RNAPII serine 2 and serine 5 phosphorylation by ChIP qPCR or ChIP-seq. As many of these transcription peaks also coincide with novel or enhanced ATAC peaks, and particularly given the appearance of a possible separate TSS 55kb upstream at *4 (Figure 4.9), it would be interesting to interrogate these loci for the presence of H3K27ac and/or H3K4me1 or H3K4me3, to determine if the activity of TetR-TA at *TetO* has somehow uncovered cryptic promoters in this region.

Using the mESC TetR system, our observations of Myb directed activity have so far been limited to loci containing the TetO sequence. Building on these observations, it would be worthwhile to study its effect at endogenous gene regulatory elements. A useful approach would be to fuse the Myb TA domain or full length Myb to de-activated Cas9 nuclease (dCas9), which can be directed by specific guide RNAs (gRNAs) to transcriptionally silent endogenous enhancers. It would be particularly significant if the activity of Myb at endogenous enhancers recapitulates our observations in the TetR system.

Finally, it would be extremely interesting to perform Capture-C from *TetO* to interrogate its interaction domain. Not least because we have observed active transcription in the context of Myb-TA domain dependent recruitment of co-activating complexes and enhancer-associated histone modifications at this locus, both of which have been shown to be important for mediating enhancer-promoter proximity, but particularly if any ATAC peaks distal to the *TetO* locus transpire to possess promoter associated histone modifications. Theoretically, if chromatin looping with *TetO* is observed, this would be exciting evidence that the Myb TA domain is sufficient for establishing enhancer-promoter interactions. Alternatively, bearing in mind that locus specific enhancer-dependent gene activation can occur independent of direct enhancer-promoter proximity (Alexander *et al.* 2019; Benabdallah *et al.* 2019), it may be just as likely that chromatin looping between *TetO* and any proven cryptic promoters may be absent. In this instance, this would imply that Myb mediated transcription does not require enhancer-promoter proximity. In either scenario, these findings would contribute important information to further our understanding of the molecular function of Myb.

# 5

# MYB dependent enhancer function is distinct from the maintenance of enhancer-promoter proximity

## Contents


**5.1 Introduction . . . . . 102**
**5.2 Aims . . . . . 103**
**5.3 Results . . . . . 103**
5.3.1 MYB binding correlates with CBP and P300 at enhancers and promoters . . . . . 103
5.3.2 Identifying a set of MYB sensitive gene targets in MLL-AF9 leukaemia . . . . . 110
5.3.3 MYB dependent gene targets have higher levels of MYB binding and are important for cell proliferation and replication . . . . . 114
5.3.4 H3K27ac at enhancers and promoters is sensitive to MYB knockdown . . . . . 115
5.3.5 Loss of MYB binding results in differential changes to chromatin accessibility . . . . . 117
5.3.6 Loss of MYB binding does not affect TAD structure or enhancer-promoter looping . . . . . 119
5.3.7 MYB is necessary for eRNA transcription at the enhancers of MYB dependent gene targets . . . . . 123
**5.4 Discussion . . . . . 127**


*A large part of the bioinformatic analysis in this chapter was performed in collaboration with Dr Nicholas Crump and Mr Joe Harman (Milne Lab, Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, Oxford).*

## 5.1 Introduction

As discussed in Chapter 1.3.3, MYB is a direct target of MLL-FPs and is overexpressed in MLLr leukaemias, including MLL-AF9 leukaemia. It is thought to be a key mediator of MLL-AF9 oncogene addiction and is required for leukaemia maintenance (Zuber *et al.* 2011), but its molecular function in such circumstances remains unclear.

In the previous chapter (Chapter 4), we established that the recruitment of transcriptional co-activators (Cbp, p300, Brd4, Med1, Tbp), as well as, activating histone modifications (H3K27ac, H3K4me1), is dependent on the presence of the Myb TA domain. To ensure that these are functional interactions and not simply locus specific observations, it would be important to study them in a biological context at native gene regulatory elements e.g. in an MLL-AF9 leukaemia model. For this purpose, the global binding distribution of these proteins may be mapped using ChIP-sequencing, and correlational analyses of their co-occupancy with MYB performed. The power of such an approach is that association patterns between multiple proteins may be observed on a genome-wide scale through the strength of their correlation, and at different genomic loci e.g. enhancers vs promoters. Furthermore, any spurious associations may be falsified by negative correlations.

At the same time, it was also observed in the TetR system that the Myb TA domain alone is sufficient for establishing enhancer-like activity when localised at TetR binding sites (Chapter 4). This raises the question of the reciprocal requirement for Myb in the maintenance of enhancer function, and if this could be the possible mechanism through which loss of Myb expression mediates abrogation of leukaemia phenotypes (Zuber *et al.* 2011). Therefore, by identifying MYB sensitive gene targets in an MLL-AF9 leukaemia model, this chapter further investigates the molecular events leading up to MYB dependent loss of gene expression.

## 5.2 Aims

In this chapter, an MLL-AF9 leukaemia model was used to establish the following aims:

1. To map the genome-wide co-occupancy of MYB with transcription related factors in MLL-AF9 leukaemia
2. To identify a set of MYB sensitive gene targets in MLL-AF9 leukaemia
3. To dissect the implications of loss of MYB on enhancer function

## 5.3 Results

### 5.3.1 MYB binding correlates with CBP and P300 at enhancers and promoters

The following experiments were performed in THP-1 cells as an MLL-AF9 leukaemia model, given the relative abundance of genomic data already held in our lab and in published datasets. Furthermore, THP-1 cells are sensitive to loss of MYB (Zuber *et al.* 2011), providing a well characterised system with which to study its molecular function in greater detail.

In order to understand the role of MYB in a cellular context, it was first necessary to map its genome-wide distribution using ChIP-sequencing (ChIP-seq). CBP, P300, H3K27ac, H3K4me1, H3K4me3 and H3K27me3 ChIP-seq data sets were also generated for the purposes of downstream analysis, together with pre-existing ChIP-seq data sets in our lab (BRD4, H3K79me2, MLL-N, CTCF and EZH2). Example tracks at a well described MYB target gene, *BCL2,* (Zuber *et al.* 2011; Godfrey *et al.* 2016; Tzelepis *et al.* 2018) are shown in Figure 5.1.

Somewhat surprisingly, MYB ChIP-seq in THP-1 cells revealed promiscuous binding, with >61000 discrete peaks called. Of these, 23.6% were found at putative enhancers (E), 17.6% at promoters of expressed genes (P-E), 4.2% at promoters of non-expressed genes (P-NE), and the remaining at un-annotated genomic loci (O) (Figure 5.2A). Enhancers were defined as H3K27ac regions overlapping with

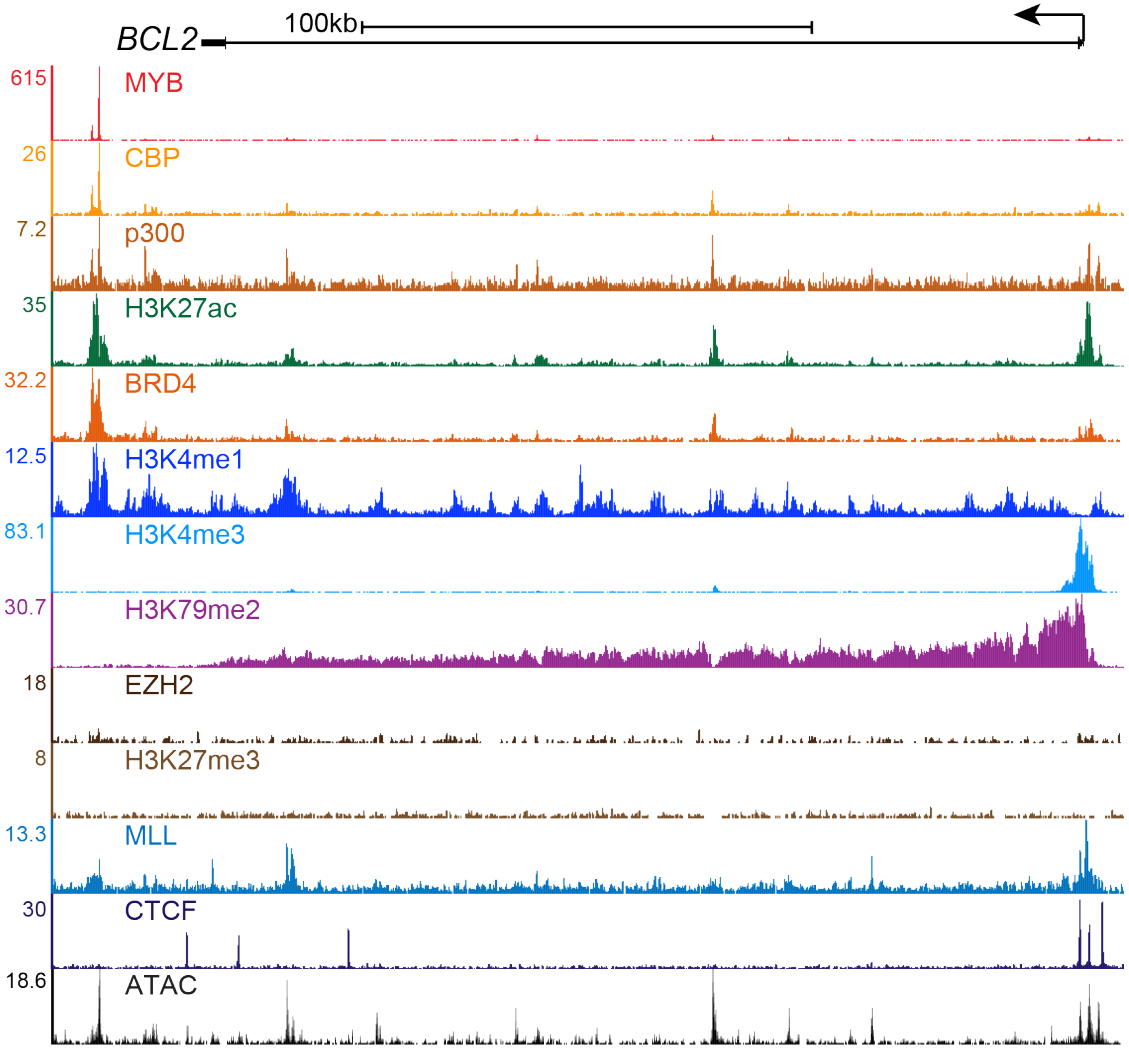



**Figure 5.1: ChIP-seq tracks at the *BCL2* locus.** ChIP-seq tracks at *BCL2* of MYB, CBP, P300, H3K27ac, H3K4me1, H3K4me3, H3K27me3 in THP-1 cells, including ChIP-seq generated in the lab by E.Ballabio, N.Crump, J.Kerry and T.Milne (BRD4, H3K79me2, EZH2, MLL-N and CTCF). Numbers represent the scale shown for each track, normalised to read count per 10 million total reads.

H3K4me1, with regions <1kb apart stitched together, while promoter regions were defined as a 2kb window centred on annotated TSSs (Homer analysis software, Heinz *et al.* 2010), and expressed promoters defined by nascent RNA-seq in THP-1 cells. Overall, at MYB peaks, levels of MYB binding was enriched at enhancers compared to other loci (Figure 5.2B).

Consistent with the fact that MYB physically interacts with CBP and P300 via its TA domain (Chapter 4 & Oelgeschläger *et al.* 1996; Kasper *et al.* 2002; Sandberg *et al.* 2005; Kiewitz *et al.* 1997), MYB binding positively correlated with

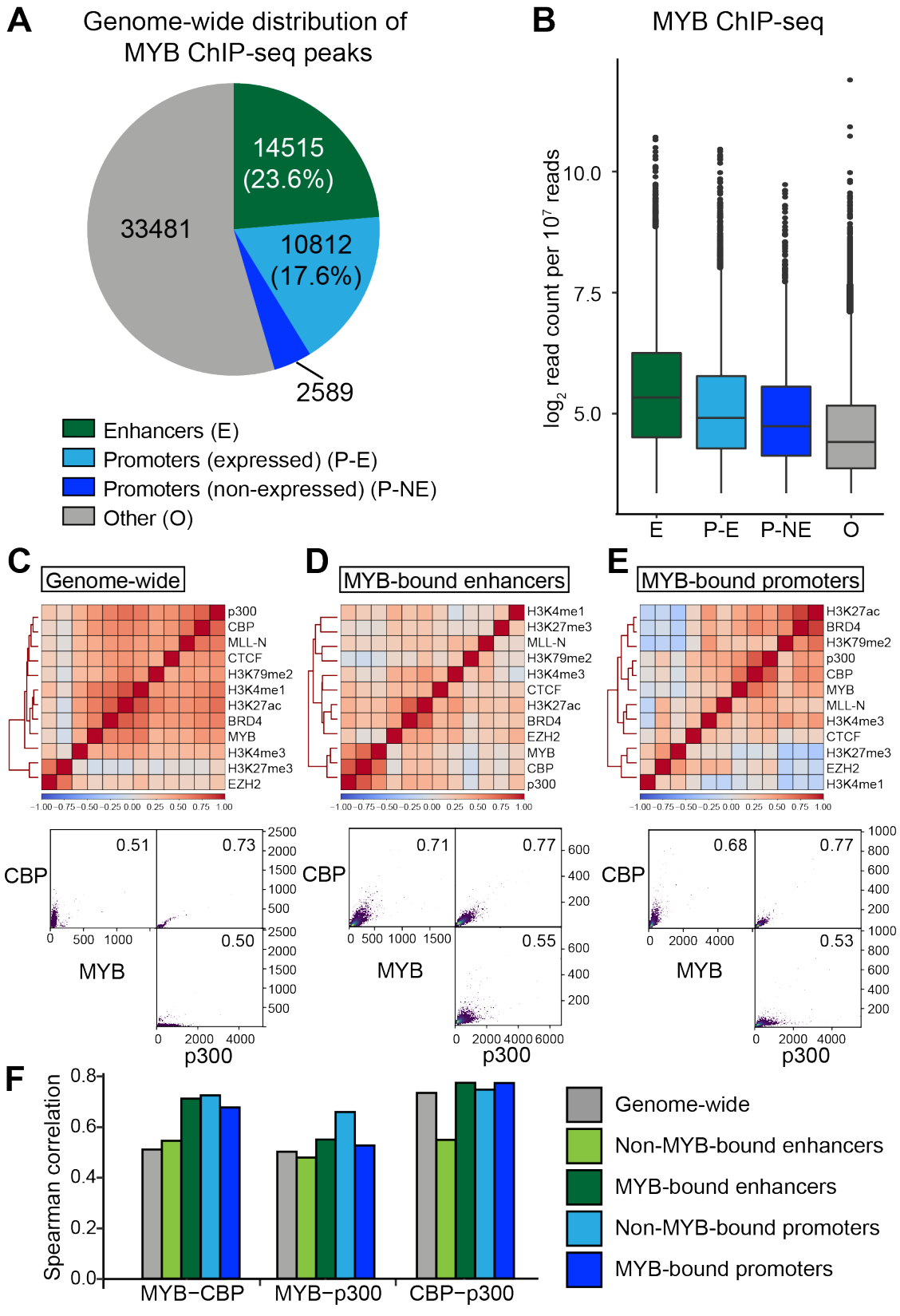



**Figure 5.2: Genome-wide distribution and correlations of MYB ChIP-seq peaks.**

both these proteins. Interestingly, depending on the regions sampled, MYB binding demonstrated two distinct patterns of clustering. At enhancers and promoters, MYB showed consistent clustering and positive correlation coefficients with CBP and P300 (Figures 5.2D-F). In contrast, when sampled genome-wide, although MYB still showed a positive correlation with CBP and P300 (Figures 5.2C & F), it clustered separately from CBP and P300, instead showing a stronger association with BRD4, H3K27ac and H3K4me1 (Figure 5.2C). This suggests that MYB recruitment of CBP and P300 may be more common at gene regulatory elements, suggesting an important role at these loci.

To investigate this relationship further, heatmaps were generated to reveal the distribution of MYB, CBP, P300 and H3K27ac at MYB-bound and non-MYB bound enhancers and promoters (Figure 5.3). Surprisingly, MYB was detectable at some enhancers and promoters categorised as non-MYB-bound, albeit at lower levels. This is likely a consequence of the peak-calling algorithm used, where low-level binding is not identified as a significant peak. However, this emphasises the point that MYB binding is associated with the vast majority of enhancers and promoters.

Consistent with the observed correlation coefficients at both enhancers and promoters, the levels of CBP and P300, and indeed H3K27ac were proportional to the level of MYB binding (Figure 5.3 and 5.4). Importantly, there was little detectable CBP and P300 binding at non-MYB-bound regions, and non-MYB-bound

---

**Figure 5.2:** continued. (A) Genome wide distribution of MYB ChIP-seq peaks, where enhancers (E) are represented in dark green, expressed promoters (P-E) in light blue, non-expressed promoters (P-NE) in royal blue and other unannotated genomic loci (O) in grey. (B) Box plots of $\log_2$ read count per 10 million ChIP-seq reads, where the lower and upper hinges correspond to the 1st and 3rd quartiles respectively and whiskers extend to the smallest and largest values no further than 1.5X the inter-quartile range from the hinge. Individual data points outside the whisker range are defined as outliers. (C-E) Above: Heat maps of Spearman correlation scores between ChIP-seq data sets ordered by unsupervised hierarchical clustering, comparing the distribution of reads genome-wide (C), at MYB-bound enhancers (D) or MYB-bound promoters (E). Below: Scatter plots of binned ChIP-seq reads for MYB, CBP and P300 at the indicated regions. (F) Bar graph visualisation of pairwise Spearman correlation of ChIP-seq reads, comparing MYB, CBP and P300 at the indicated genomic loci.

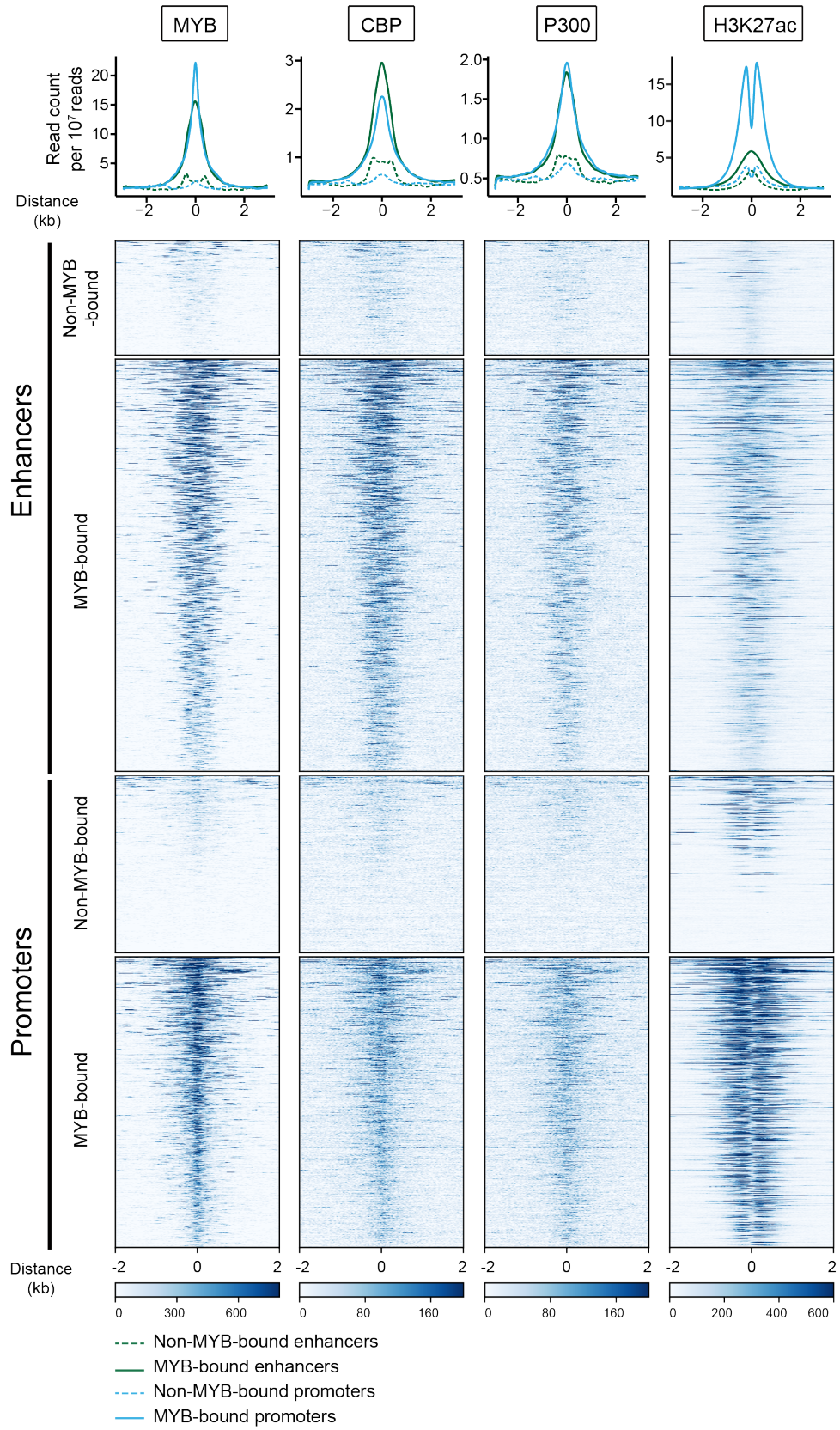



**Figure 5.3: Correlational metaprofiles and heatmaps of MYB, CBP, P300 and H3K27ac at gene regulatory elements.**

enhancers were also associated with a near absence of H3K27ac (Figure 5.3). Of note, there appears to be a subset of promoters marked by H3K27ac (Figure 5.3) and H3K4me3 (Figure 5.4) which are not bound by MYB. In particular, H3K4me3 recruitment appears to be indifferent to the presence or absence of MYB. Also, with the caveat that this could be an artefact of low-level MYB binding, the observation of H3K27ac at some non-MYB-bound promoters suggests that other factors may be able to stabilise its recruitment, consistent with the fact that other transcription factors are known to interact with CBP and P300 (Dai *et al.* 1996; Eckner 1996; Oliner *et al.* 1996; Yuan *et al.* 1996). Together, these data indicate that MYB may play a role in recruiting H3K27ac to a large proportion of enhancers in THP-1 cells.

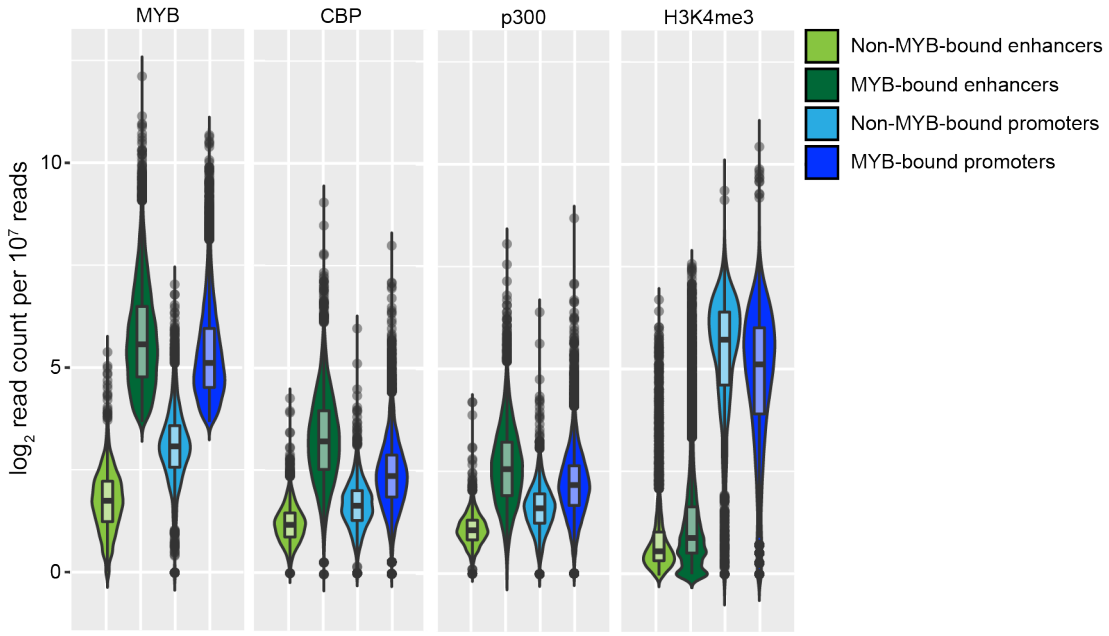



**Figure 5.4: Correlational violin plots of MYB, CBP, P300 and H3K4me3 at gene regulatory elements.** Fixed width violin plots of $\log_2$ read counts per 10 million ChIP-seq reads at MYB-bound and non-MYB-bound enhancer and promoters, with overlaying box-plot, where the lower and upper hinges correspond to the 1st and 3rd quartiles respectively and whiskers extend to the smallest and largest values no further than 1.5X the inter-quartile range from the hinge. Individual data points outside the whisker range are defined as outliers.

---

**Figure 5.3:** continued. Metaprofiles (above) and heat maps (below) generated using DeepTools (v3.0.1) computeMatrix centred on enhancers or promoters, separated based on the presence of a MYB peak, with a 6kb window. Each ChIP-seq is colour-scaled separately. Rows were ordered by average intensity of all four ChIP-seq tracks.

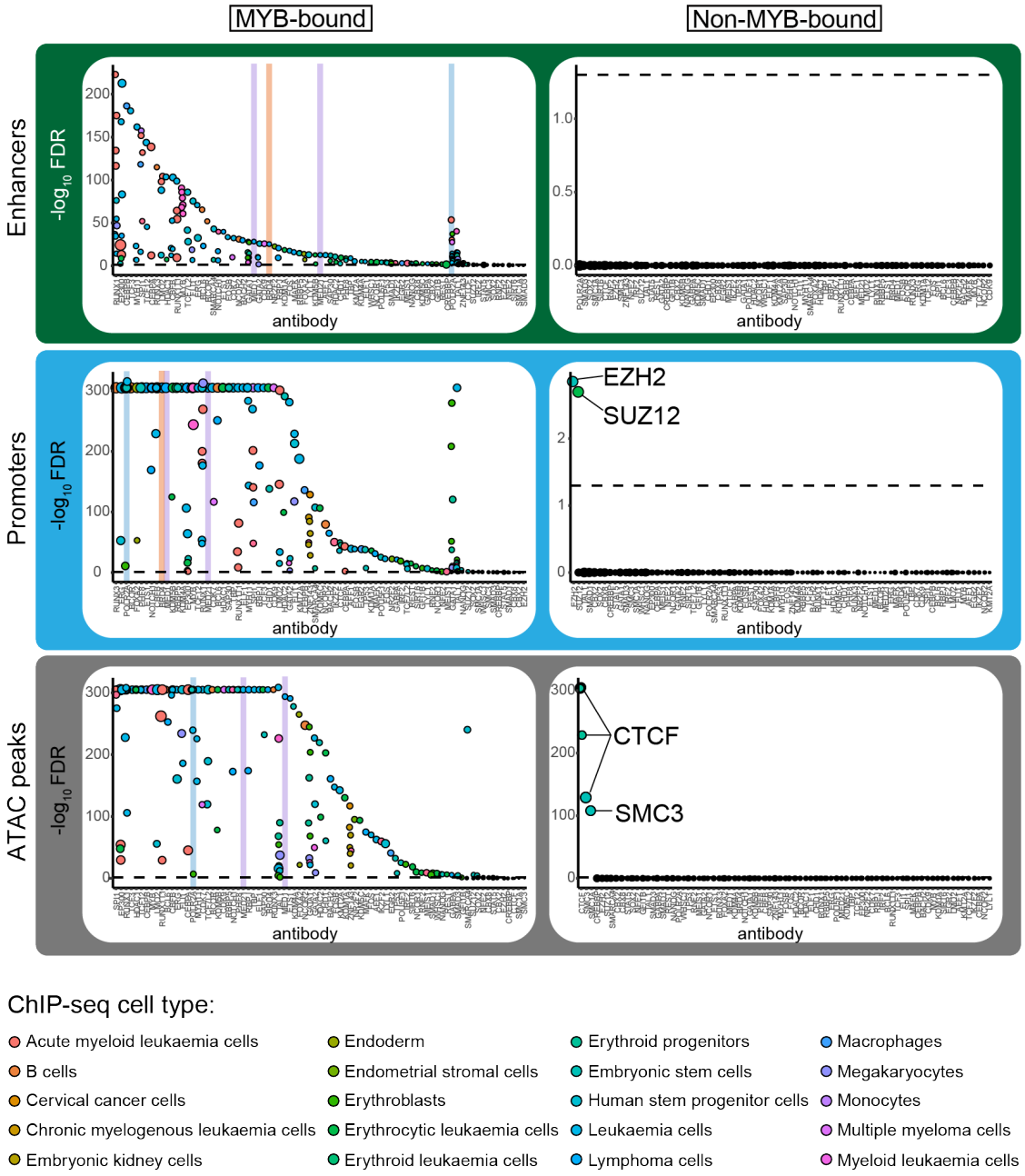



**Figure 5.5: LOLA analysis at MYB-bound vs non-MYB-bound peaks.** MYB-bound and non-MYB-bound enhancers, promoters and ATAC peaks were tested for enrichment of published ChIP-seq datasets using the R package LOLA. Individual dots represent different ChIP-seq datasets and are colour coded according to cell type. Dot size corresponds to the odds ratio of the statistical test. Data for Mediator subunits (MEDI & MED21) highlighted by purple bars, BRD4 by orange bars and the RNAPII subunit POLR2A by blue bars.

We next asked whether the genomic regions identified were characterised by the presence of other chromatin factors. Locus overlap analysis (LOLA) (Sheffield *et al.*

2016) was used to look for enrichment of published ChIP-seq data sets for chromatin-associated factors across a variety of cell types, which could suggest the classes of protein expected to be bound in THP-1 cells. Firstly, MYB-bound enhancer, promoter and ATAC peaks were found to correlate significantly with a wide range of TFs, transcriptional co-activators e.g. Mediator (MED1 & MED21) and BRD4, as well as RNAPII (Figure 5.5). Conversely, non-MYB-bound promoter peaks only showed significant enrichment for Polycomb Repressive Complex 2 (PRC2) proteins EZH2 and SUZ12, while non-MYB-bound ATAC peaks only correlated with structural proteins CTCF and the Cohesin subunit SMC3. This provides supportive evidence that MYB plays a predominantly activating role and may be functioning at gene regulatory elements by facilitating or participating in co-operative binding with other TFs, stabilising activating complexes and RNAPII at these loci.

### 5.3.2 Identifying a set of MYB sensitive gene targets in MLL-AF9 leukaemia

Having identified a strong correlation between MYB binding and enhancer/promoter-related factors, it was necessary to test whether this relationship was causal, i.e. whether MYB binding was required for enhancer or promoter function. As a convenient and reproducible method to deplete MYB levels, MYB knockdown (KD) was used. KD was optimised using two different short interfering RNAs (siRNAs) to achieve maximum KD efficiency in THP-1 cells. There was no significant difference in efficacy between 10µM or 20µM concentrations of either siRNA on RNA or protein levels of MYB or of the known MYB target genes *BCL2, MYC and SMYD2* (Figure 5.6A-C). As has been previously observed, MYB was essential for cell growth, as KD almost completely eliminated the clonogenic potential of THP-1 cells (Figure 5.6D). As siRNA 1 appeared to be marginally more effective both in KD efficiency and in reducing colony formation, this was used in downstream experiments.

Next, a time course experiment was performed to ascertain the optimal time point for analysis. Based on qRT-PCR and western blot data, 24 hours post-transfection was chosen, as this was one of the earliest time points where loss of

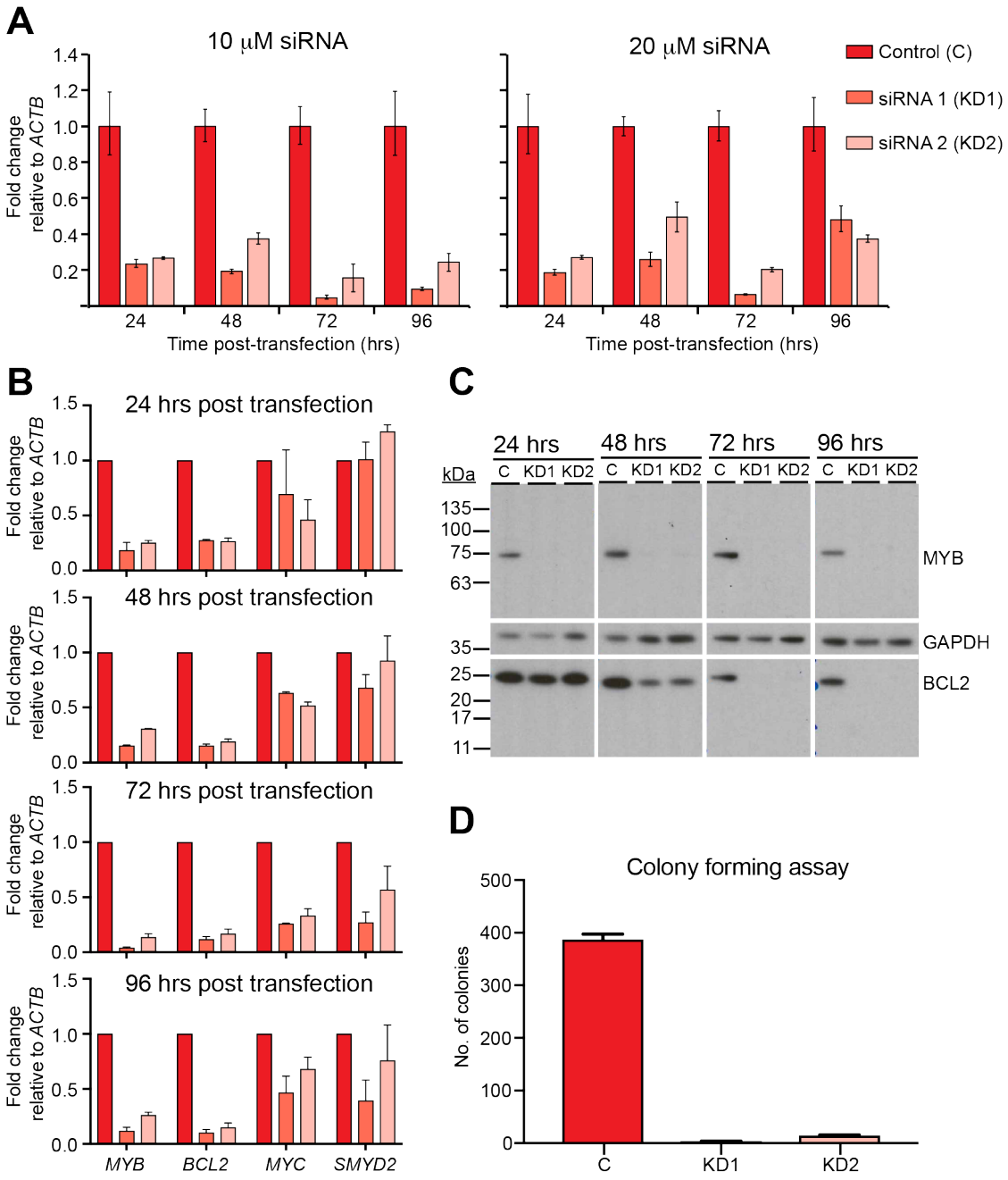



**Figure 5.6: Optimisation of MYB knockdown using targeting siRNA.** (A) qRT-PCR analysis of *MYB* expression following transfection with 10µL of 10µM or 20µM siRNA per 2 x $10^7$ cells in 800µL media. (B) Time-course qRT-PCR analysis of the expression of known *MYB* target genes following transfection with 10µL of 10µM of siRNA 1. Values in (A) and (B) are normalised to *ACTB* expression, relative to non-targeting siRNA control. Mean of three biological replicates; error bars represent SD.

MYB protein was confirmed, and additionally, allowed the cells sufficient time to recover from the effects of electroporation (Figure 5.7A and B).

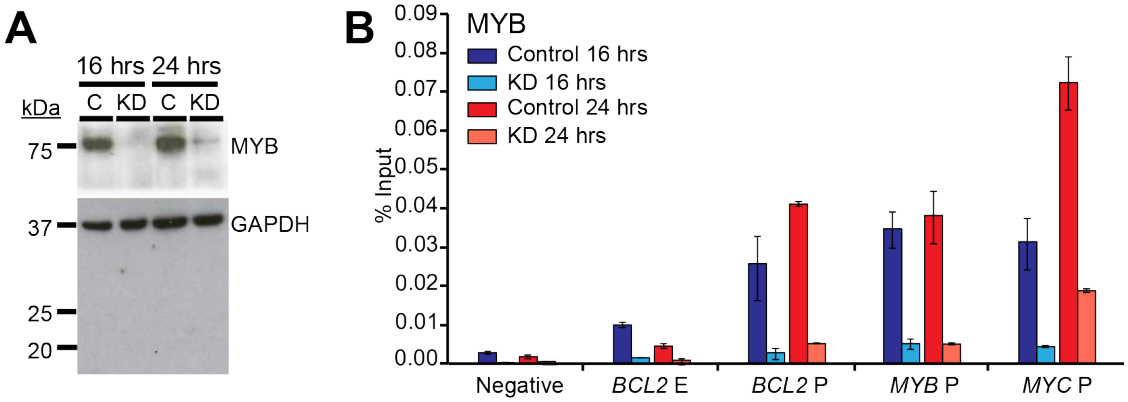



**Figure 5.7: Time course analysis of MYB KD.** (A) Representative Western blot analysis of WCE of THP-1 cells harvested at 16 and 24 hours following transfection with 10µL of 10µM of siRNA 1. Western blot membrane probed using antibodies against MYB and GAPDH (as loading control). (B) MYB ChIP qRT-PCR analysis at the *BCL2* enhancer (*BCL2* E), *BCL2* promoter (*BCL2* P), *MYB* promoter (*MYB* P), *MYC* promoter (*MYC* P) and negative control region in THP-1 cells harvested at 16 and 24 hours following transfection as above. Mean of three biological replicates; error bars represent standard deviation.

In order to investigate the effect of loss of MYB on transcription, reference normalised MYB ChIP-seq (ChIP-rx) (as per Orlando *et al.* 2014) and nascent RNA-seq were performed 24 hrs after siRNA transfection (Figure 5.8). ChIP-rx was used as it allows relative quantitative comparisons in MYB occupancy genome-wide, while nascent RNA-seq was chosen to avoid the masking of dynamic differential changes occurring at the level of transcription by steady state RNA levels.

ChIP-rx confirmed global loss of MYB binding (Figure 5.8B) and this corresponded to changes in nascent transcript levels of >4000 genes (Figure 5.8C-D). Differentially expressed genes were defined as those whose expression showed a $\log_2$

**Figure 5.6:** continued. (C) Representative Western blot analysis of WCE of THP-1 cells following transfection with 10µL of 10µM of siRNA 1 harvested at various time points following transfection. Western blot membrane probed using antibodies against MYB, BCL2 and GAPDH (as loading control). (D) Colony counts 14 days after plating 500 transfected cells onto 1mL of semisolid methylcellulose medium. Data are the mean ± SD of three independent experiments. Three replicates were plated per experiment.

fold change of >0.5 and an adjusted p-Value (FDR) of <0.05. Of 15087 expressed genes in THP-1 cells, there were 1699 downregulated and 2305 upregulated genes following MYB KD.

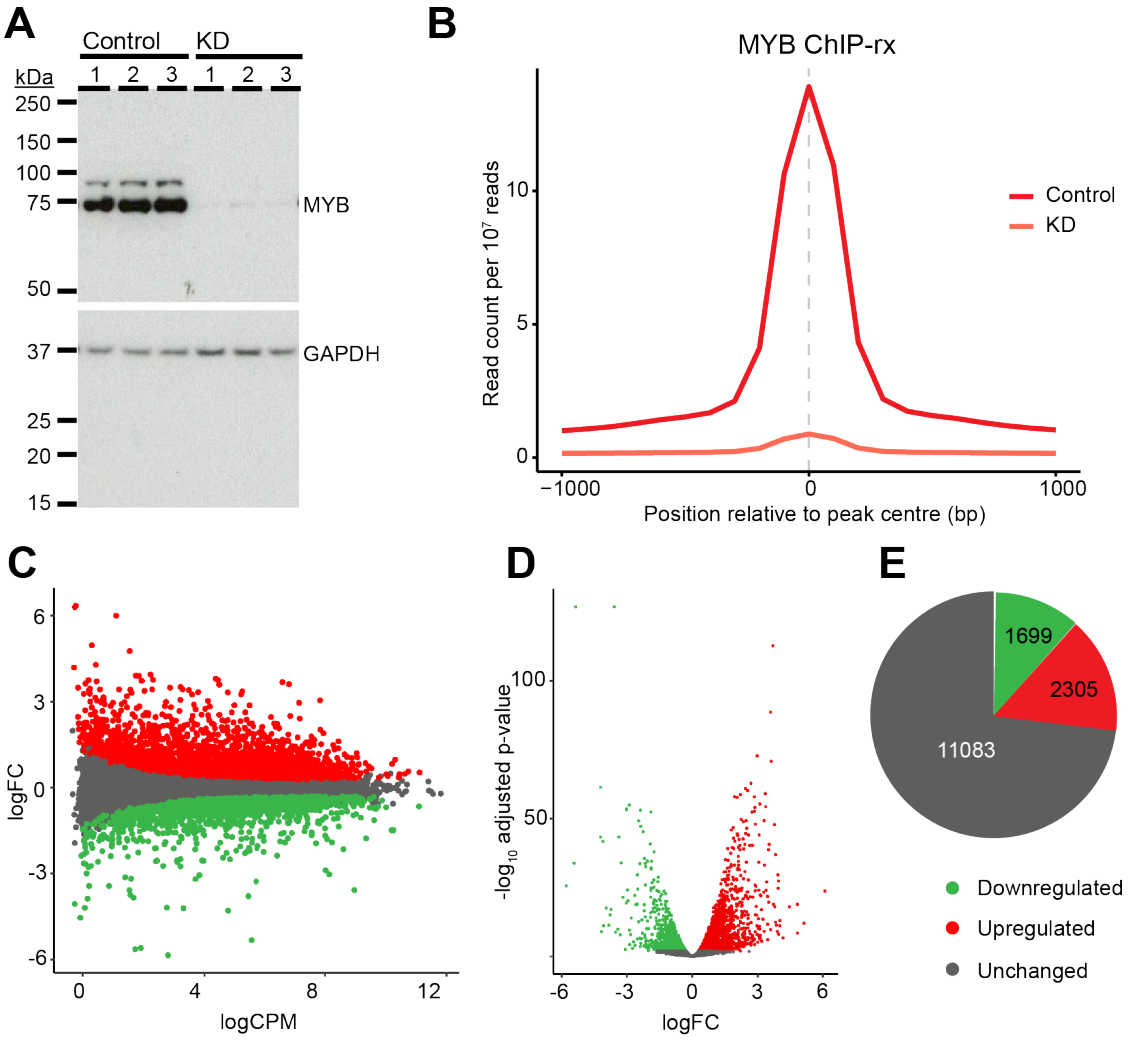



**Figure 5.8: Loss of MYB binding results in changes in transcription.** (A) Western blot validation of MYB KD in the three replicates used for ChIP-rx. (B) Metaprofiles of MYB ChIP-rx signal at all MYB peaks 24 hours following siRNA transfection (pink) compared to control (red) in THP-1 cells. Plot generated using DeepTools (v3.0.1) computeMatrix centred on MYB peaks with a 2kb window. (C) MA plot and (D) Volcano plot of nascent RNA-seq data showing differential gene expression in THP-1 cells following MYB knockdown - downregulated (green), upregulated (red) and unchanged (grey). Differential expression = FDR (false discovery rate) < 0.05 from three biological replicates. log CPM (counts per million)/FC (fold-change) is averaged over all samples. (E) Proportion of differentially expressed and unchanged genes, excluding genes not expressed in THP-1 cells.

### 5.3.3 MYB dependent gene targets have higher levels of MYB binding and are important for cell proliferation and replication

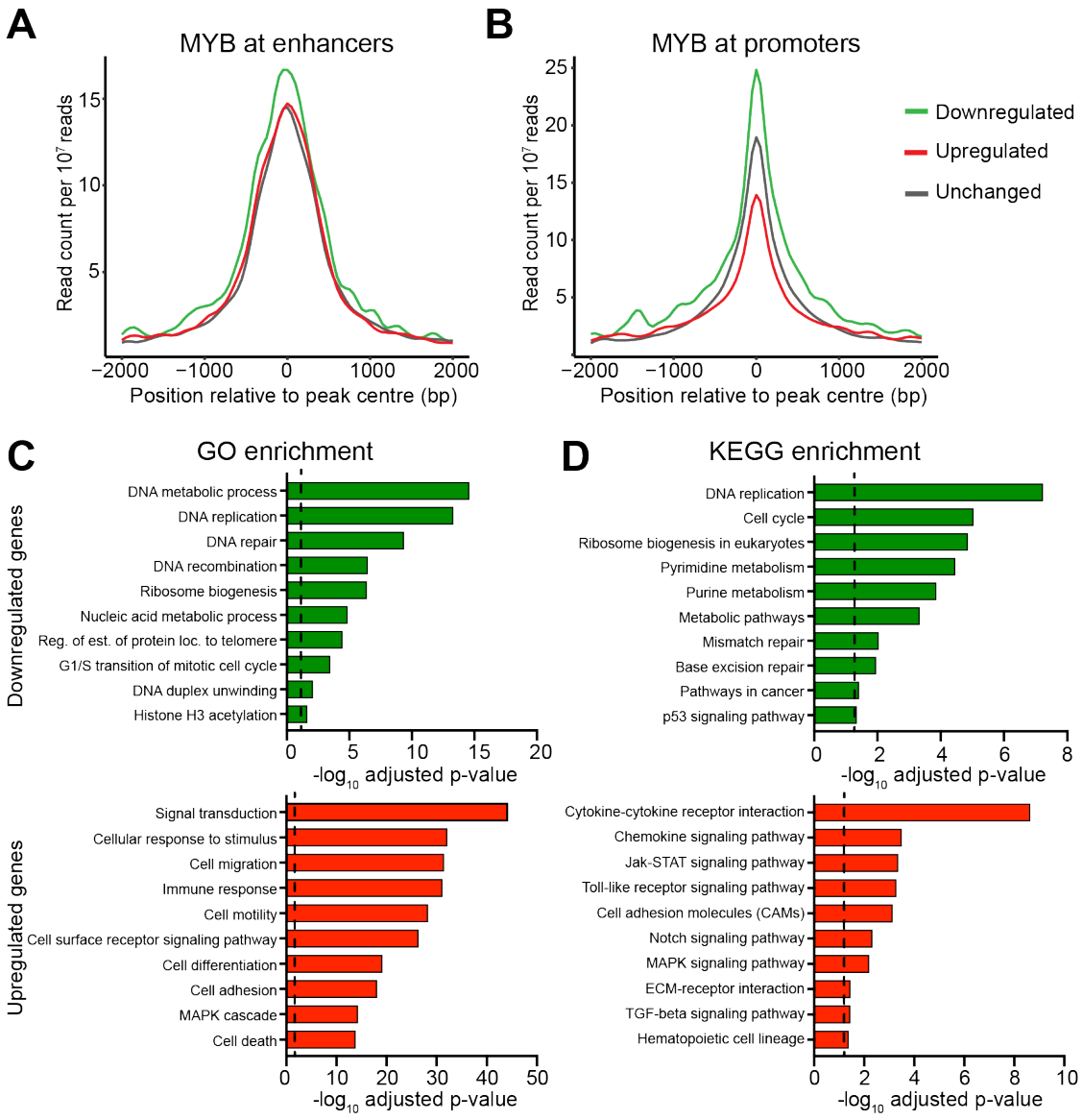



**Figure 5.9: Loss of MYB binding results in transcriptional changes affecting leukaemia maintenance.** Metaprofiles of MYB ChIP-seq signal centred on (A) enhancers and (B) promoters of downregulated (green), upregulated (red) and unchanged (grey) genes with a 4kb window. Biological pathway (C) GO and (D) KEGG terms enriched in downregulated and upregulated genes (logFC <-1 or >1, FDR <0.05) following MYB KD.

In order to understand the role of MYB at enhancers in transcription, it was necessary to assign each enhancer to a target gene. We applied the commonly-used nearest

neighbour approach (Godfrey *et al.* 2019), where enhancers are annotated with the nearest expressed promoter. While this approach may not accurately capture all enhancer-promoter relationships, it provides a reasonable approximation to provide a genome-wide picture that can be verified downstream. Thus, a comparison of MYB levels between downregulated, upregulated and unchanged genes, revealed that downregulated genes appear to have higher levels of MYB both at their promoters, as well as their assigned putative enhancers (Figure 5.9 A and B). This suggests that genes associated with higher levels of MYB are more likely to be sensitive to loss of MYB, arguing that MYB may be directly regulating these targets.

Furthermore, downregulated genes were enriched for biological pathways associated with replication and cell cycle progression, consistent with the biological role of MYB in maintaining a proliferative state in leukaemia. Upregulated genes were associated with differentiation, haematopoietic lineage, and cell death, supportive of the effect observed in leukaemia cell lines upon loss of MYB (Figure 5.9C). Note that some of these transcriptional changes may be a consequence of the cellular response to MYB KD, rather than being direct targets of the transcription factor.

### 5.3.4 H3K27ac at enhancers and promoters is sensitive to MYB knockdown

As MYB directly interacts with the lysine acetyltransferases CBP and P300, and H3K27ac marks active enhancers and promoters, it was important to understand the effect of loss of MYB on H3K27ac levels genome-wide and how this relates to changes in transcription.

To address this question, H3K27ac ChIP-rx was performed 24 hours after MYB KD. This revealed an overall reduction in the amount of H3K27ac at MYB-bound enhancers and promoters compared to other MYB-peaks (Figure 5.10A). Surprisingly, despite the fact that MYB levels are higher at enhancers compared to promoters (Figure 5.2B), the opposite is true for H3K27ac, where higher levels are observed at promoters compared to enhancers (Figure 5.10A). A likely explanation for the higher level of H3K27ac at promoters, is that at these regions, there may

be many other TFs contributing to histone acetyltransferase (HAT) recruitment and possibly higher levels of stabilising/protective protein complexes. Conversely, while enhancers appear to require less H3K27ac for their activity, it may be the case that there is less redundancy in TF binding at these elements, therefore higher levels of MYB are required to maintain H3K27ac levels. This supported by LOLA analysis (Figure 5.5), where more proteins co-localise with MYB-bound promoters than at MYB-bound enhancers.

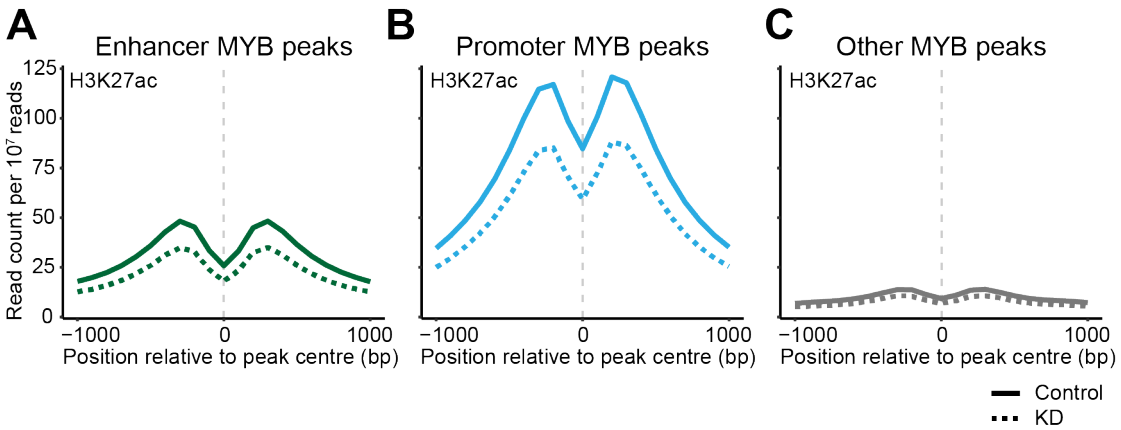



**Figure 5.10: Loss of MYB binding results in decreased H3K27ac at gene regulatory elements.** Metaprofiles of H3K27ac ChIP-rx signal at MYB-bound (A) enhancer, (B) promoter and (C) non-enhancer/promoter peaks, 24 hours after MYB KD (dotted lines) compared to control (solid lines) in THP-1 cells. Plot generated using DeepTools (v3.0.1) computeMatrix centred on MYB peaks with a 2kb window.

Consistent with this concept, MYB appears to be important for the maintenance of H3K27ac at some loci more than others. An example of a MYB dependent gene is *BCL2*, which has a 3' enhancer that has been described across leukaemia cell types (Godfrey *et al.* 2019; Godfrey *et al.* 2016; Tzelepis *et al.* 2018). MYB is bound at high levels at the *BCL2* 3' enhancer, as well as at lower levels at its promoter and loss of MYB binding at both these loci (Figure 5.11A) was associated with concurrent loss of transcription at this gene (Figure 5.11B). Interestingly, there was a disproportionate loss of H3K27ac at the enhancer compared with its promoter (Figure 5.10C), and most MYB-bound intragenic H3K27ac peaks at this locus in fact remain largely unchanged. This indicates that MYB is more important for the maintenance of H3K27ac at particular loci, and is necessary for H3K27ac at the *BCL* enhancer.

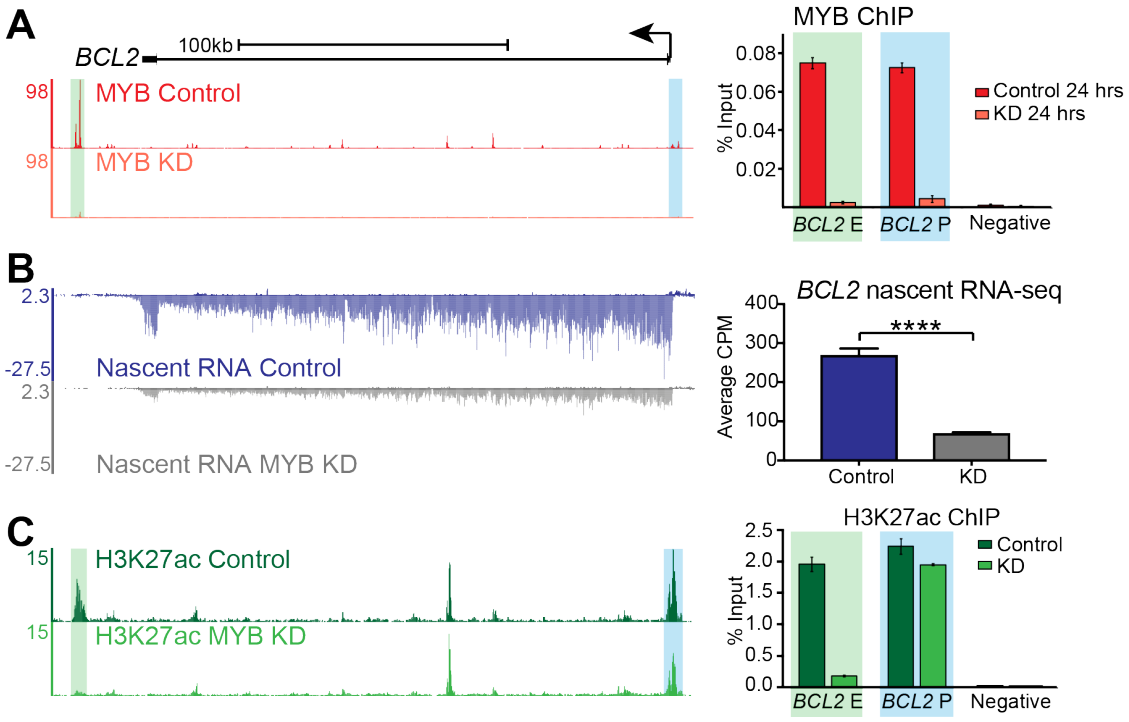



**Figure 5.11: *BCL2* is an example of a MYB dependent gene target.** (A) MYB ChIP-rx tracks at the *BCL2* locus 24 hours after MYB KD (pink) compared to control (red) and ChIP qPCR validation at the *BCL2* enhancer (*BCL2* E) (green window) and promoter (*BCL2* P) (blue window). Mean of three biological replicates; error bars represent SD. (B) Strand specific nascent RNA-seq tracks at the *BCL2* locus following MYB KD (grey) compared to control (blue) and bar graph plotting average CPM of three biological replicates. (C) H3K27ac ChIP-rx tracks at the *BCL2* locus 24 hours after MYB KD (light green) compared to control (dark green) and ChIP qPCR validation at the *BCL2* enhancer (*BCL2* E) (green window) and promoter (*BCL2* P) (blue window).

### 5.3.5 Loss of MYB binding results in differential changes to chromatin accessibility

In order to understand the role of MYB on maintaining chromatin accessibility and facilitating cooperative binding, ATAC-seq was performed following MYB KD. This revealed that MYB was particularly important for chromatin accessibility at a relatively small number of loci. For example, the near absence of H3K27ac at the *BCL2* enhancer (Figure 5.10B and C) was accompanied by a significant loss of ATAC-signal (Figure 5.12C and D) at this locus. Importantly, the small reduction in H3K27ac observed at the *BCL2* promoter was not associated with a loss of open chromatin (Figure 5.12C and D). Bearing in mind that *BCL2* expression is strongly decreased, this suggests that the activity of the *BCL2* enhancer is

necessary for transcription, and this enhancer activity is largely MYB dependent. It also indicates that loss of active transcription can precede loss of chromatin accessibility at the promoter.

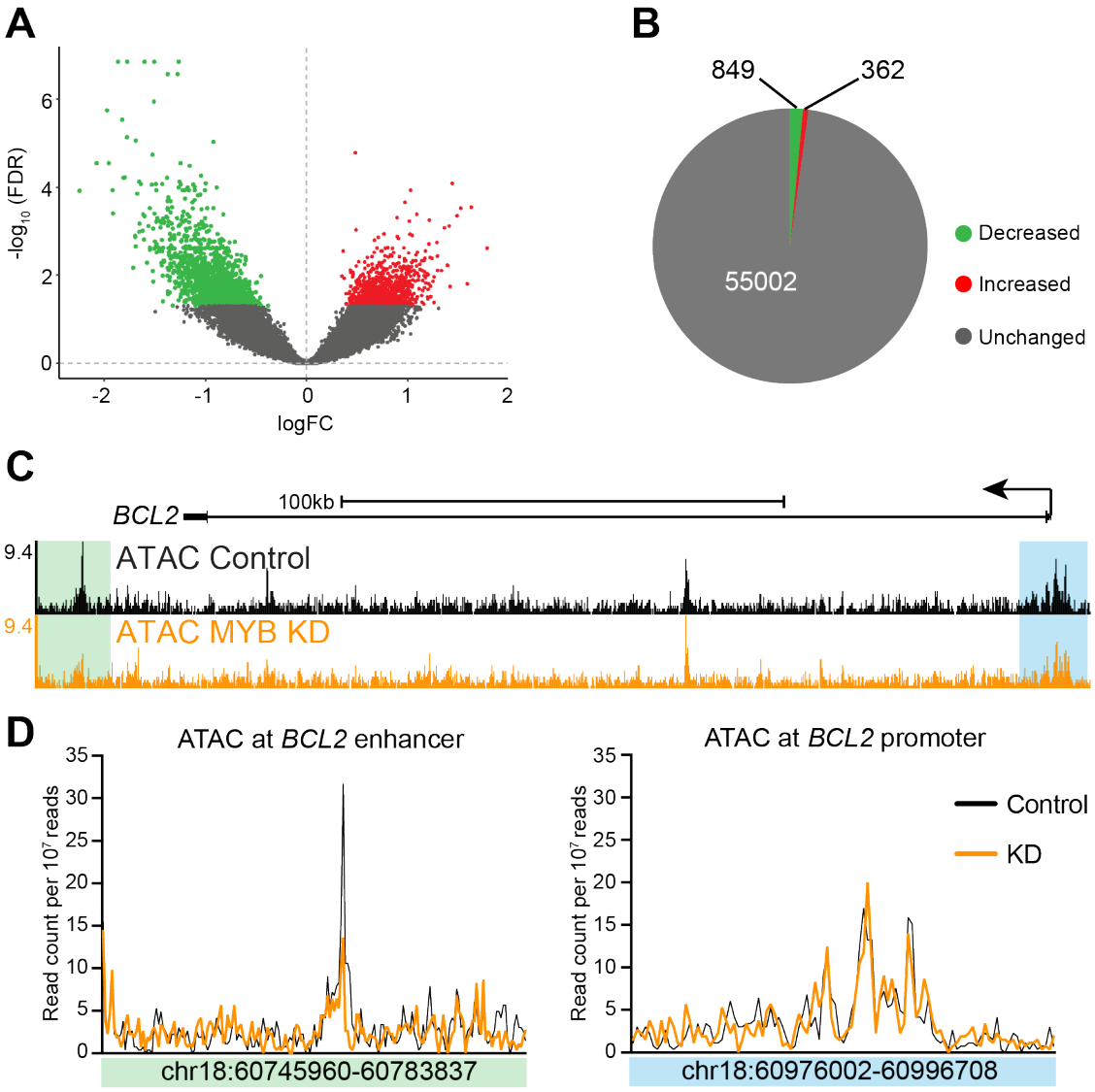



**Figure 5.12: Loss of MYB binding results in sporadic ATAC changes.** (A) Volcano plot of ATAC-seq data showing differential peaks (increased (red), decreased (green), and unchanged (grey)) in THP-1 cells following MYB knockdown. Differential expression = FDR (false discovery rate) $< 0.05$ from three biological replicates. log FC is averaged over all samples. (B) Proportion of differentially expressed and unchanged ATAC peaks. (C) ATAC-seq tracks at the *BCL2* locus 24 hours after MYB KD (yellow) compared to control (black). (D) Histogram of ATAC-seq signal at the *BCL2* enhancer (*BCL2* E) (green window) and promoter (*BCL2* P). (blue window). Representative of three biological replicates.

To determine if these findings could be applied more generally, the distribution

of ATAC changes was analysed genome wide. Taking into account the number of dysregulated genes, a comparatively small number of ATAC peaks were differentially affected by MYB KD. Within the differential peaks, while there did appear to be a trend towards loss of chromatin accessibility, with 849 decreased peaks compared with 362 increased peaks (Figure 5.12A and B), there was no convincing correlation between changes in accessibility with changes in gene expression (data not shown). This could be attributed not only to the challenges of annotating distal regulatory elements to the correct gene, but also to the inherent noise associated with ATAC-seq, with large inter-sample variation within the same experimental conditions masking subtle changes. Thus, the above results indicate that overall, loss of MYB binding leads to loss of transcription at MYB dependent gene targets, but without predictable changes in chromatin accessibility.

### 5.3.6 Loss of MYB binding does not affect TAD structure or enhancer-promoter looping

As discussed in Chapter 1.1.3.1, for the majority of genes, enhancer and promoter proximity has been shown to be essential for enhancer function and gene activation (Brown *et al.* 2018; H. Chen *et al.* 2018). The mechanisms that drive this proximity are unclear, but it is likely that enhancer associated proteins have a role to play in chromatin looping. For example, TFs such as MYB are thought to facilitate the assembly of phase condensates at gene regulatory elements, which may mediate and/or stabilise enhancer-promoter interactions (Boija *et al.* 2018; Nolis *et al.* 2009; Sabari *et al.* 2018; Shrinivas *et al.* 2019).

Given that loss of MYB binding was associated with decreased transcription at a subset of genes, one possibility is that MYB was contributing to gene activation by maintaining enhancer-promoter interactions. Therefore, Next Generation Capture-C (Davies *et al.* 2016) was performed to capture interactions with the promoters of selected downregulated genes (n = 5), and as controls, from the promoters of upregulated (n = 7) and unchanged genes (not significantly downregulated (n =

7), not significantly upregulated (n = 10)), using a previously designed probe set available in the lab (Godfrey *et al.* 2019).

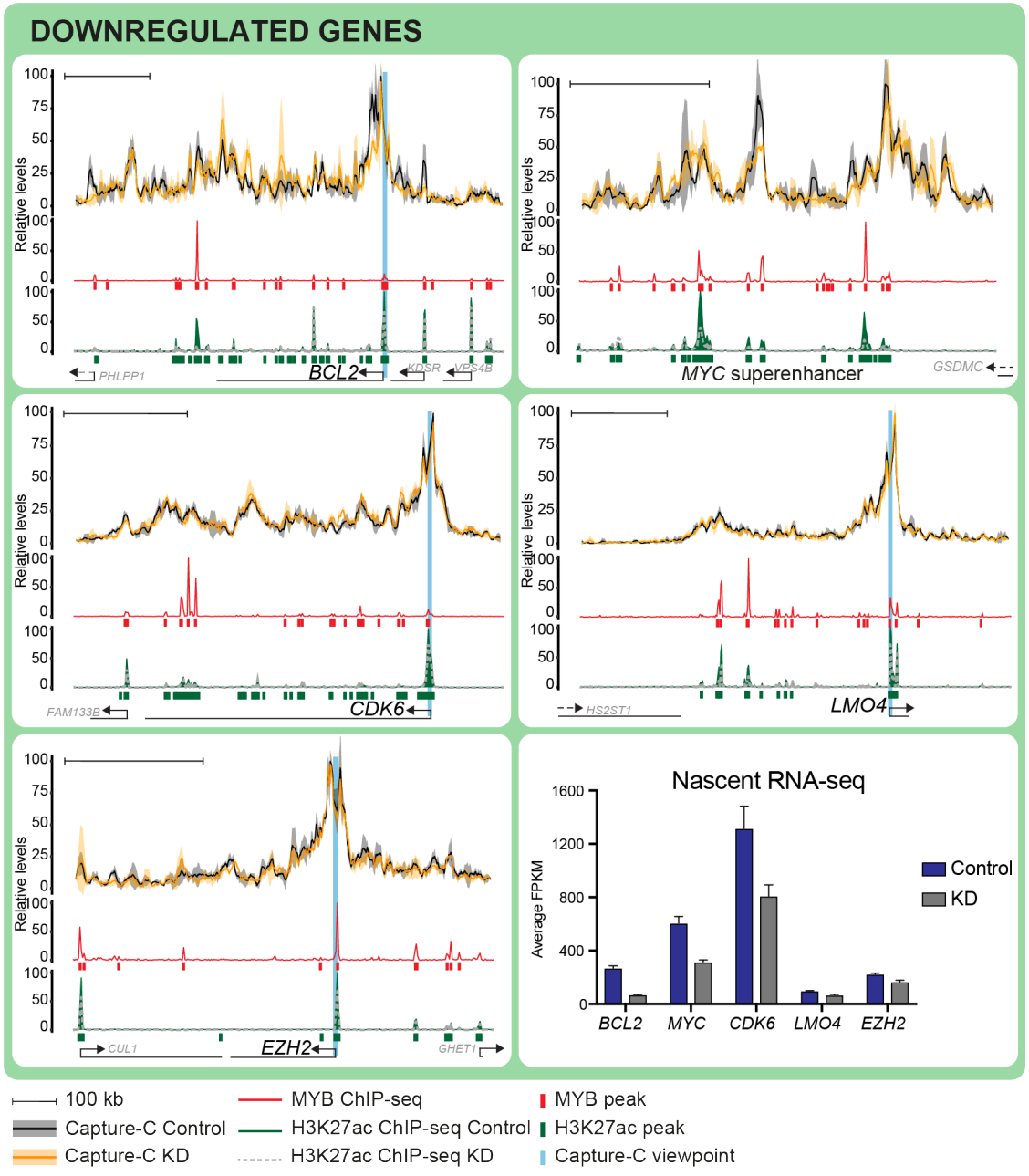



**Figure 5.13: Capture-C traces at genes which are transcriptionally downregulated following MYB knockdown.** Capture-C traces at gene loci 24 hours after MYB KD (yellow line) compared to control (black line). Ribbons shows +/-1 SD for three replicates. Vertical blue bar indicates the capture point for each gene. The MYC superenhancer is ~1.7 Mb downstream of the MYC promoter, where the capture viewpoint is. MYB ChIP-seq traces at the same locus are shown in red. H3K27ac ChIP-rx traces 24 hours after MYB KD are shown as grey dotted lines and control traces are underlaid in green. Scale bars represent 100 kb. Nascent RNA-seq levels are shown for each gene in control cells (blue) or following MYB KD (grey).

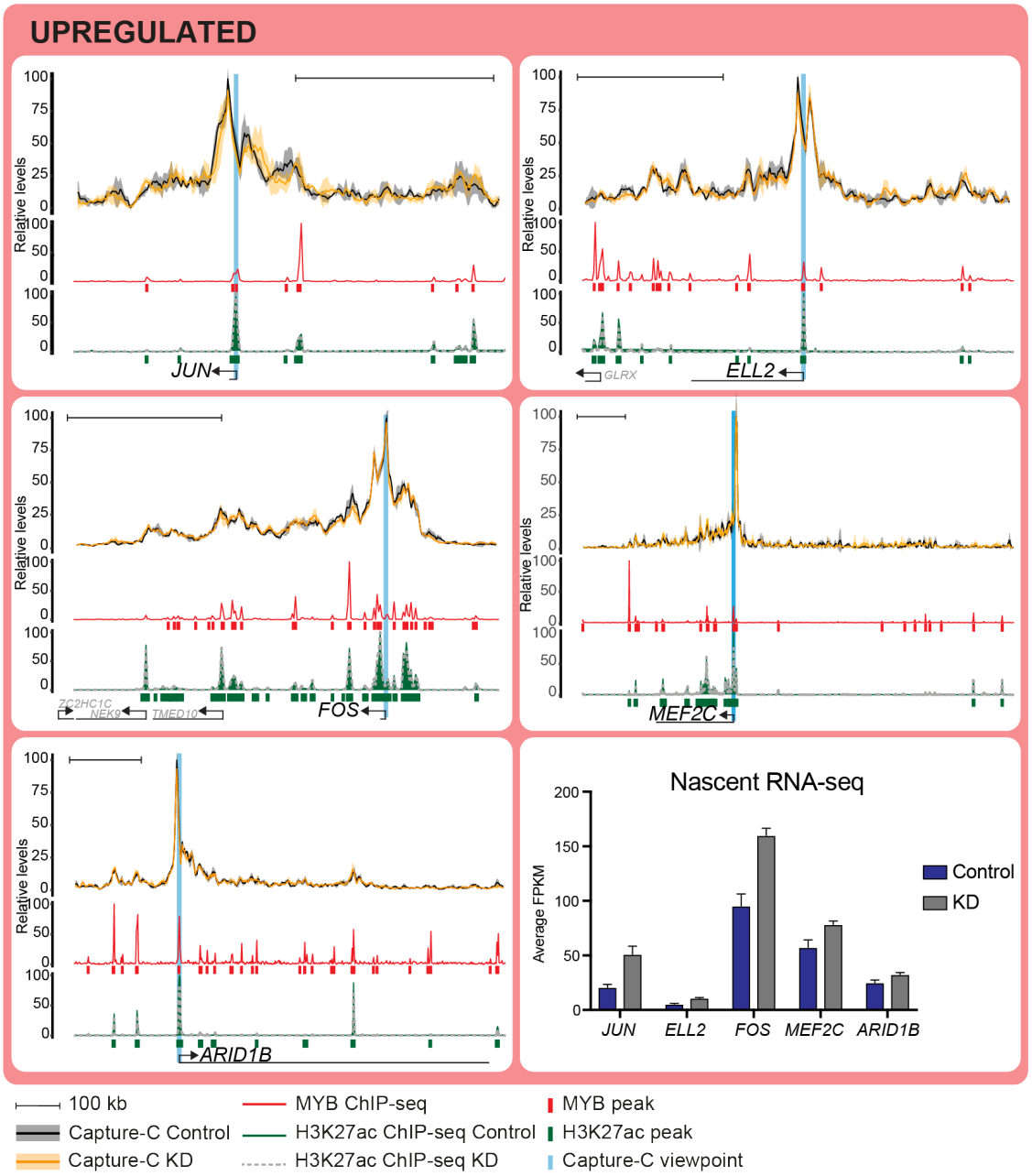



**Figure 5.14: Capture-C traces at genes which are transcriptionally upregulated following MYB knockdown.** Capture-C traces at genes 24 hours after MYB KD (yellow line) compared to control (black line). Ribbons shows +/-1 SD for three replicates. Vertical blue bar indicates the capture point for each gene. MYB ChIP-seq traces at the same locus are shown in red. ChIP-rx traces for H3K27ac ChIP-rx traces 24 hours after MYB KD are shown as grey dotted lines and control traces are underlaid in green. Scale bars represent 100 kb. Nascent RNA-seq levels are shown for each gene in control cells (blue) or following MYB KD (grey)

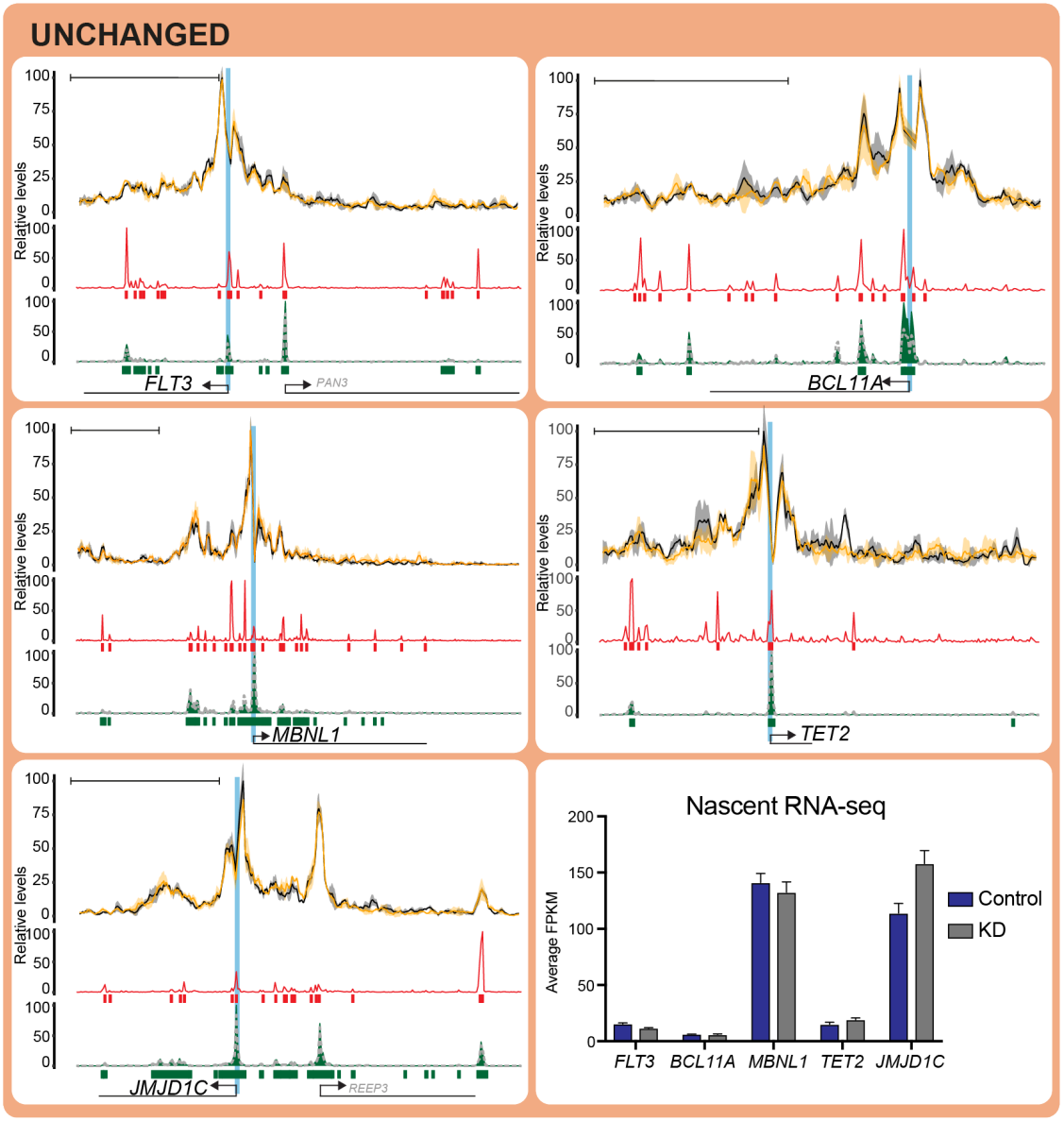



**Figure 5.15: Capture-C traces at genes which are transcriptionally unchanged following MYB knockdown.** Capture-C traces at genes 24 hours after MYB KD (yellow line) compared to control (black line). Ribbons shows +/-1 SD for three replicates. Vertical blue bar indicates the capture point for each gene. MYB ChIP-seq traces at the same locus are shown in red. ChIP-rx traces for H3K27ac ChIP-rx traces 24 hours after MYB KD are shown as grey dotted lines and control traces are underlaid in green. Scale bars represent 100 kb. Nascent RNA-seq levels are shown for each gene in control cells (blue) or following MYB KD (grey)

Strikingly, loss of MYB binding had little or no effect on the frequency of enhancer-promoter interactions at all genes assayed, irrespective of their

transcriptional response to MYB KD (Figures 5.13-5.15, Appendix 2-4). For example, at *BCL2*, interaction frequency between the enhancer and promoter was maintained despite the loss of transcription, as well as reduction in H3K27ac and chromatin accessibility at the enhancer. Similarly at the MYC enhancer, while subtle reductions in interactions with the promoter were observed, these were not statistically significant, despite loss of H3K27ac and transcription from the promoter (Figure 5.13). It is possible that 24 hours after siRNA transfection is too early a time point to generate changes in chromatin structure, or that their detection is limited by the resolution of Capture-C, and perhaps more measurable changes may be apparent with more sensitive techniques. However, what these findings do suggest, is that whilst MYB binding may be necessary to maintain enhancer activity, it does not function primarily through stabilising physical proximity with the gene promoter. Furthermore, activation of transcription and maintenance of enhancer-promoter interactions appear to be temporally separable events, and factors mediating the two are likely to be distinct from each other.

### 5.3.7 MYB is necessary for eRNA transcription at the enhancers of MYB dependent gene targets

As mentioned in Chapter 1.1.3.2, several studies have found that eRNA expression is highly correlated with changes in the expression of nearby genes (T.-K. Kim *et al.* 2010; Lam *et al.* 2013; Wang *et al.* 2011). Therefore, cognate enhancers of the MYB dependent gene targets *BCL2*, *MYC* and *EZH2* were identified by Capture-C analysis, and eRNA transcription from these loci was examined after MYB KD using nascent RNA-seq and poly(A) minus RNA-seq (Figure 5.16-5.18).

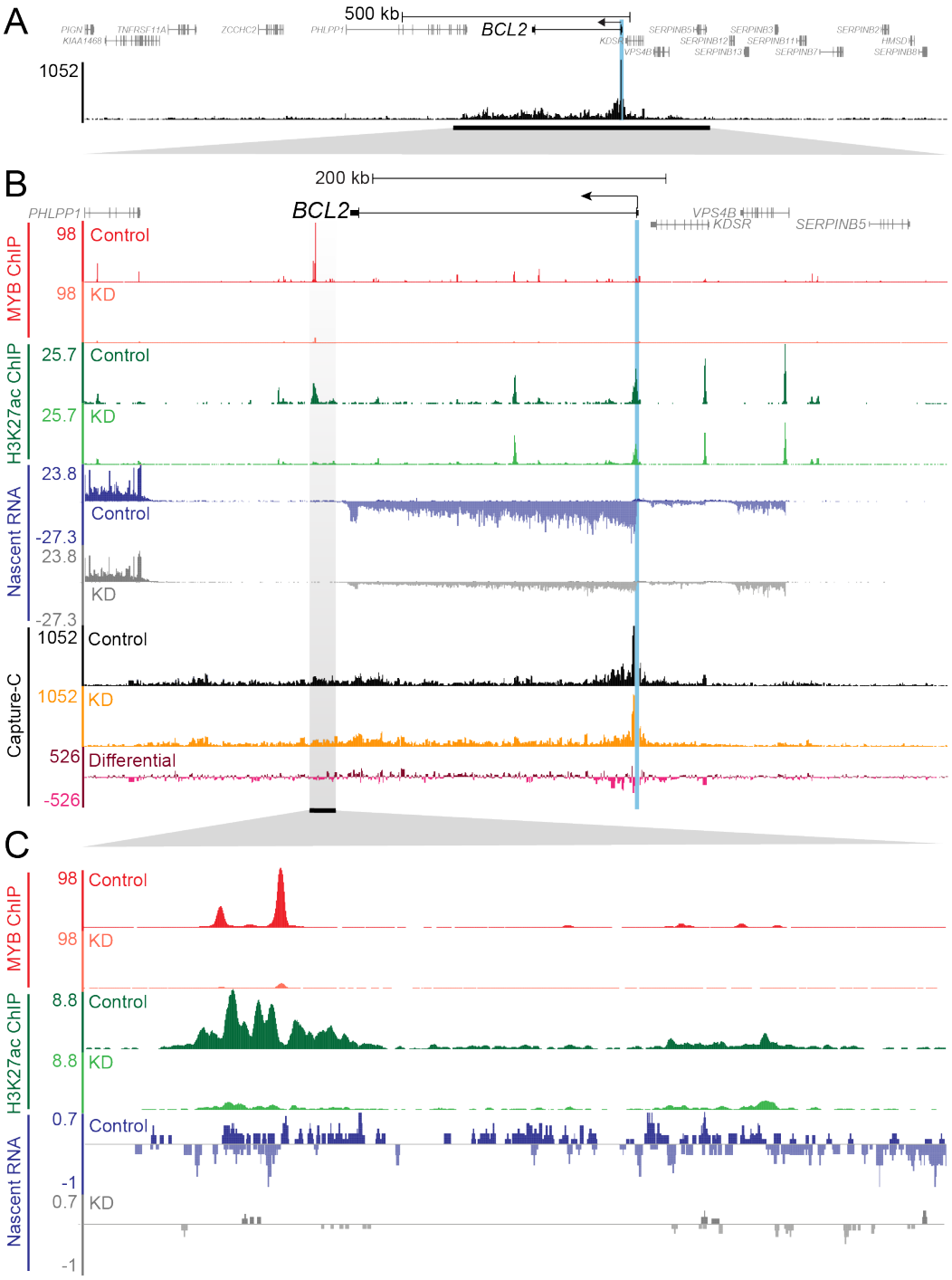



**Figure 5.16: MYB is required for eRNA transcription at *BCL2*.** (A) Interaction domain of the *BCL2* promoter. Vertical blue bar indicates the capture point. (B) Windowed view of Capture-C from the *BCL2* promoter following MYB KD (yellow) compared to control (black). Differential tracks show the change in Capture-C profile in MYB KD samples compared to control. Accompanying MYB ChIP-rx, H3K27ac ChIP-rx and stranded nascent RNA-seq tracks are displayed. (C) Windowed view of stranded nascent RNA-seq at the *BCL2* enhancer following MYB KD (grey) compared to control (blue).

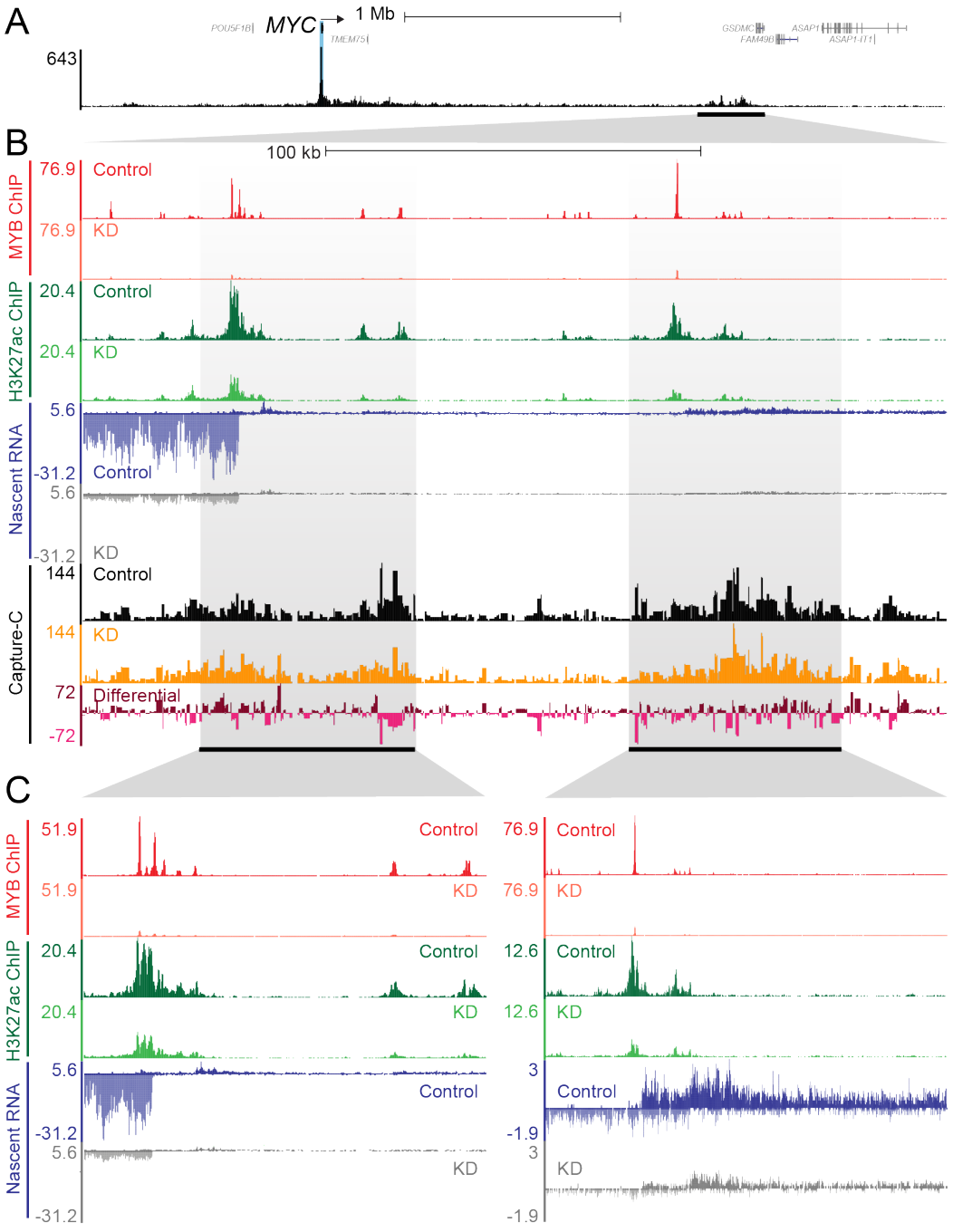



**Figure 5.17: MYB is required for eRNA transcription at *MYC*.** (A) Interaction domain of the *MYC* promoter. Vertical blue bar indicates the capture point. (B) Windowed view of Capture-C from the *MYC* promoter following MYB KD (yellow) compared to control (black). Differential tracks show the change in Capture-C profile in MYB KD samples compared to control. Accompanying MYB ChIP-rx, H3K27ac ChIP-rx and stranded nascent RNA-seq tracks are displayed. (C) Windowed view of stranded nascent RNA-seq at both *MYC* enhancers following MYB KD (grey) compared to control (blue).

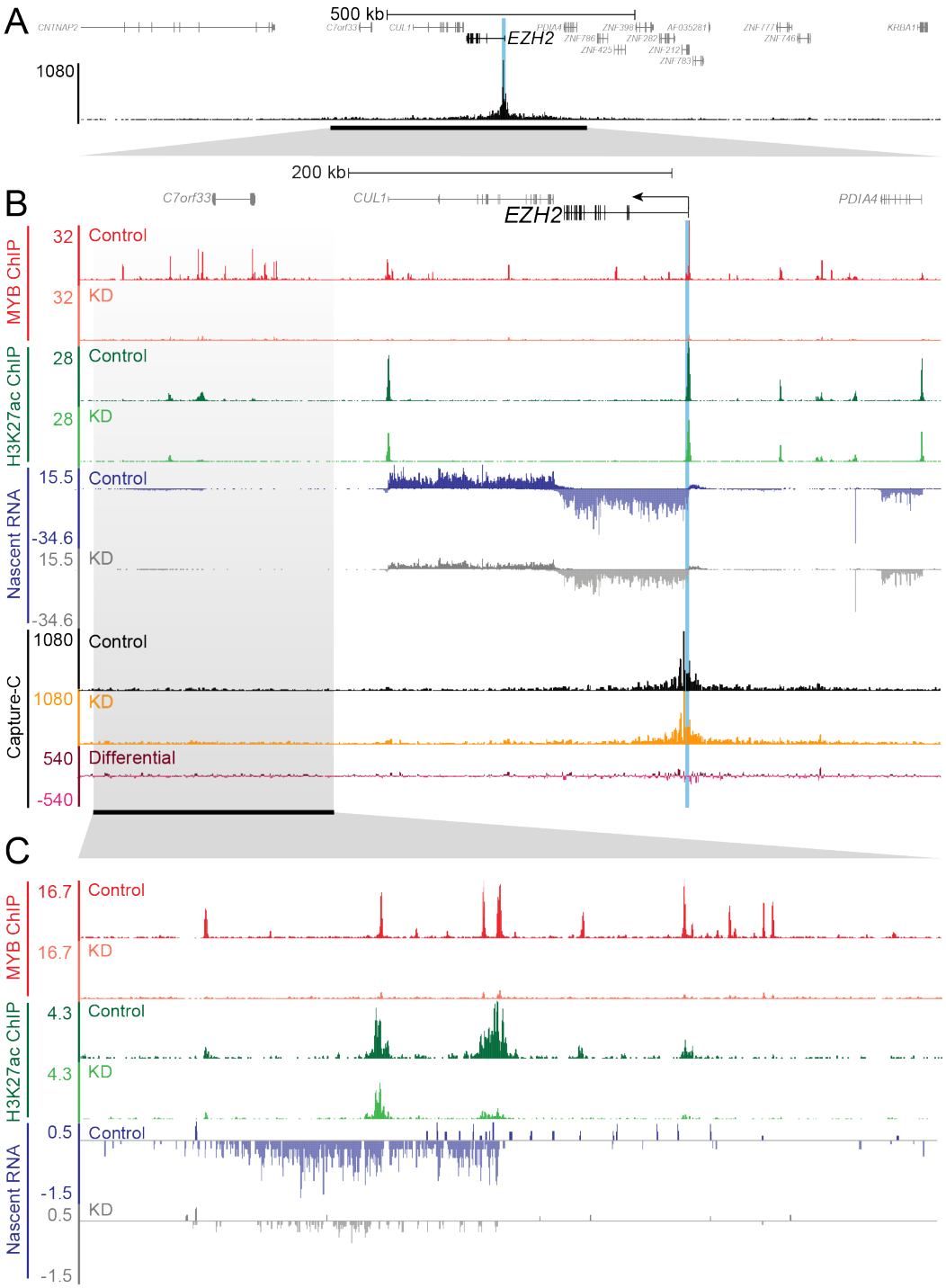



**Figure 5.18: MYB is required for eRNA transcription at *EZH2*.** (A) Interaction domain of the *EZH2* promoter. Vertical blue bar indicates the capture point. (B) Windowed view of Capture-C from the *EZH2* promoter following MYB KD (yellow) compared to control (black). Differential tracks show the change in Capture-C profile in MYB KD samples compared to control. Accompanying MYB ChIP-rx, H3K27ac ChIP-rx and stranded nascent RNA-seq tracks are displayed. (C) Windowed view of stranded nascent RNA-seq at the *EZH2* enhancer following MYB KD (grey) compared to control (blue).

In the control condition, low level transcription was detected at each of these enhancers, consistent with eRNA synthesis. Remarkably, there was a significant reduction in eRNA transcription following MYB KD. Furthermore, the decrease in eRNA transcription at these loci was consistently associated with loss of MYB binding, loss of H3K27ac and loss of chromatin accessibility (Figures 5.13, 5.16-5.18). Together, this suggests a mechanistic role for MYB at a subset of MYB dependent genes, where continued transcription from these gene promoters is dependent on MYB binding at enhancers and maintaining H3K27ac, chromatin accessibility and eRNA transcription. Importantly, MYB does not appear to play a structural role in stabilising enhancer-promoter interactions, or at least changes in transcription due to loss of MYB appear to precede changes in enhancer-promoter interactions.

## 5.4 Discussion

In summary, using an MLL-AF9 leukaemia model, we have demonstrated by ChIP-seq that MYB binding correlates strongly with that of CBP and P300 at enhancers (Figure 5.1-5.4). In addition, we have shown that MYB is necessary for the maintenance of nascent transcription at MYB sensitive gene targets (Figure 5.8-5.9). One caveat of RNA-seq analysis is the assumption that the vast majority of gene expression remains constant, and normalisation is adjusted to the total library size. Therefore, more subtle changes in gene expression between samples could be masked by redistribution of the data. Through normalisation artefact, many downregulated genes may be classified as unchanged, and many unchanged genes may be classified as upregulated. Therefore, although genes classified as downregulated almost certainly are, the difficulty lies in being confident in which genes are upregulated and unchanged. To control for this, exogenous normalisation would be required. However, this was not performed in this thesis. Taking the above into consideration, and the fact that we are interested in the role of MYB in gene activation, the definition of MYB dependent gene targets has been restricted to high confidence downregulated genes that have MYB bound at their promoters and putative enhancers.

MYB dependent gene activation appears to be mediated through its maintenance of enhancer activity, as demonstrated by the loss of H3K27ac, eRNA transcription and chromatin accessibility at the cognate intergenic enhancers of specific MYB sensitive genes (Figures 5.10, 5.12 & 5.16). Interestingly, these observations are in keeping with recent evidence that suggests an interconnected relationship between CBP/P300, H3K27ac and eRNA function. Firstly, targeting the P300-core domain (consisting of the bromo- and acetyltransferase domains) alone to enhancers using dCas9 and enhancer-specific gRNAs has been shown to increase the expression of enhancer-associated genes, suggesting that localisation of HAT activity to enhancers is sufficient for endogenous gene activation (Hilton *et al.* 2015). At the same time, small molecule inhibition (GNE-049) of the CBP/P300 bromodomain has been shown to result in the loss of H3K27ac specifically from enhancers, leading to decreased P-TEFb and RNAPII localisation, suppressed eRNA production and significant downregulation of nearby genes, despite the continued presence of CBP/P300, BRD4 and H3K18ac at chromatin (Raisner *et al.* 2018). On the other hand, it has been shown by Bose *et al.*, that CBP/P300 can interact with RNAs via a regulatory region in its HAT domain. Where CBP/P300 is recruited to enhancer chromatin, eRNA binding in *cis* directly stimulates its acetyltransferase activity, increasing H3K27ac (and H3K18ac) at the enhancer and associated promoter, thus promoting gene expression. Complementary to these findings, eRNA depletion using antisense oligonucleotides leads to loss of H3K27ac at the target enhancer and its promoter, resulting in a concurrent decrease in mRNA expression despite maintained CBP/P300 occupancy at these loci (Bose *et al.* 2017a; Bose *et al.* 2017b). Taken together, our data strongly suggest that MYB dependent maintenance of enhancer function is mediated at least in part by the activity of its co-activators CBP/P300. At these loci, loss of MYB binding results in loss of enhancer function, likely due to the direct loss of CBP/P300 localisation, but also indirectly through concurrent RNAPII destabilisation and the subsequent loss of CBP/P300 activity, which as published data suggests, is possibly related to the reduction in eRNA transcription.

Of note, besides the maintenance of activating histone marks and eRNA transcription at MYB dependent enhancers, we have observed that MYB is required for the maintenance of chromatin accessibility at specific intergenic enhancers such as the BCL2 enhancer. This strongly suggests that the presence of MYB is either required for the co-operative binding of other TFs and/or the stabilisation of activating protein complexes, or that it alone is sufficient and necessary for enhancer activity at these loci. Based on these observations, it would be beneficial to perform Capture-C to identify the cognate enhancers of more MYB sensitive genes, and to then test if enhancer features are also lost at these loci with loss of MYB. This will allow us to understand if there is a common mechanism through which MYB regulates gene expression.

On the other hand, it is important to note that the majority of MYB bound ATAC peaks were not observed to change in a MYB dependent manner. The reason for this is likely due to the presence of compensatory mechanisms such as the presence of independently bound TFs or protein complexes, or that the time point assayed (24 hours post transfection) is too early to capture changes at less responsive loci. Of interest, the fact that we observe maintained chromatin accessibility at the gene regulatory elements of downregulated genes, fits with the observation that open chromatin per se is not sufficient for enhancer or promoter activity (R. Andersson *et al.* 2014a; Klemm *et al.* 2019; Corces *et al.* 2016; Lara-Astiaso *et al.* 2014; Dogan *et al.* 2015). At MYB sensitive genes, it would be interesting to attempt to identify the characteristics that determine which ATAC peaks decrease or remain unchanged. The most likely hypothesis would be the presence of other bound TFs and further LOLA analysis comparing decreased vs unchanged/increased ATAC peaks may help elucidate this.

Several studies have suggested a role for eRNAs in maintaining gene expression through enhancer-promoter looping (Arnold *et al.* 2019). Significantly, our data indicates that at a given time point target gene transcription is impaired after MYB knockdown, eRNAs are reduced, but enhancer-promoter interaction frequency by Capture-C is not altered (Figure 5.13). It is possible that enhancer-promoter

looping interactions are relatively stable and require a longer treatment period for such changes to be observed, or that our ability to observe such changes is limited by the sensitivity and resolution of this technique. This may be partially addressed by performing Capture-C at a later time point and also to increase the number of Myb sensitive enhancer loci assayed. Under these conditions, if changes in enhancer-promoter interaction frequency are observed, it is possible that MYB may contribute to the maintenance of enhancer-promoter interactions, either through direct or indirect mechanisms. However, if no changes are observed even at a later time point, a role for MYB in this would be very unlikely. One caveat to this would be that there is data to suggest that enhancer-promoter interactions are broken and reformed when cells replicate, and as replication and cell cycle progression are processes that are inhibited by MYB KD, it is possible that the preservation of enhancer-promoter interactions is due to loss of cell division in response to loss of MYB. At the same time, while we observed a massive reduction in MYB protein levels and binding 24 hours following siRNA, it is important to note that MYB ChIP signal was not completely abolished. This raises the possibility that there may be a threshold requirement for MYB in the maintenance of different features of enhancer function. Specifically, where high levels of MYB binding are required for the maintenance of transcription, but low levels are still sufficient for the maintenance of enhancer-promoter interactions. To test this requires a system where MYB protein levels or binding may be complete abolished, such using inducible degron protein degradation technology or the TetR system.

By and large, these data and emerging evidence suggest that stabilisation of enhancer-promoter interactions is not sufficient for the maintenance of gene activation (El Khattabi *et al.* 2019). Rather, enhancer-promoter looping likely creates a permissive environment to facilitate contact of transcription activating factors, and hence transcriptional regulation between these two loci. More significantly, our data suggests that modulation of enhancer-promoter interactions is not a contributing factor in the immediate transcriptional regulation of MYB dependent genes. Importantly, we show that MYB dependent gene activation can

be disrupted with minimal changes in enhancer-promoter interaction frequency, suggesting that these are functionally and temporally separable events.

# 6
# Discussion and future work

## Contents


**6.1 General discussion . . . . . 132**
**6.2 An *in vivo* model for understanding disease resistance 133**
**6.3 A model for MYB function at enhancers . . . . . 133**
6.3.1 MYB function in establishing enhancer activity . . . . . 133
6.3.2 MYB function in maintaining enhancer activity . . . . . 135
**6.4 Future work . . . . . 136**


## 6.1 General discussion

MYB is a key transcription factor required for the development and maintenance of normal haematopoiesis. In MLLr leukaemia, *MYB* is a direct downstream target common to several MLL-FPs including MLL-AF9, the most common MLL-FP in acute myeloid leukaemia. MYB has been shown to be necessary for the maintenance of MLL-AF9 leukaemia, and even partial and transient suppression of MYB has been shown to abrogate MLL-AF9 leukaemia *in vivo* (Zuber *et al.* 2011). Taken together with the observation that normal haematopoiesis can tolerate a larger relative reduction in MYB levels compared to leukaemia cells (Zuber *et al.* 2011), these data suggest inhibition of MYB activity may be clinically tolerated, making it an attractive therapeutic target. This has particular clinical significance as despite risk-adapted treatment approaches, 5-year overall survival rates for MLLr AML remain poor - between 4 and 39% (Grimwade *et al.* 2010; Papaemmanuil *et al.* 2016; Estey 2018). Mortality is usually due to relapsed or refractory disease (H. Döhner *et al.* 2015). Therefore, understanding the mechanisms by which MYB regulates

transcription and the pathways that contribute to resistance to MYB inhibition will allow us to rationalise and design better therapeutic strategies to tackle this disease. Given the critical role of MYB in multiple cancer tissue types, such discoveries may have even broader clinical and scientific implications.

## 6.2 An *in vivo* model for understanding disease resistance

In Chapter 3, we engineered a transgenic mouse model to allow us to study the effect of *Myb* repression on leukaemogenesis (Figure 3.2). To our knowledge, this is the first *in vivo* model system with an endogenous, inducible circuit that can be used to fine tune gene expression. We observed that reduced *Myb* expression resulted in a significant attenuation of leukaemia fitness early in leukaemogenesis. However, the same cells were able to evolve a more aggressive phenotype over time. Thus, this transgenic mouse model can also serve as an excellent system for observing the evolution of alternative, Myb-independent pathways that mediate leukaemogenesis in an organism in real time. As explored in Chapter 3.4, future work should include dissecting the molecular pathways that contribute to this phenomenon.

## 6.3 A model for MYB function at enhancers

Our data strongly suggest that at a subset of MYB sensitive genes, MYB dependent enhancer function is necessary and may be sufficient for gene activation. Integrating data from this thesis with published evidence, we propose a working model for MYB function in 1) establishing and 2) maintaining enhancer activity (Figures 6.1 & 6.2).

### 6.3.1 MYB function in establishing enhancer activity

At gene regulatory elements, MYB binding at chromatin directs the localisation of CBP/P300, as well as stabilisation of Mediator and the RNAPII pre-initiation complex via its interaction with TAF12 (Xu *et al.* 2018). CBP/P300 acetylate surrounding lysine residues, including histone H3K27, which results in chromatin

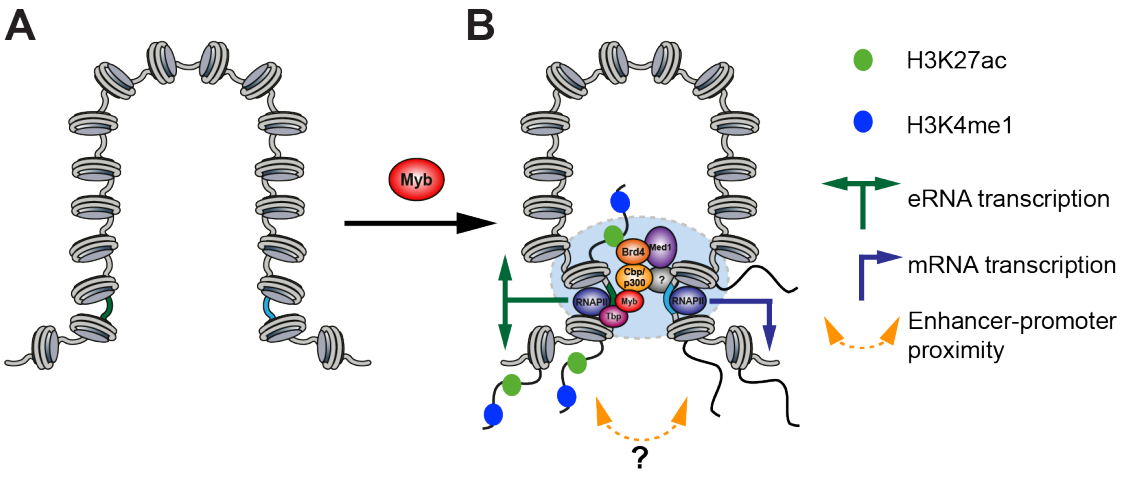



**Figure 6.1: Model for MYB function in establishing enhancer activity.** (A) Schematic of inactive chromatin. Green and blue highlighted DNA regions represent enhancer elements and promoter elements respectively. (B) MYB binding to enhancers stabilises its known interaction partners and co-activating complexes at these loci, resulting in recruitment of H3K27ac and H3K4me1, chromatin accessibility and active transcription at the enhancer and promoter.

relaxation and also acts as a binding substrate for co-activators such as BRD4 (Wu *et al.* 2007; Zeng *et al.* 2002; Filippakopoulos *et al.* 2012; Gilan *et al.* 2016). BRD4 is known to recruit P-TEFb, which is responsible for the pause release of RNAPII and productive transcription elongation (Marshall *et al.* 1995; Marshall *et al.* 1996; Yamaguchi *et al.* 1999; Wada *et al.* 1998). The subsequent transcriptional activity at enhancers produces eRNAs which interact with the regulatory region of the CBP/P300 HAT domain, enhancing its acetyltransferase activity and thus promoting chromatin accessibility, further transcription and gene activation (Bose *et al.* 2017a; Bose *et al.* 2017b). It is unclear how MYB binding results in increased H3K4me1 at enhancers, as we observed in our TetR-TA system. H3K4me1 is deposited at enhancers by MLL3/4 *in vivo* (Crump *et al.* 2019), suggesting that these enzymes may be localised here by MYB. Interaction between MYB and MLL3/4 has not been reported suggesting that this may be an indirect recruitment possibly following the establishment of a complex of enhancer-associated proteins and bears further investigation (Figure 6.1).

### 6.3.2 MYB function in maintaining enhancer activity

In the absence of other compensatory TFs or co-factors, loss of MYB binding at the enhancers of actively transcribing genes results in loss of enhancer function. This is likely through a direct reduction in CBP/P300 recruitment, leading to loss of H3K27ac, co-activating proteins and RNAPII destabilisation, but also possibly indirectly through loss of CBP/P300 activity in the absence of eRNA transcription. Importantly, MYB dependent gene activation can be disrupted with minimal changes in enhancer-promoter interaction frequency, suggesting that MYB acts as a functional rather than architectural bridge between enhancers and promoters to maintain gene activation. It is unclear what facilitates enhancer-promoter interactions in the absence of MYB, but likely candidates include other transcription factors or associated proteins, for example Ldb1 (Deng *et al.* 2012, cohesin (El Khattabi *et al.* 2019) or other structural factors such as ncRNA (Hsieh *et al.* 2014) (Figure 6.2).

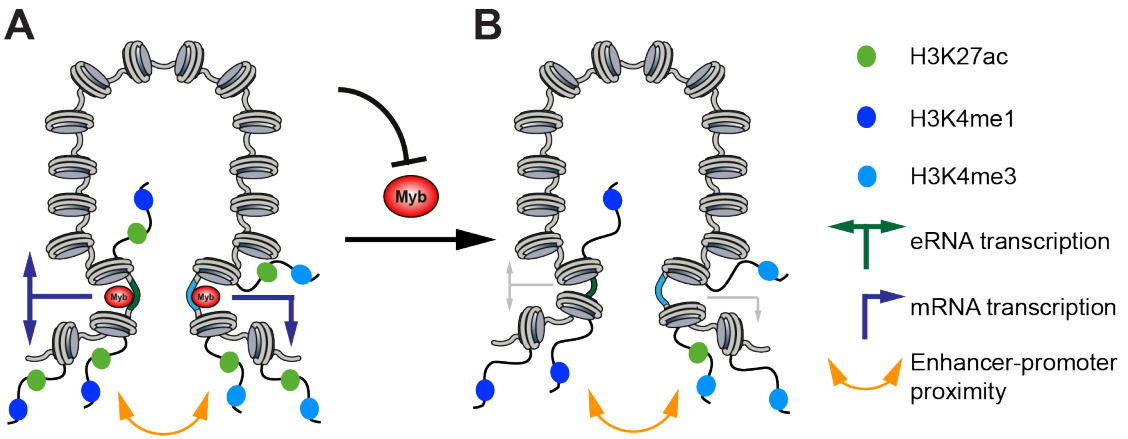



**Figure 6.2: Model for MYB function in maintaining enhancer activity.** Green and blue highlighted DNA regions represent enhancer elements and promoter elements respectively. (A) At MYB dependent enhancers, MYB binding is required for the maintenance of H3K27ac, chromatin accessibility and eRNA transcription (B) Loss of MYB binding leads to a reduction in H3K27ac, chromatin accessibility and eRNA transcription at the enhancer, with an associated reduction in transcription of the associated gene. Enhancer-promoter interactions are unperturbed.

## 6.4 Future work

Building on the above model, it would be important to identify which MYB TA domain interactions are required for establishing enhancer function. The MYB TA domain is known to interact directly with CBP/P300 and TAF12. The requirement for each of these interactions may be interrogated using the TetR system to dissect the molecular events leading to enhancer activity. The requirement for Cbp/p300 may be tested using the Myb TA domain mutants L302A and M303V, which disrupt its interaction with Cbp and p300 respectively (Kasper *et al.* 2013; Pattabiraman *et al.* 2009; Sandberg *et al.* 2005; Ramaswamy *et al.* 2018). Mutant Myb TA domains could be fused to the TetR and the presence of enhancer features at the *TetO* locus could be tested for by ChIP qPCR. Complementary to this, the core HAT domain of Cbp/p300 could also be fused to the TetR and similar experiments performed (Hilton *et al.* 2015). To test for the requirement for TAF12, TAF12 squelching (using the TAF4 histone fold domain as bait) (Xu *et al.* 2018) may be used to perturb its interaction with the Myb TA domain in TetR-TA cells and the consequences of this disruption observed at the *TetO* locus. Finally, by exploiting fact that the interaction between TetR and *TetO* is disrupted upon doxycycline treatment, a time-based dissection of what enhancer features are lost upon TetR-TA dissociation from chromatin can be performed.

Taking this a step further, our KD data in THP-1 cells shows that MYB is not required for the maintenance of enhancer-promoter interactions. What is yet unclear is if MYB is required for their establishment. If not, how does it activate transcription from distal loci as observed in TetR-TA cells? To address this, Capture-C could first be performed from *TetO* in TetR-TA cells to test for the presence of enhancer-promoter interactions. If enhancer-promoter interactions are observed, a complementary experiment using a Dox-on TetR system could be used to induce binding of TetR-TA to *TetO* in the presence of doxycycline and a time-course analysis of the appearance of specific enhancer features identified to eludicate the hierarchy of recruitment of enhancer features.

It is clear that there is still much work to be done to fully comprehend the functional role of MYB in leukaemia and gene regulation. This thesis has attempted to make some headway in addressing these questions. Ongoing and future work will hopefully offer greater insight into these mechanisms and add to the understanding of the molecular events underlying leukaemogenesis.

# 7
# Appendices

**Table 7.1:** Antibodies used for flow cytometry analysis and sorting

| Antibody | Conjugate | Dilution | Clone | Supplier | Catalogue numer | Application |
|---|---|---|---|---|---|---|
| B220 | PECy5 | 1/200 | RA3-6B2 | BioLegend | 103210 | Lineage staining |
| B220 | PECy7 | 1/200 | RA3-6B2 | BioLegend | 103221 | BM staining |
| CD105 | Bio-TexR | 1/100 | MJ7/18 | eBioscience | 13-1051-82 | BM staining |
| CD135 (Flt3) | Bio-TexR | 1/100 | A2F10 | eBioscience | 13-1351-82 | BM staining |
| CD150 | APC | 1/100 | TC15-12F12.2 | BioLegend | 115910 | BM staining |
| CD16/32 | PeCy7 | 1/150 | 93 | eBioscience | 25-0161-82 | BM staining |
| CD16/32 | Unconjucated | 1/50 | 93 | eBioscience | 14-0161-81 | Fc receptor block |
| CD19 | PECy5 | 1/800 | eBio1D3 | eBioscience | 15-0193-8 | Lineage staining |
| CD19 | PECy7 | 1/400 | eBio1D3 | eBioscience | 25-0193-82 | Peripheral blood staining |
| CD3 | PECy5 | 1/200 | | | | Lineage staining |
| CD34 | AF700 | 1/100 | RAM34 | eBioscience | 56-0341-82 | BM staining |
| CD34 | FITC | 1/25 | RAM34 | eBioscience | 56-0341-82 | BM staining |
| CD4 | PECy5 | 1/1600 | RM4-5 | BioLegend | 100514 | Lineage staining |
| CD4 | PE | 1/800 | GK1.5 | eBioscience | 12-0041-85 | Peripheral blood staining |
| CD4 | APC ef780 | 1/800 | GK1.5 | eBioscience | | Peripheral blood staining |
| CD41 | BV421 | 1/200 | MWRes30 | BioLegend | 133911 | BM staining |
| CD45.1 | AF700 | 1/100 | A20 | BioLegend | 110724 | Allotype discrimination |
| CD45.2 | PB | 1/100 | 104 | BioLegend | 109819 | Allotype discrimination |
| CD45.2 | FITC | 1/50 | 104 | eBioscience | 11-0454-85 | Allotype discrimination |
| CD48 | PE | 1/400 | | | | |
| CD49b | APC | 1/100 | | | | |
| CD5 | PECy5 | 1/1200 | 53-7.3 | BioLegend | 100610 | Lineage staining |
| CD8a | PE | 1/200 | 53-6.7 | eBioscience | 12-0081-85 | Peripheral blood staining |
| CD8 | ef780 | 1/800 | 53-6.7 | BioLegend | | Lineage staining |
| CD8a | PECy5 | 1/1600 | 53-6.7 | BioLegend | 100716 | Lineage staining |

**Table 7.1 continued from previous page**

| Antibody | Conjugate | Dilution | Clone | Supplier | Catalogue numer | Application |
|---|---|---|---|---|---|---|
| cKit | APC ef780 | 1/1200 | 2B8 | eBioscience | 47-1171-82 | BM staining |
| Gr1 | PECy5 | 1/400 | | | | Lineage staining |
| Gr1 | PO | 1/100 | | | | |
| Ly6D | PB | 1/400 | | | | |
| Mac1 | PECy5 | 1/400 | M1/70 | BioLegend | 101210 | Lineage staining |
| Mac1 | APC | 1/400 | M1/70 | eBioscience | 17-0112-83 | Peripheral blood staining |
| NK1.1 | APC | 1/400 | PK136 | BioLegend | | Peripheral blood staining |
| NK1.1 | BV650 | 1/50 | PK136 | BioLegend | 108735 | Peripheral blood staining |
| Sca1 | PECy7 | 1/200 | D7 | BD Horizon | | BM staining |
| Sca1 | BV786 | 1/200 | D7 | BD Horizon | 563991 | BM staining |
| Sca1 | PB | 1/150 | E13-161.7 | BioLegend | 122520 | BM staining |
| Streptavidin | BV421 | 1/200 | NA | BioLegend | 405226 | Secondary staining |
| Ter119 | PECy5 | 1/600 | TER-119 | BioLegend | 116210 | Lineage staining |
| CD127 | Bio-TexR | 1/100 | | | | |

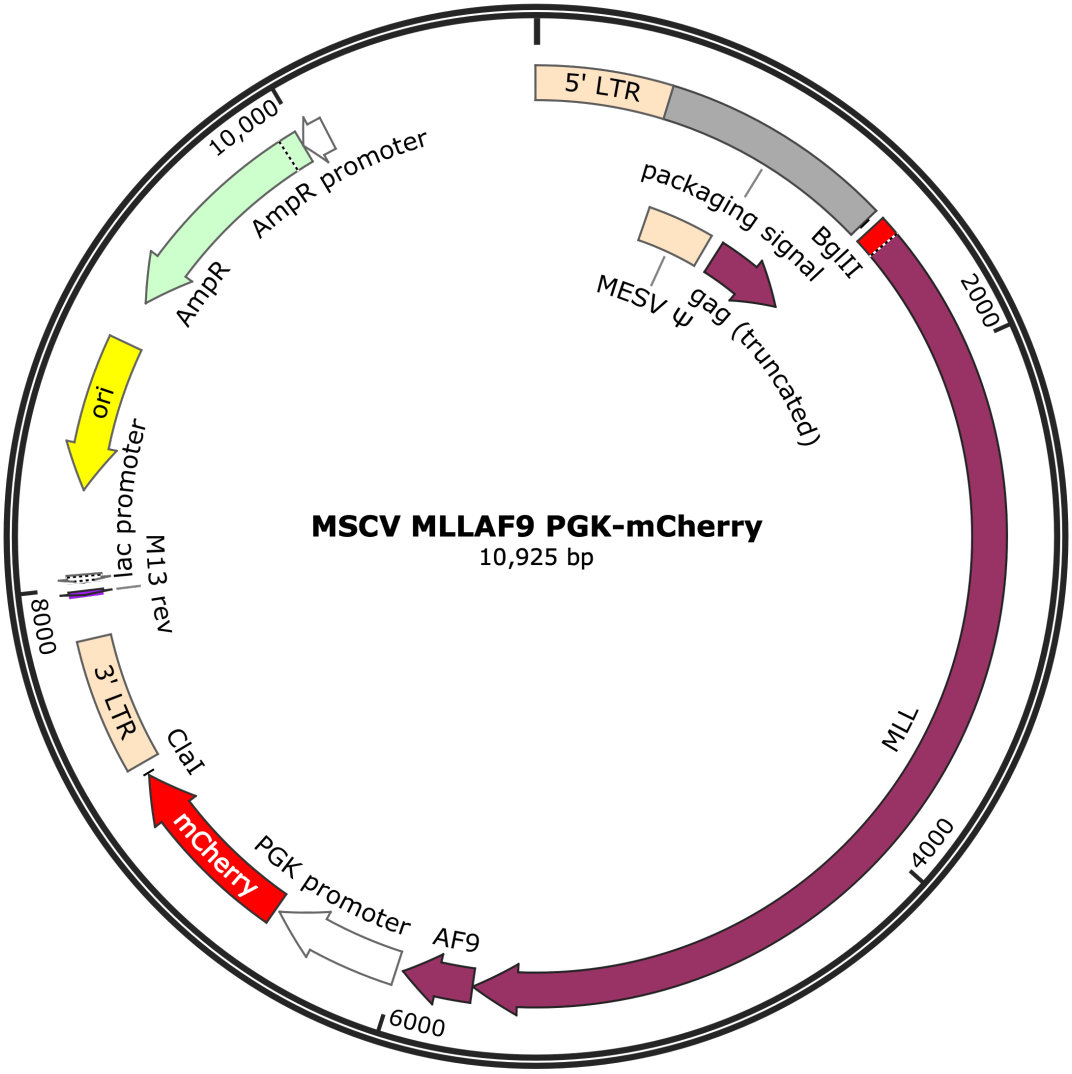



**Figure 7.1: MSCV MLL-AF9/mCherry plasmid map**

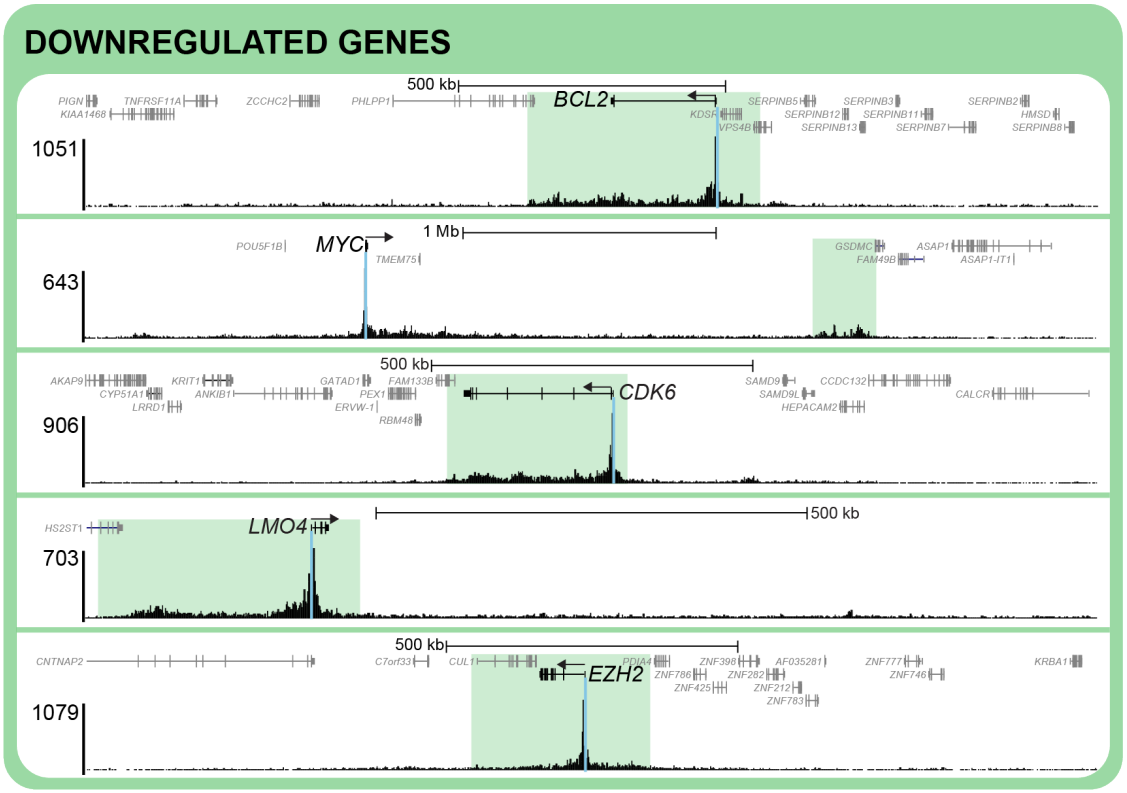



**Figure 7.2: Domain view of Capture-C traces at genes which are transcriptionally downregulated.** The interaction domain of each promoter is displayed. Vertical blue bar indicates the capture point.

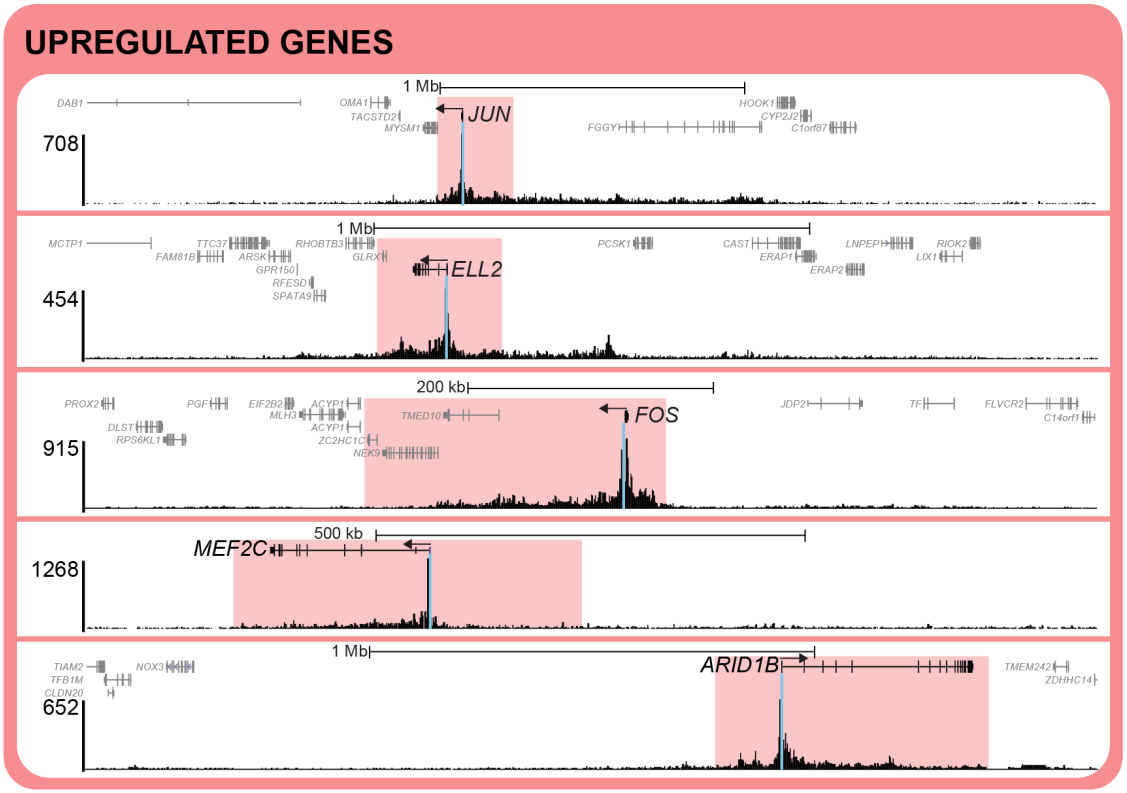



**Figure 7.3: Capture-C traces at genes which are transcriptionally upregulated.** The interaction domain of each promoter is displayed. Vertical blue bar indicates the capture point.

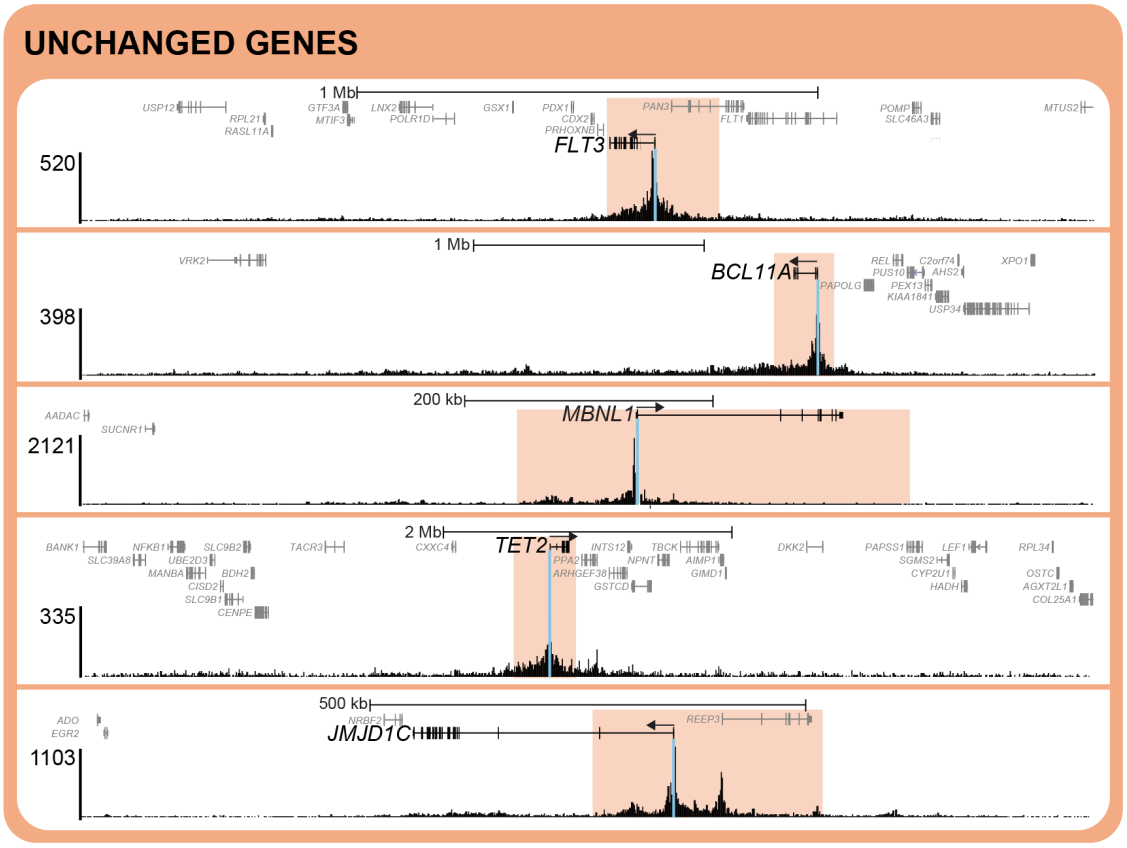



**Figure 7.4: Capture-C traces at genes which are transcriptionally unchanged.** The interaction domain of each promoter is displayed. Vertical blue bar indicates the capture point.

# References


Abraham, B. J., D. Hnisz, A. S. Weintraub, N. Kwiatkowski, C. H. Li, Z. Li, *et al.* (Feb. 2017). "Small genomic insertions form enhancers that misregulate oncogenes." *Nature Communications* 8, p. 14385.

Alexander, J. M., J. Guan, B. Li, L. Maliskova, M. Song, Y. Shen, *et al.* (May 2019). "Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity." *eLife* 8.

Allen, B. L. and D. J. Taatjes (Mar. 2015). "The Mediator complex: a central integrator of transcription." *Nature Reviews Molecular Cell Biology* 16.3, pp. 155–166.

Allen, R. D., T. P. Bender, and G. Siu (May 1999). "c-Myb is essential for early T cell development." *Genes & Development* 13.9, pp. 1073–1078.

Andersen, M. K., D. H. Christiansen, M. Kirchhoff, and J. Pedersen-Bjergaard (May 2001). "Duplication or amplification of chromosome band 11q23, including the unrearranged MLL gene, is a recurrent abnormality in therapy-related MDS and AML, and is closely related to mutation of the TP53 gene and to previous therapy with alkylating agents." *Genes, chromosomes & cancer* 31.1, pp. 33–41.

Andersson, A. K., J. Ma, J. Wang, X. Chen, A. L. Gedman, J. Dang, *et al.* (Apr. 2015). "The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias." *Nature Publishing Group* 47.4, pp. 330–337.

Andersson, K. B., E. Kowenz-Leutz, E. M. Brendeford, A.-H. H. Tygsett, A. Leutz, and O. S. Gabrielsen (Feb. 2003). "Phosphorylation-dependent down-regulation of c-Myb DNA binding is abrogated by a point mutation in the v-myb oncogene." *The Journal of biological chemistry* 278.6, pp. 3816–3824.

Andersson, R., C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt, M. Boyd, *et al.* (Mar. 2014a). "An atlas of active enhancers across human cell types and tissues." *Nature* 507.7493, pp. 455–461.

Andersson, R., P. Refsing Andersen, E. Valen, L. J. Core, J. Bornholdt, M. Boyd, *et al.* (2014b). "Nuclear stability and transcriptional directionality separate functionally distinct RNA species". *Nature Communications* 5.1, p. 5336.

Anfossi, G., A. M. Gewirtz, and B. Calabretta (May 1989). "An oligomer complementary to c-myb-encoded mRNA inhibits proliferation of human myeloid leukemia cell lines." *Proceedings of the National Academy of Sciences of the United States of America* 86.9, pp. 3379–3383.

Arnold, P. R., A. D. Wells, and X. C. Li (2019). "Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate." *Frontiers in cell and developmental biology* 7, p. 377.

Arnosti, D. N., S. Barolo, M. Levine, and S. Small (Jan. 1996). "The eve stripe 2 enhancer employs multiple modes of transcriptional synergy." *Development* 122.1, pp. 205–214.

Arnosti, D. N. and M. M. Kulkarni (Apr. 2005). "Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards?" *Journal of Cellular Biochemistry* 94.5, pp. 890–898.

Arvey, A., E. Larsson, C. Sander, C. S. Leslie, and D. S. Marks (Apr. 2010). "Target mRNA abundance dilutes microRNA and siRNA activity." *Molecular systems biology* 6, p. 363.

Ashe, H. L., J. Monks, M. Wijgerde, P. Fraser, and N. J. Proudfoot (Oct. 1997). "Intergenic transcription and transinduction of the human beta-globin locus." *Genes & Development* 11.19, pp. 2494–2509.

Aziz, N., M. R. Miglarese, R. C. Hendrickson, J. Shabanowitz, T. W. Sturgill, D. F. Hunt, *et al.* (July 1995). "Modulation of c-Myb-induced transcription activation by a phosphorylation site near the negative regulatory domain." *Proceedings of the National Academy of Sciences of the United States of America* 92.14, pp. 6429–6433.

Aziz, N., J. Wu, J. W. Dubendorff, J. S. Lipsick, T. W. Sturgill, and T. P. Bender (Aug. 1993). "c-Myb and v-Myb are differentially phosphorylated by p42mapk in vitro." *Oncogene* 8.8, pp. 2259–2265.

Bach, C., D. Mueller, S. Buhl, M. P. Garcia-Cuellar, and R. K. Slany (Feb. 2009). "Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2." *Oncogene* 28.6, pp. 815–823.

Badiani, P., P. Corbella, D. Kioussis, J. Marvel, and K. Weston (Apr. 1994). "Dominant interfering alleles define a role for c-Myb in T-cell development." *Genes & Development* 8.7, pp. 770–782.

Banerji, J., S. Rusconi, and W. Schaffner (Dec. 1981). "Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences." *Cell* 27.2 Pt 1, pp. 299–308.

Barabe, F., J. A. Kennedy, K. J. Hope, and J. E. Dick (Apr. 2007). "Modeling the initiation and progression of human acute leukemia in mice." *Science* 316.5824, pp. 600–604.

Bartel, D. P. (Jan. 2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." *Cell* 116.2, pp. 281–297.

— (Jan. 2009). "MicroRNAs: target recognition and regulatory functions." *Cell* 136.2, pp. 215–233.

— (Mar. 2018). "Metazoan MicroRNAs." *Cell* 173.1, pp. 20–51.

Behm-Ansmant, I., J. Rehwinkel, and E. Izaurralde (2006a). "MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay." *Cold Spring Harbor symposia on quantitative biology* 71, pp. 523–530.

Behm-Ansmant, I., J. Rehwinkel, T. Doerks, A. Stark, P. Bork, and E. Izaurralde (July 2006b). "mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes." *Genes & Development* 20.14, pp. 1885–1898.

Beltran, M., M. Tavares, N. Justin, G. Khandelwal, J. Ambrose, B. M. Foster, *et al.* (Oct. 2019). "G-tract RNA removes Polycomb repressive complex 2 from genes." *Nature structural & molecular biology* 26.10, pp. 899–909.

Beltran, M., C. M. Yates, L. Skalska, M. Dawson, F. P. Reis, K. Viiri, *et al.* (July 2016). "The interaction of PRC2 with RNA or chromatin is mutually antagonistic." *Genome Research* 26.7, pp. 896–907.

Benabdallah, N. S., I. Williamson, R. S. Illingworth, L. Kane, S. Boyle, D. Sengupta, *et al.* (Aug. 2019). "Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation." *Molecular cell*.

Bender, T. P., C. S. Kremer, M. Kraus, T. Buch, and K. Rajewsky (June 2004). "Critical functions for c-Myb at three checkpoints during thymocyte development." *Nature Immunology* 5.7, pp. 721–729.

Benoist, C. and P. Chambon (Mar. 1981). "In vivo sequence requirements of the SV40 early promotor region." *Nature* 290.5804, pp. 304–310.

Bernt, K. M., N. Zhu, A. U. Sinha, S. Vempati, J. Faber, A. V. Krivtsov, *et al.* (July 2011). "MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L." *Cancer Cell* 20.1, pp. 66–78.

Biddie, S. C., S. John, P. J. Sabo, R. E. Thurman, T. A. Johnson, R. L. Schiltz, *et al.* (July 2011). "Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding." *Molecular cell* 43.1, pp. 145–155.

Biedenkapp, H., U. Borgmeyer, A. E. Sippel, and K. H. Klempnauer (Oct. 1988). "Viral myb oncogene encodes a sequence-specific DNA-binding activity." *Nature* 335.6193, pp. 835–837.

Bies, J., J. Markus, and L. Wolff (Mar. 2002). "Covalent attachment of the SUMO-1 protein to the negative regulatory domain of the c-Myb transcription factor modifies its stability and transactivation capacity." *The Journal of biological chemistry* 277.11, pp. 8999–9009.

Bies, J., M. Sramko, and L. Wolff (Dec. 2013). "Stress-induced phosphorylation of Thr486 in c-Myb by p38 mitogen-activated protein kinases attenuates conjugation of SUMO-2/3." *The Journal of biological chemistry* 288.52, pp. 36983–36993.

Birney, E., J. A. Stamatoyannopoulos, A. Dutta, R. Guigó, T. R. Gingeras, E. H. Margulies, *et al.* (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project". *Nature* 447.7146, pp. 799–816.

Biroccio, A., B. Benassi, I. D'Agnano, C. D'Angelo, S. Buglioni, M. Mottolese, *et al.* (Apr. 2001). "c-Myb and Bcl-x Overexpression Predicts Poor Prognosis in Colorectal Cancer : Clinical and Experimental Findings". *The American Journal of Pathology* 158.4, p. 1289.

Biswas, D., T. A. Milne, V. Basrur, J. Kim, K. S. J. Elenitoba-Johnson, C. D. Allis, *et al.* (Sept. 2011). "Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes." *Proceedings of the National Academy of Sciences of the United States of America* 108.38, pp. 15751–15756.

Bitoun, E., P. L. Oliver, and K. E. Davies (Jan. 2007). "The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling." *Human molecular genetics* 16.1, pp. 92–106.

Blackledge, N. P., A. M. Farcas, T. Kondo, H. W. King, J. F. McGouran, L. L. P. Hanssen, *et al.* (June 2014). "Variant PRC1 complex-dependent H2A

ubiquitylation drives PRC2 recruitment and polycomb domain formation." *Cell* 157.6, pp. 1445–1459.

Boija, A., I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, *et al.* (Dec. 2018). "Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains." *Cell* 175.7, 1842–1855.e16.

Bose, D. A. and S. L. Berger (Dec. 2017a). "eRNA binding produces tailored CBP activity profiles to regulate gene expression." *RNA biology* 14.12, pp. 1655–1659.

Bose, D. A., G. Donahue, D. Reinberg, R. Shiekhattar, R. Bonasio, and S. L. Berger (Jan. 2017b). "RNA Binding to CBP Stimulates Histone Acetylation and Transcription." *Cell* 168.1-2, 135–149.e22.

Bosson, A. D., J. R. Zamudio, and P. A. Sharp (Nov. 2014). "Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition." *Molecular cell* 56.3, pp. 347–359.

Bousquet, M., G. Zhuang, C. Meng, W. Ying, P. S. Cheruku, A. T. Shie, *et al.* (Aug. 2013). "miR-150 blocks MLL-AF9-associated leukemia through oncogene repression." *Molecular cancer research : MCR* 11.8, pp. 912–922.

Boyle, W. J., M. A. Lampert, J. S. Lipsick, and M. A. Baluda (July 1984). "Avian myeloblastosis virus and E26 virus oncogene products are nuclear proteins." *Proceedings of the National Academy of Sciences of the United States of America* 81.14, pp. 4265–4269.

Brennecke, J., A. Stark, R. B. Russell, and S. M. Cohen (Mar. 2005). "Principles of microRNA-target recognition." *PLoS biology* 3.3, e85.

Brent, R. and M. Ptashne (Dec. 1985). "A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor." *Cell* 43.3 Pt 2, pp. 729–736.

Broeker, P. L., A. Harden, J. D. Rowley, and N. Zeleznik-Le (1996). "The mixed lineage leukemia (MLL) protein involved in 11q23 translocations contains a domain that binds cruciform DNA and scaffold attachment region (SAR) DNA." *Current topics in microbiology and immunology* 211, pp. 259–268.

Brown, J. M., N. A. Roberts, B. Graham, D. Waithe, C. Lagerholm, J. M. Telenius, *et al.* (Sept. 2018). "A tissue-specific self-interacting chromatin domain forms independently of enhancer-promoter interactions". *Nature Communications* 9.1, p. 3849.

Buechele, C., E. H. Breese, D. Schneidawind, C.-H. Lin, J. Jeong, J. Duque-Afonso, *et al.* (Oct. 2015). "MLL leukemia induction by genome editing of human CD34+ hematopoietic cells." *Blood* 126.14, pp. 1683–1694.

Buecker, C. and J. Wysocka (June 2012). "Enhancers as information integration hubs in development: lessons from genomics." *Trends in Genetics* 28.6, pp. 276–284.

Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang, and W. J. Greenleaf (Oct. 2013). "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position". *Nature Methods* 10.12, pp. 1213–1218.

Buono, M., R. Facchini, S. Matsuoka, S. Thongjuea, D. Waithe, T. C. Luis, *et al.* (Feb. 2016). "A dynamic niche provides Kit ligand in a stage-specific manner to the earliest thymocyte progenitors." *Nature cell biology* 18.2, pp. 157–167.

Cabezas-Wallscheid, N., D. Klimmeck, J. Hansson, D. B. Lipka, A. Reyes, Q. Wang, *et al.* (Oct. 2014). "Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis." *Cell Stem Cell* 15.4, pp. 507–522.

Calabretta, B., R. B. Sims, M. Valtieri, D. Caracciolo, C. Szczylik, D. Venturelli, *et al.* (Mar. 1991). "Normal and leukemic hematopoietic cells manifest differential sensitivity to inhibitory effects of c-myb antisense oligodeoxynucleotides: an in vitro study relevant to bone marrow purging." *Proceedings of the National Academy of Sciences of the United States of America* 88.6, pp. 2351–2355.

Campanini, E. B., M. W. Vandewege, N. E. Pillai, B.-H. Tay, J. L. Jones, B. Venkatesh, *et al.* (Nov. 2015). "Early Evolution of Vertebrate Mybs: An Integrative Perspective Combining Synteny, Phylogenetic, and Gene Expression Analyses." *Genome biology and evolution* 7.11, pp. 3009–3021.

Cancer Genome Atlas Research Network, T. J. Ley, C. Miller, L. Ding, B. J. Raphael, A. J. Mungall, *et al.* (May 2013). "Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia." *New England Journal of Medicine* 368.22, pp. 2059–2074.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, *et al.* (2005). "The Transcriptional Landscape of the Mammalian Genome". *Science* 309.5740, pp. 1559–1563.

Carpinelli, M. R., D. J. Hilton, D. Metcalf, J. L. Antonchuk, C. D. Hyland, S. L. Mifsud, *et al.* (Apr. 2004). "Suppressor screen in Mpl-/- mice: c-Myb mutation causes supraphysiological production of platelets in the absence of thrombopoietin signaling." *Proceedings of the National Academy of Sciences of the United States of America* 101.17, pp. 6553–6558.

Caslini, C., Z. Yang, M. El-Osta, T. A. Milne, R. K. Slany, and J. L. Hess (Aug. 2007). "Interaction of MLL amino terminal sequences with menin is required for transformation." *Cancer research* 67.15, pp. 7275–7283.

Chandradoss, S. D., N. T. Schirle, M. Szczepaniak, I. J. MacRae, and C. Joo (July 2015). "A Dynamic Search Process Underlies MicroRNA Targeting." *Cell* 162.1, pp. 96–107.

Chen, H., M. Levo, L. Barinov, M. Fujioka, J. B. Jaynes, and T. Gregor (Sept. 2018). "Dynamic interplay between enhancer-promoter topology and gene activity." *Nature Genetics* 50.9, pp. 1296–1303.

Chen, W., A. R. Kumar, W. A. Hudson, Q. Li, B. Wu, R. A. Staggs, *et al.* (May 2008). "Malignant transformation initiated by Mll-AF9: gene dosage and critical target cells." *Cancer Cell* 13.5, pp. 432–440.

Chen, W., M. G. O'Sullivan, W. Hudson, and J. Kersey (Mar. 2011). "Modeling human infant MLL leukemia in mice: leukemia from fetal liver differs from that originating in postnatal marrow." *Blood* 117.12, pp. 3474–3475.

Chen, Y.-X., J. Yan, K. Keeshan, A. T. Tubbs, H. Wang, A. Silva, *et al.* (Jan. 2006). "The tumor suppressor menin regulates hematopoiesis and myeloid transformation by influencing Hox gene expression." *Proceedings of the National Academy of Sciences of the United States of America* 103.4, pp. 1018–1023.

Cho, W.-K., J.-H. Spille, M. Hecht, C. Lee, C. Li, V. Grube, *et al.* (July 2018). "Mediator and RNA polymerase II clusters associate in transcription-dependent condensates." *Science* 361.6400, pp. 412–415.

Chu, C.-y. and T. M. Rana (July 2006). "Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54." *PLoS biology* 4.7, e210.

Cimino, G., D. T. Moir, O. Canaani, K. Williams, W. M. Crist, S. Katzav, *et al.* (Dec. 1991). "Cloning of ALL-1, the locus involved in leukemias with the t(4;11)(q21;q23), t(9;11)(p22;q23), and t(11;19)(q23;p13) chromosome translocations." *Cancer research* 51.24, pp. 6712–6714.

Cirillo, L. A., F. R. Lin, I. Cuesta, D. Friedman, M. Jarnik, and K. S. Zaret (Feb. 2002). "Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4." *Molecular cell* 9.2, pp. 279–289.

Clappier, E., W. Cuccuini, A. Kalota, A. Crinquette, J.-M. Cayuela, W. A. Dik, *et al.* (Aug. 2007). "The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children." *Blood* 110.4, pp. 1251–1261.

Collins, E. C., R. Pannell, E. M. Simpson, A. Forster, and T. H. Rabbitts (Aug. 2000). "Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development." *EMBO reports* 1.2, pp. 127–132.

Corces, M. R., J. D. Buenrostro, B. Wu, P. G. Greenside, S. M. Chan, J. L. Koenig, *et al.* (Oct. 2016). "Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution." *Nature Genetics* 48.10, pp. 1193–1203.

Corral, J., I. Lavenir, H. Impey, A. J. Warren, A. Forster, T. A. Larson, *et al.* (June 1996). "An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes." *Cell* 85.6, pp. 853–861.

Creyghton, M. P., A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, *et al.* (Dec. 2010). "Histone H3K27ac separates active from poised enhancers and predicts developmental state." *Proceedings of the National Academy of Sciences of the United States of America* 107.50, pp. 21931–21936.

Crump, N. T. and T. A. Milne (Aug. 2019). "Why are so many MLL lysine methyltransferases required for normal mammalian development?" *Cellular and molecular life sciences : CMLS* 76.15, pp. 2885–2898.

Dai, P., H. Akimaru, Y. Tanaka, D. X. Hou, T. Yasukawa, C. Kanei-Ishii, *et al.* (Mar. 1996). "CBP as a transcriptional coactivator of c-Myb." *Genes & Development* 10.5, pp. 528–540.

Dash, A. B., F. C. Orrico, and S. A. Ness (Aug. 1996). "The EVES motif mediates both intermolecular and intramolecular regulation of c-Myb." *Genes & Development* 10.15, pp. 1858–1869.

Davies, J. O. J., J. M. Telenius, S. J. McGowan, N. A. Roberts, S. Taylor, D. R. Higgs, *et al.* (Jan. 2016). "Multiplexed analysis of chromosome conformation at vastly improved sensitivity." *Nature Methods* 13.1, pp. 74–80.

De Santa, F., I. Barozzi, F. Mietton, S. Ghisletti, S. Polletti, B. K. Tusi, *et al.* (May 2010). "A large fraction of extragenic RNA pol II transcription sites overlap enhancers." *PLoS biology* 8.5, e1000384.

Debernardi, S., S. Skoulakis, G. Molloy, T. Chaplin, A. Dixon-McIver, and B. D. Young (May 2007). "MicroRNA miR-181a correlates with morphological sub-class of acute myeloid leukaemia and the expression of its target genes in global genome-wide analysis." *Leukemia* 21.5, pp. 912–916.

Deng, W., J. Lee, H. Wang, J. Miller, A. Reik, P. D. Gregory, *et al.* (June 2012). "Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor." *Cell* 149.6, pp. 1233–1244.

Denny, S. K., D. Yang, C.-H. Chuang, J. J. Brady, J. S. Lim, B. M. Gruner, *et al.* (July 2016). "Nfib Promotes Metastasis through a Widespread Increase in Chromatin Accessibility." *Cell* 166.2, pp. 328–342.

Di Stefano, B., S. Collombet, J. S. Jakobsen, M. Wierer, J. L. Sardina, A. Lackner, *et al.* (Apr. 2016). "C/EBPalpha creates elite cells for iPSC reprogramming by upregulating Klf4 and increasing the levels of Lsd1 and Brd4." *Nature cell biology* 18.4, pp. 371–381.

Dixon, J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, *et al.* (Apr. 2012). "Topological domains in mammalian genomes identified by analysis of chromatin interactions." *Nature* 485.7398, pp. 376–380.

Djabali, M., L. Selleri, P. PARRY, M. Bower, B. D. Young, and G. A. EVANS (Oct. 1992). "A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias." *Nature Genetics* 2.2, pp. 113–118.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, *et al.* (Jan. 2013). "STAR: ultrafast universal RNA-seq aligner." *Bioinformatics* 29.1, pp. 15–21.

Dobson, C. L., A. J. Warren, R. Pannell, A. Forster, I. Lavenir, J. Corral, *et al.* (July 1999). "The mll-AF9 gene fusion in mice controls myeloproliferation and specifies acute myeloid leukaemogenesis." *Embo Journal* 18.13, pp. 3564–3574.

Dogan, N., W. Wu, C. S. Morrissey, K.-B. Chen, A. Stonestrom, M. Long, *et al.* (2015). "Occupancy by key transcription factors is a more accurate predictor of enhancer activity than histone modifications or chromatin accessibility." *Epigenetics & chromatin* 8, p. 16.

Döhner, H., E. Estey, D. Grimwade, S. Amadori, F. R. Appelbaum, T. Büchner, *et al.* (Jan. 2017). "Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel." *Blood* 129.4, pp. 424–447.

Döhner, H., D. J. Weisdorf, and C. D. Bloomfield (Sept. 2015). "Acute Myeloid Leukemia." *New England Journal of Medicine* 373.12, pp. 1136–1152.

Döhner, K., K. Tobis, R. Ulrich, S. Fröhling, A. Benner, R. F. Schlenk, *et al.* (Aug. 2002). "Prognostic Significance of Partial Tandem Duplications of the MLLGene in Adult Patients 16 to 60 Years Old With Acute Myeloid Leukemia and Normal Cytogenetics: A Study of the Acute Myeloid Leukemia Study Group Ulm". *Journal of Clinical Oncology* 20.15, pp. 3254–3261.

Driessen, E. M. C., E. H. J. van Roon, J. A. P. Spijkers-Hagelstein, P. Schneider, P. De Lorenzo, M. G. Valsecchi, *et al.* (June 2013). "Frequencies and prognostic impact of RAS mutations in MLL-rearranged acute lymphoblastic leukemia in infants." *Haematologica* 98.6, pp. 937–944.

Drynan, L. F., R. Pannell, A. Forster, N. M. M. Chan, F. Cano, A. Daser, *et al.* (Sept. 2005). "Mll fusions generated by Cre-loxP-mediated de novo translocations can induce lineage reassignment in tumorigenesis." *Embo Journal* 24.17, pp. 3136–3146.

Eckner, R. (Nov. 1996). "p300 and CBP as transcriptional regulators and targets of oncogenic events." *Biological chemistry* 377.11, pp. 685–688.

El Khattabi, L., H. Zhao, J. Kalchschmidt, N. Young, S. Jung, P. Van Blerkom, *et al.* (Aug. 2019). "A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers." *Cell* 178.5, 1145–1158.e20.

Emambokus, N., A. Vegiopoulos, B. Harman, E. Jenkinson, G. Anderson, and J. Frampton (Sept. 2003). "Progression through key stages of haemopoiesis is dependent on distinct threshold levels of c-Myb." *Embo Journal* 22.17, pp. 4478–4488.

Emerenciano, M., T. d. C. Barbosa, B. de Almeida Lopes, C. Meyer, R. Marschalek, and M. S. Pombo-de-Oliveira (July 2015). "Subclonality and prenatal origin of RAS mutations in KMT2A (MLL)-rearranged infant acute lymphoblastic leukaemia." *British Journal of Haematology* 170.2, pp. 268–271.

Erson-Bensan, A. E. and T. Can (June 2016). "Alternative Polyadenylation: Another Foe in Cancer." *Molecular cancer research : MCR* 14.6, pp. 507–517.

Estey, E. H. (Oct. 2018). "Acute myeloid leukemia: 2019 update on risk-stratification and management." *American Journal of Hematology* 93.10, pp. 1267–1291.

Fahl, S. P., R. B. Crittenden, D. Allman, and T. P. Bender (Oct. 2009). "c-Myb Is Required for Pro-B Cell Differentiation". *The Journal of Immunology* 183.9, pp. 5582–5592.

Fan, Y., T. Nikitina, J. Zhao, T. J. Fleury, R. Bhattacharyya, E. E. Bouhassira, *et al.* (Dec. 2005). "Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation." *Cell* 123.7, pp. 1199–1212.

Favier, D. and T. J. Gonda (Jan. 1994). "Detection of proteins that bind to the leucine zipper motif of c-Myb." *Oncogene* 9.1, pp. 305–311.

Feng, Q., H. Wang, H. H. Ng, H. Erdjument-Bromage, P. Tempst, K. Struhl, *et al.* (June 2002). "Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain." *Current biology : CB* 12.12, pp. 1052–1058.

Filippakopoulos, P., S. Picaud, M. Mangos, T. Keates, J.-P. Lambert, D. Barsyte-Lovejoy, *et al.* (Mar. 2012). "Histone recognition and large-scale structural analysis of the human bromodomain family." *Cell* 149.1, pp. 214–231.

Fleischmann, K. K., P. Pagel, J. von Frowein, T. Magg, A. A. Roscher, and I. Schmid (Jan. 2016). "The leukemogenic fusion gene MLL-AF9 alters microRNA expression pattern and inhibits monoblastic differentiation via miR-511 repression." *Journal of experimental & clinical cancer research : CR* 35, p. 9.

Forster, A., R. Pannell, L. DRYNAN, F. Cano, N. CHAN, R. CODRINGTON, *et al.* (2005). "Chromosomal translocation engineering to recapitulate primary events of human cancer." *Cold Spring Harbor symposia on quantitative biology* 70, pp. 275–282.

Frampton, J., T. J. Gibson, S. A. Ness, G. Döderlein, and T. Graf (Dec. 1991). "Proposed structure for the DNA-binding domain of the Myb oncoprotein based on model building and mutational analysis." *Protein engineering* 4.8, pp. 891–901.

Friedman, R. C., K. K.-H. Farh, C. B. Burge, and D. P. Bartel (Jan. 2009). "Most mammalian mRNAs are conserved targets of microRNAs." *Genome Research* 19.1, pp. 92–105.

Fu, S. L. and J. S. Lipsick (Aug. 1996). "FAETL motif required for leukemic transformation by v-Myb." *Journal of virology* 70.8, pp. 5600–5610.

Fudenberg, G., M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny (May 2016). "Formation of Chromosomal Domains by Loop Extrusion." *CELREP* 15.9, pp. 2038–2049.

Fuglerud, B. M., R. B. Lemma, P. Wanichawan, A. Y. M. Sundaram, R. Eskeland, and O. S. Gabrielsen (July 2017). "A c-Myb mutant causes deregulated differentiation due to impaired histone binding and abrogated pioneer factor function." *Nucleic Acids Research* 45.13, pp. 7681–7696.

Fulton, D. L., S. Sundararajan, G. Badis, T. R. Hughes, W. W. Wasserman, J. C. Roach, *et al.* (2009). "TFCat: the curated catalog of mouse and human transcription factors." *Genome Biology* 10.3, R29.

Furlong, E. E. M. and M. Levine (Sept. 2018). "Developmental enhancers and chromosome topology". *Science* 361.6409, pp. 1341–1345.

Garcia, P., M. Clarke, A. Vegiopoulos, O. Berlanga, A. Camelo, M. Lorvellec, *et al.* (May 2009). "Reduced c-Myb activity compromises HSCs and leads to a myeloproliferation with a novel stem cell basis." *The EMBO Journal* 28.10, pp. 1492–1504.

Gilan, O., E. Y. N. Lam, I. Becher, D. Lugo, E. Cannizzaro, G. Joberty, *et al.* (July 2016). "Functional interdependence of BRD4 and DOT1L in MLL leukemia." *Nature structural & molecular biology* 23.7, pp. 673–681.

Godfrey, L., N. T. Crump, R. Thorne, I.-J. Lau, E. Repapi, D. Dimou, *et al.* (June 2019). "DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation." *Nature Communications* 10.1, p. 2803.

Godfrey, L., J. Kerry, R. Thorne, E. Repapi, J. O. J. Davies, M. Tapia, *et al.* (Nov. 2016). "MLL-AF4 binds directly to a BCL-2 specific enhancer and modulates H3K27 acetylation." *Experimental Hematology*.

Golay, J., E. Erba, S. Bernasconi, G. Peri, and M. Introna (July 1994). "The A-myb gene is preferentially expressed in tonsillar CD38+, CD39-, and sIgM- B lymphocytes and in Burkitt's lymphoma cell lines." *Journal of immunology (Baltimore, Md. : 1950)* 153.2, pp. 543–553.

Gomez-Diaz, E. and V. G. Corces (Nov. 2014). "Architectural proteins: regulators of 3D genome organization in cell fate." *Trends in cell biology* 24.11, pp. 703–711.

Gonda, T. J., D. Favier, P. Ferrao, E. M. Macmillan, R. Simpson, and F. Tavner (1996). "The c-myb negative regulatory domain." *Current topics in microbiology and immunology* 211, pp. 99–109.

Göran Stenman, M. K. A. Y. A. (Aug. 2010). "New tricks from an old oncogene: Gene fusion and copy number alterations of MYB in human cancer". *Cell Cycle* 9.15, p. 2986.

Grasser, F. A., T. Graf, and J. S. Lipsick (Aug. 1991). "Protein truncation is required for the activation of the c-myb proto-oncogene." *Molecular and Cellular Biology* 11.8, pp. 3987–3996.

Grimwade, D., R. K. Hills, A. V. Moorman, H. Walker, S. Chatters, A. H. Goldstone, *et al.* (July 2010). "Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials." *Blood* 116.3, pp. 354–365.

Gu, Y., T. Nakamura, H. Alder, R. Prasad, O. Canaani, G. Cimino, *et al.* (Nov. 1992). "The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene." *Cell* 71.4, pp. 701–708.

Guarente, L., J. S. Nye, A. Hochschild, and M. Ptashne (Apr. 1982). "Mutant lambda phage repressor with a specific defect in its positive control function." *Proceedings of the National Academy of Sciences of the United States of America* 79.7, pp. 2236–2239.

Guertin, M. J. and J. T. Lis (Sept. 2010). "Chromatin landscape dictates HSF binding to target DNA elements." *PLoS genetics* 6.9, e1001114.

Hall, W. J., C. W. Bean, and M. Pollard (July 1941). "Hall: Transmission of fowl leucosis through chick". *Am. J. Vet. Res* 2, pp. 272–279.

Hansen, A. S., T.-H. S. Hsieh, C. Cattoglio, I. Pustova, R. Saldana-Meyer, D. Reinberg, *et al.* (Sept. 2019). "Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF." *Molecular cell*.

Hedge, S. P., A. Kumar, C. Kurschner, and L. H. Shapiro (May 1998). "c-Maf interacts with c-Myb to regulate transcription of an early myeloid gene during differentiation." *Molecular and Cellular Biology* 18.5, pp. 2729–2737.

Hegyi, H. (Mar. 2015). "Enhancer-promoter interaction facilitated by transiently forming G-quadruplexes." *Scientific reports* 5, p. 9165.

Heintzman, N. D., G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark, L. F. Harp, *et al.* (May 2009). "Histone modifications at human enhancers reflect global cell-type-specific gene expression." *Nature* 459.7243, pp. 108–112.

Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, *et al.* (Mar. 2007). "Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome." *Nature Genetics* 39.3, pp. 311–318.

Heinz, S., C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, *et al.* (May 2010). "Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities". *Molecular cell* 38.4, pp. 576–589.

Hess, J. L., C. B. Bittner, D. T. Zeisig, C. Bach, and U. Fuchs (July 2006). "c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells". *Blood* 108.1, pp. 297–304.

Hilton, I. B., A. M. D'Ippolito, C. M. Vockley, P. I. Thakore, G. E. Crawford, T. E. Reddy, *et al.* (May 2015). "Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers." *Nature Biotechnology* 33.5, pp. 510–517.

Hnisz, D., B. J. Abraham, T. I. Lee, A. Lau, V. Saint-Andre, A. A. Sigova, *et al.* (Nov. 2013). "Super-enhancers in the control of cell identity and disease." *Cell* 155.4, pp. 934–947.

Ho, Y., F. Elefant, S. A. Liebhaber, and N. E. Cooke (Aug. 2006). "Locus control region transcription plays an active role in long-range gene activation." *Molecular cell* 23.3, pp. 365–375.

Hochschild, A., N. Irwin, and M. Ptashne (Feb. 1983). "Repressor structure and the mechanism of positive control." *Cell* 32.2, pp. 319–325.

Holmqvist, P.-H., S. Belikov, K. S. Zaret, and O. Wrange (Apr. 2005). "FoxA1 binding to the MMTV LTR modulates chromatin structure and transcription." *Experimental cell research* 304.2, pp. 593–603.

Hou, Y., F. Li, R. Zhang, S. Li, H. Liu, Z. S. Qin, *et al.* (Sept. 2019). "Integrative characterization of G-Quadruplexes in the three-dimensional chromatin structure." *Epigenetics* 14.9, pp. 894–911.

Howe, K. M. and R. J. Watson (July 1991). "Nucleotide preferences in sequence-specific recognition of DNA by c-myb protein." *Nucleic Acids Research* 19.14, pp. 3913–3919.

Hsieh, C.-L., T. Fei, Y. Chen, T. Li, Y. Gao, X. Wang, *et al.* (May 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* 111.20, pp. 7319–7324.

Hu, Y. L., R. G. Ramsay, C. Kanei-Ishii, S. Ishii, and T. J. Gonda (Sept. 1991). "Transformation by carboxyl-deleted Myb reflects increased transactivating

capacity and disruption of a negative regulatory domain." *Oncogene* 6.9, pp. 1549–1553.

Hughes, C. M., O. Rozenblatt-Rosen, T. A. Milne, T. D. Copeland, S. S. Levine, J. C. Lee, *et al.* (Feb. 2004). "Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus." *Molecular cell* 13.4, pp. 587–597.

Hume, M. A., L. A. Barrera, S. S. Gisselbrecht, and M. L. Bulyk (Jan. 2015). "UniPROBE, update 2015: new tools and content for the online database of protein-binding microarray data on protein-DNA interactions." *Nucleic Acids Research* 43.Database issue, pp. D117–22.

Huret, J. L., P. Dessen, and A. Bernheim (Jan. 2001). "Atlas of Genetics and Cytogenetics in Oncology and Haematology, updated." *Nucleic Acids Research* 29.1, pp. 303–304.

Hutvagner, G. and P. D. Zamore (Sept. 2002). "A microRNA in a multiple-turnover RNAi enzyme complex." *Science* 297.5589, pp. 2056–2060.

Hyle, J., Y. Zhang, S. Wright, B. Xu, Y. Shao, J. Easton, *et al.* (July 2019). "Acute depletion of CTCF directly affects MYC regulation through loss of enhancer-promoter looping." *Nucleic Acids Research* 47.13, pp. 6699–6713.

Inoue, F., M. Kircher, B. Martin, G. M. Cooper, D. M. Witten, M. T. McManus, *et al.* (Jan. 2017). "A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity". *Genome Research* 27.1, pp. 38–52.

Ivanov, X., Z. Mladenov, S. T. Nedyalkov, T. G. Todorov, and M. Yakimov (Oct. 1964). "Experimental investigations into avian leucoses. V. Transmission, haematology and morphology of avian myelocytomatosis". *Bull. Inst. Pathol. Comp. Animaux Domest. (Sofia)* 10, pp. 5–38.

Iwafuchi-Doi, M., G. Donahue, A. Kakumanu, J. A. Watts, S. Mahony, B. F. Pugh, *et al.* (Apr. 2016a). "The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation." *Molecular cell* 62.1, pp. 79–91.

Iwafuchi-Doi, M. and K. S. Zaret (June 2016b). "Cell fate control by pioneer transcription factors." *Development* 143.11, pp. 1833–1837.

Jamin, N., O. S. Gabrielsen, N. Gilles, P. N. Lirsac, and F. Toma (Aug. 1993). "Secondary structure of the DNA-binding domain of the c-Myb oncoprotein in solution. A multidimensional double and triple heteronuclear NMR study." *European journal of biochemistry* 216.1, pp. 147–154.

Jeong, J., A. Jager, P. Domizi, M. Pavel-Dinu, L. Gojenola, M. Iwasaki, *et al.* (Oct. 2019). "High-efficiency CRISPR induction of t(9;11) chromosomal translocations and acute leukemias in human blood stem cells." *Blood advances* 3.19, pp. 2825–2835.

Jin, S., H. Zhao, Y. Yi, Y. Nakata, A. Kalota, and A. M. Gewirtz (Feb. 2010). "c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis." *Journal of Clinical Investigation* 120.2, pp. 593–606.

Jolma, A., J. Yan, T. Whitington, J. Toivonen, K. R. Nitta, P. Rastas, *et al.* (Jan. 2013). "DNA-binding specificities of human transcription factors." *Cell* 152.1-2, pp. 327–339.

Jongen-Lavrencic, M., S. M. Sun, M. K. Dijkstra, P. J. M. Valk, and B. Löwenberg (May 2008). "MicroRNA expression profiling in relation to the genetic heterogeneity of acute myeloid leukemia." *Blood* 111.10, pp. 5078–5085.

Kalkbrenner, F., S. Guehmann, and K. Moelling (May 1990). "Transcriptional activation by human c-myb and v-myb genes." *Oncogene* 5.5, pp. 657–661.

Kanei-Ishii, C., E. M. Macmillan, T. Nomura, A. Sarai, R. G. Ramsay, S. Aimoto, *et al.* (Apr. 1992). "Transactivation and transformation by Myb are negatively regulated by a leucine-zipper structure." *Proceedings of the National Academy of Sciences of the United States of America* 89.7, pp. 3088–3092.

Kanei-Ishii, C., J. Tanikawa, A. Nakai, R. I. Morimoto, and S. Ishii (July 1997). "Activation of heat shock transcription factor 3 by c-Myb in the absence of cellular stress." *Science* 277.5323, pp. 246–248.

Kanhere, A., K. Viiri, C. C. Araújo, J. Rasaiyaah, R. D. Bouwman, W. A. Whyte, *et al.* (June 2010). "Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2." *Molecular cell* 38.5, pp. 675–688.

Kapranov, P., J. Cheng, S. Dike, D. A. Nix, R. Duttagupta, A. T. Willingham, *et al.* (June 2007). "RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription". *Science* 316.5830, p. 1484.

Kasper, L. H., F. Boussouar, P. A. Ney, C. W. Jackson, J. Rehg, J. M. van Deursen, *et al.* (Oct. 2002). "A transcription-factor-binding surface of coactivator p300 is required for haematopoiesis." *Nature* 419.6908, pp. 738–743.

Kasper, L. H., T. Fukuyama, S. Lerach, Y. Chang, W. Xu, S. Wu, *et al.* (2013). "Genetic interaction between mutations in c-Myb and the KIX domains of CBP and p300 affects multiple blood cell lineages and influences both gene activation and repression." *PLoS ONE* 8.12, e82684.

Keegan, L., G. Gill, and M. Ptashne (Feb. 1986). "Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein." *Science* 231.4739, pp. 699–704.

Kiewitz, A. and H. Wolfes (Oct. 1997). "Mapping of protein-protein interactions between c-myb and its coactivator CBP by a new phage display technique." *FEBS Letters* 415.3, pp. 258–262.

Kim, J., A. J. Woo, J. Chu, J. W. Snow, Y. Fujiwara, C. G. Kim, *et al.* (Oct. 2010). "A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs." *Cell* 143.2, pp. 313–324.

Kim, T.-K., M. Hemberg, J. M. Gray, A. M. Costa, D. M. Bear, J. Wu, *et al.* (May 2010). "Widespread transcription at neuronal activity-regulated enhancers." *Nature* 465.7295, pp. 182–187.

Kim, T.-K. and R. Shiekhattar (Aug. 2015). "Architectural and Functional Commonalities between Enhancers and Promoters." *Cell* 162.5, pp. 948–959.

King, H. W. and R. J. Klose (Mar. 2017). "The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells". *eLife* 6, e22631.

Kitagawa, K., Y. Kotake, Y. Hiramatsu, N. Liu, S. Suzuki, S. Nakamura, *et al.* (Nov. 2010). "GSK3 regulates the expressions of human and mouse c-Myb via different mechanisms". *Cell Division* 5.1, pp. 1–10.

Klemm, S. L., Z. Shipony, and W. J. Greenleaf (Apr. 2019). "Chromatin accessibility and the regulatory epigenome." *Nature Reviews Genetics* 20.4, pp. 207–220.

Klempnauer, K. H., T. J. Gonda, and J. M. Bishop (Dec. 1982). "Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene." *Cell* 31.2 Pt 1, pp. 453–463.

Koch, F., R. Fenouil, M. Gut, P. Cauchy, T. K. Albert, J. Zacarias-Cabeza, *et al.* (2011). "Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters". *Nature structural & molecular biology* 18.8, pp. 956–963.

Krivtsov, A. V., M. E. Figueroa, A. U. Sinha, M. C. Stubbs, Z. Feng, P. J. M. Valk, *et al.* (Apr. 2013). "Cell of origin determines clinically relevant subtypes of MLL-rearranged AML." *Leukemia* 27.4, pp. 852–860.

Krivtsov, A. V. and S. A. Armstrong (Nov. 2007). "MLL translocations, histone modifications and leukaemia stem-cell development." *Nature Reviews Cancer* 7.11, pp. 823–833.

Krivtsov, A. V., Z. Feng, M. E. Lemieux, J. Faber, S. Vempati, A. U. Sinha, *et al.* (Nov. 2008). "H3K79 methylation profiles define murine and human MLL-AF4 leukemias." *Cancer Cell* 14.5, pp. 355–368.

Krivtsov, A. V., D. Twomey, Z. Feng, M. C. Stubbs, Y. Wang, J. Faber, *et al.* (Aug. 2006). "Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9." *Nature* 442.7104, pp. 818–822.

Krogan, N. J., J. Dover, A. Wood, J. Schneider, J. Heidt, M. A. Boateng, *et al.* (Mar. 2003). "The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation." *Molecular cell* 11.3, pp. 721–729.

Lacoste, N., R. T. Utley, J. M. Hunter, G. G. Poirier, and J. Cote (Aug. 2002). "Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase." *The Journal of biological chemistry* 277.34, pp. 30421–30424.

Lai, B., J.-E. Lee, Y. Jang, L. Wang, W. Peng, and K. Ge (June 2017). "MLL3/MLL4 are required for CBP/p300 binding on enhancers and super-enhancer formation in brown adipogenesis." *Nucleic Acids Research* 45.11, pp. 6388–6403.

Lai, F., U. A. Orom, M. Cesaroni, M. Beringer, D. J. Taatjes, G. A. Blobel, *et al.* (Feb. 2013). "Activating RNAs associate with Mediator to enhance chromatin architecture and transcription." *Nature* 494.7438, pp. 497–501.

Lam, M. T. Y., H. Cho, H. P. Lesch, D. Gosselin, S. Heinz, Y. Tanaka-Oishi, *et al.* (June 2013). "Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription." *Nature* 498.7455, pp. 511–515.

Landgraf, P., M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, *et al.* (June 2007). "A mammalian microRNA expression atlas based on small RNA library sequencing." *Cell* 129.7, pp. 1401–1414.

Lara-Astiaso, D., A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, *et al.* (Aug. 2014). "Chromatin state dynamics during blood formation." *Science* 345.6199, pp. 943–949.

Lavallee, V.-P., I. Baccelli, J. Krosl, B. Wilhelm, F. Barabe, P. Gendron, *et al.* (Sept. 2015). "The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias." *Nature Genetics* 47.9, pp. 1030–1037.

Lawrence, H. J., G. Sauvageau, R. K. Humphries, and C. Largman (May 1996). "The role of HOX homeobox genes in normal and leukemic hematopoiesis." *Stem cells (Dayton, Ohio)* 14.3, pp. 281–291.

Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, *et al.* (Sept. 2003). "The nuclear RNase III Drosha initiates microRNA processing." *Nature* 425.6956, pp. 415–419.

Lee, Y., K. Jeon, J.-T. Lee, S. Kim, and V. N. Kim (Sept. 2002). "MicroRNA maturation: stepwise processing and subcellular localization." *Embo Journal* 21.17, pp. 4663–4670.

Lettice, L. A., S. J. H. Heaney, L. A. Purdie, L. Li, P. de Beer, B. A. Oostra, *et al.* (July 2003). "A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly." *Human molecular genetics* 12.14, pp. 1725–1735.

Lettice, L. A., T. Horikoshi, S. J. H. Heaney, M. J. van Baren, H. C. van der Linde, G. J. Breedveld, *et al.* (May 2002). "Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly." *Proceedings of the National Academy of Sciences of the United States of America* 99.11, pp. 7548–7553.

Leverson, J. D., P. J. Koskinen, F. C. Orrico, E. M. Rainio, K. J. Jalkanen, A. B. Dash, *et al.* (Oct. 1998a). "Pim-1 kinase and p100 cooperate to enhance c-Myb activity." *Molecular cell* 2.4, pp. 417–425.

Leverson, J. D. and S. A. Ness (Jan. 1998b). "Point mutations in v-Myb disrupt a cyclophilin-catalyzed negative regulatory mechanism." *Molecular cell* 1.2, pp. 203–211.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, *et al.* (Aug. 2009). "The Sequence Alignment/Map format and SAMtools." *Bioinformatics* 25.16, pp. 2078–2079.

Li, W., D. Notani, Q. Ma, B. Tanasa, E. Nunez, A. Y. Chen, *et al.* (June 2013). "Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation." *Nature* 498.7455, 516 EP–-.

Li, Y., H. Wen, Y. Xi, K. Tanaka, H. Wang, D. Peng, *et al.* (Oct. 2014). "AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation." *Cell* 159.3, pp. 558–571.

Liao, Y., G. K. Smyth, and W. Shi (Apr. 2014). "featureCounts: an efficient general purpose program for assigning sequence reads to genomic features." *Bioinformatics* 30.7, pp. 923–930.

Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, *et al.* (Oct. 2009). "Comprehensive mapping of long-range interactions reveals folding principles of the human genome." *Science* 326.5950, pp. 289–293.

Lieu, Y. K. and E. P. Reddy (2009). "Conditional c-myb knockout in adult hematopoietic stem cells leads to loss of self-renewal due to impaired proliferation and accelerated differentiation". *Proc Natl Acad Sci USA*.

Lieu, Y. K., A. Kumar, A. G. Pajerowski, T. J. Rogers, and E. P. Reddy (Oct. 2004). "Requirement of c-myb in T cell development and in mature T cell function." *Proceedings of the National Academy of Sciences of the United States of America* 101.41, pp. 14853–14858.

Lim, B., T. Heist, M. Levine, and T. Fukaya (Apr. 2018). "Visualization of Transvection in Living Drosophila Embryos". *Molecular cell* 70.2, 287–296.e6.

Lin, C. Y., J. Loven, P. B. Rahl, R. M. Paranal, C. B. Burge, J. E. Bradner, *et al.* (Sept. 2012). "Transcriptional amplification in tumor cells with elevated c-Myc." *Cell* 151.1, pp. 56–67.

Lin, C., E. R. Smith, H. Takahashi, K. C. Lai, S. Martin-Brown, L. Florens, *et al.* (Feb. 2010). "AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia." *Molecular cell* 37.3, pp. 429–437.

Ling, J., B. Baibakov, W. Pi, B. M. Emerson, and D. Tuan (July 2005). "The HS2 enhancer of the beta-globin locus control region initiates synthesis of non-coding, polyadenylated RNAs independent of a cis-linked globin promoter." *Journal of molecular biology* 350.5, pp. 883–896.

Lipsick, J. S. (July 1996). "One billion years of Myb." *Oncogene* 13.2, pp. 223–235.

Liu, X. and N. D. Clarke (Oct. 2002). "Rationalization of gene regulation by a eukaryotic transcription factor: calculation of regulatory region occupancy from predicted binding affinities." *Journal of molecular biology* 323.1, pp. 1–8.

Liu, Y., M. Ding, Q. Gao, A. He, Y. Liu, and H. Mei (2018). "Current Advances on the Important Roles of Enhancer RNAs in Gene Regulation and Cancer." *BioMed Research International* 2018, p. 2405351.

Loven, J., H. A. Hoke, C. Y. Lin, A. Lau, D. A. Orlando, C. R. Vakoc, *et al.* (Apr. 2013). "Selective inhibition of tumor oncogenes by disruption of super-enhancers." *Cell* 153.2, pp. 320–334.

Lüscher, B., E. Christenson, D. W. Litchfield, E. G. Krebs, and R. N. Eisenman (Apr. 1990). "Myb DNA binding inhibited by phosphorylation at a site deleted during oncogenic activation." *Nature* 344.6266, pp. 517–522.

Malaterre, J., T. Mantamadiotis, S. Dworkin, S. Lightowler, Q. Yang, M. I. Ransome, *et al.* (Jan. 2008). "c-Myb is required for neural progenitor cell proliferation and maintenance of the neural stem cell niche in adult brain." *Stem cells (Dayton, Ohio)* 26.1, pp. 173–181.

Maniatis, T., S. Goodbourn, and J. A. Fischer (June 1987). "Regulation of inducible and tissue-specific gene expression." *Science* 236.4806, pp. 1237–1245.

Mansour, M. R., B. J. Abraham, L. Anders, A. Berezovskaya, A. Gutierrez, A. D. Durbin, *et al.* (Dec. 2014). "An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element". *Science* 346.6215, pp. 1373–1377.

Marschalek, R., I. Nilson, K. Lochner, R. Greim, G. Siegler, J. Greil, *et al.* (Nov. 1997). "The structure of the human ALL-1/MLL/HRX gene." *Leukemia & lymphoma* 27.5-6, pp. 417–428.

Marshall, N. F., J. Peng, Z. Xie, and D. H. Price (Oct. 1996). "Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase." *The Journal of biological chemistry* 271.43, pp. 27176–27183.

Marshall, N. F. and D. H. Price (May 1995). "Purification of P-TEFb, a transcription factor required for the transition into productive elongation." *The Journal of biological chemistry* 270.21, pp. 12335–12338.

Martinez, N., B. Drescher, H. Riehle, C. Cullmann, H.-P. Vornlocher, A. Ganser, *et al.* (Aug. 2004). "The oncogenic fusion protein RUNX1-CBFA2T1 supports proliferation and inhibits senescence in t(8;21)-positive leukaemic cells." *BMC cancer* 4, p. 44.

Masamha, C. P. and E. J. Wagner (Jan. 2018). "The contribution of alternative polyadenylation to the cancer phenotype." *Carcinogenesis* 39.1, pp. 2–10.

Masternak, K., N. Peyraud, M. Krawczyk, E. Barras, and W. Reith (Feb. 2003). "Chromatin remodeling and extragenic transcription at the MHC class II locus control region." *Nature Immunology* 4.2, pp. 132–137.

Maurano, M. T., E. Haugen, R. Sandstrom, J. Vierstra, A. Shafer, R. Kaul, *et al.* (Dec. 2015). "Large-scale identification of sequence variants influencing human transcription factor occupancy in vivo." *Nature Genetics* 47.12, pp. 1393–1401.

Mayr, C. and D. P. Bartel (Aug. 2009). "Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells." *Cell* 138.4, pp. 673–684.

Mbangkollo, D., R. Burnett, N. McCabe, M. Thirman, H. Gill, H. Yu, *et al.* (June 1995). "The human MLL gene: nucleotide sequence, homology to the Drosophila trx zinc-finger domain, and alternative splicing." *DNA and cell biology* 14.6, pp. 475–483.

McBryant, S. J., V. H. Adams, and J. C. Hansen (2006). "Chromatin architectural proteins." *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 14.1, pp. 39–51.

Meijer, H. A., Y. W. Kong, W. T. Lu, A. Wilczynska, R. V. Spriggs, S. W. Robinson, *et al.* (Apr. 2013). "Translational repression and eIF4A2

activity are critical for microRNA-mediated gene regulation." *Science* 340.6128, pp. 82–85.

Mettus, R. V., J. Litvin, A. Wali, A. Toscani, K. Latham, K. Hatton, *et al.* (Oct. 1994). "Murine A-myb: evidence for differential splicing and tissue-specific expression." *Oncogene* 9.10, pp. 3077–3086.

Meyer, C., T. Burmeister, D. Gröger, G. Tsaur, L. Fechina, A. Renneville, *et al.* (Feb. 2018). "The MLL recombinome of acute leukemias in 2017." *Leukemia* 32.2, pp. 273–284.

Meyer, S. E., D. E. Muench, A. M. Rogers, T. J. Newkold, E. Orr, E. O'Brien, *et al.* (Aug. 2018). "miR-196b target screen reveals mechanisms maintaining leukemia stemness with therapeutic potential." *The Journal of experimental medicine* 215.8, pp. 2115–2136.

Michaels, Y. S., M. B. Barnkob, H. Barbosa, T. A. Baeumler, M. K. Thompson, V. Andre, *et al.* (Feb. 2019). "Precise tuning of gene expression levels in mammalian cells." *Nature Communications* 10.1, p. 818.

Mifsud, B., F. Tavares-Cadete, A. N. Young, R. Sugar, S. Schoenfelder, L. Ferreira, *et al.* (June 2015). "Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C." *Nature Genetics* 47.6, pp. 598–606.

Miglarese, M. R., A. F. Richardson, N. Aziz, and T. P. Bender (Sept. 1996). "Differential regulation of c-Myb-induced transcription activation by a phosphorylation site in the negative regulatory domain." *The Journal of biological chemistry* 271.37, pp. 22697–22705.

Mikhaylichenko, O., V. Bondarenko, D. Harnett, I. E. Schor, M. Males, R. R. Viales, *et al.* (Jan. 2018). "The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription." *Genes & Development* 32.1, pp. 42–57.

Milne, T. A., C. M. Hughes, R. Lloyd, Z. Yang, O. Rozenblatt-Rosen, Y. Dou, *et al.* (Jan. 2005). "Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors". *Proceedings of the National Academy of Sciences* 102.3, pp. 749–754.

Milne, T. A., J. Kim, G. G. Wang, S. C. Stadler, V. Basrur, S. J. Whitcomb, *et al.* (June 2010). "Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis." *Molecular cell* 38.6, pp. 853–863.

Mink, S., U. Kerber, and K. H. Klempnauer (Apr. 1996). "Interaction of C/EBPbeta and v-Myb is required for synergistic activation of the mim-1 gene." *Molecular and Cellular Biology* 16.4, pp. 1316–1325.

Mirny, L. A. (Dec. 2010). "Nucleosome-mediated cooperativity between transcription factors." *Proceedings of the National Academy of Sciences of the United States of America* 107.52, pp. 22534–22539.

Mitchell, P. and R. Tjian (July 1989). "Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins". *Science* 245.4916, pp. 371–378.

Mizuguchi, G., H. Nakagoshi, T. Nagase, N. Nomura, T. Date, Y. Ueno, *et al.* (June 1990). "DNA binding activity and transcriptional activator function of

the human B-myb protein compared with c-MYB." *The Journal of biological chemistry* 265.16, pp. 9280–9284.

Mohan, M., H.-M. Herz, Y.-H. Takahashi, C. Lin, K. C. Lai, Y. Zhang, *et al.* (Mar. 2010). "Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom)." *Genes & Development* 24.6, pp. 574–589.

Moretti, P., P. Simmons, P. Thomas, D. Haylock, P. Rathjen, M. Vadas, *et al.* (July 1994). "Identification of homeobox genes expressed in human haemopoietic progenitor cells." *Gene* 144.2, pp. 213–219.

Mousavi, K., H. Zare, S. Dell'orso, L. Grontved, G. Gutierrez-Cruz, A. Derfoul, *et al.* (Sept. 2013). "eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci." *Molecular cell* 51.5, pp. 606–617.

Mucenski, M. L., K. McLain, A. B. Kier, S. H. Swerdlow, C. M. Schreiner, T. A. Miller, *et al.* (May 1991). "A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis." *Cell* 65.4, pp. 677–689.

Mueller, D., C. Bach, D. Zeisig, M.-P. García-Cuéllar, S. Monroe, A. Sreekumar, *et al.* (Dec. 2007). "A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification." *Blood* 110.13, pp. 4445–4454.

Mueller, D., M.-P. García-Cuéllar, C. Bach, S. Buhl, E. Maethner, and R. K. Slany (Nov. 2009). "Misguided transcriptional elongation causes mixed lineage leukemia." *PLoS biology* 7.11, e1000249.

Muntean, A. G. and J. L. Hess (Feb. 2012). "The pathogenesis of mixed-lineage leukemia." *Annual Review of Pathology: Mechanisms of Disease* 7.1, pp. 283–301.

Nalabothula, N., G. McVicker, J. Maiorano, R. Martin, J. K. Pritchard, and Y. N. Fondufe-Mittendorf (Feb. 2014). "The chromatin architectural proteins HMGD1 and H1 bind reciprocally and have opposite effects on chromatin structure and gene regulation." *BMC genomics* 15, p. 92.

Nedyalkov, S. T., S. P. Bozhkov, and G. Todorov (Oct. 1975). "Experimental erythroblastosis in the Japanese quial (Cortunix cortunix japonica) induced by E26 leukosis strain". *Acta Vet (Bruno)* 44, pp. 75–78.

Ness, S. A. and J. D. Engel (Oct. 1994). "Vintage reds and whites: combinatorial transcription factor utilization in hematopoietic differentiation." *Current Opinion in Genetics & Development* 4.5, pp. 718–724.

Ness, S. A., E. Kowenz-Leutz, T. Casini, T. Graf, and A. Leutz (May 1993). "Myb and NF-M: combinatorial activators of myeloid genes in heterologous cell types." *Genes & Development* 7.5, pp. 749–759.

Ness, S. A. (Sept. 2003). "Myb protein specificity: evidence of a context-specific transcription factor code." *Blood Cells, Molecules, and Diseases* 31.2, pp. 192–200.

Ng, H. H., Q. Feng, H. Wang, H. Erdjument-Bromage, P. Tempst, Y. Zhang, *et al.* (June 2002). "Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association." *Genes & Development* 16.12, pp. 1518–1527.

Nolis, I. K., D. J. McKay, E. Mantouvalou, S. Lomvardas, M. Merika, and D. Thanos (Dec. 2009). "Transcription factors mediate long-range enhancer-promoter interactions." *Proceedings of the National Academy of Sciences of the United States of America* 106.48, pp. 20222–20227.

Nomura, N., M. Takahashi, M. Matsui, S. Ishii, T. Date, S. Sasamoto, *et al.* (1988). "Isolation of human cDNA clones of myb-related genes, A-myb and B-myb".

Nomura, T., N. Sakai, A. Sarai, T. Sudo, C. Kanei-Ishii, R. G. Ramsay, *et al.* (Oct. 1993). "Negative autoregulation of c-Myb activity by homodimer formation through the leucine zipper." *The Journal of biological chemistry* 268.29, pp. 21914–21923.

O'Rourke, J. P. and S. A. Ness (Mar. 2008). "Alternative RNA splicing produces multiple forms of c-Myb with unique transcriptional activities." *Molecular and Cellular Biology* 28.6, pp. 2091–2101.

Oelgeschläger, M., R. Janknecht, J. Krieg, S. Schreek, and B. Lüscher (June 1996). "Interaction of the co-activator CBP with Myb proteins: effects on Myb-specific transactivation and on the cooperativity with NF-M." *Embo Journal* 15.11, pp. 2771–2780.

Oelgeschläger, M., J. Krieg, J. M. Luscher-Firzlaff, and B. Lüscher (Nov. 1995). "Casein kinase II phosphorylation site mutations in c-Myb affect DNA binding and transcriptional cooperativity with NF-M." *Molecular and Cellular Biology* 15.11, pp. 5966–5974.

Ogata, K., H. Hojo, S. Aimoto, T. Nakai, H. Nakamura, A. Sarai, *et al.* (July 1992). "Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core." *Proceedings of the National Academy of Sciences* 89.14, pp. 6428–6432.

Ogata, K., S. Morikawa, H. Nakamura, A. Sekikawa, T. Inoue, H. Kanai, *et al.* (Nov. 1994). "Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices." *Cell* 79.4, pp. 639–648.

Oh, I. H. and E. P. Reddy (May 1999). "The myb gene family in cell growth, differentiation and apoptosis". *Oncogene* 18.19, pp. 3017–3033.

Okada, Y., Q. Feng, Y. Lin, Q. Jiang, Y. Li, V. M. Coffield, *et al.* (Apr. 2005). "hDOT1L links histone methylation to leukemogenesis." *Cell* 121.2, pp. 167–178.

Okuda, H., A. Kanai, S. Ito, H. Matsui, and A. Yokoyama (Nov. 2015). "AF4 uses the SL1 components of RNAP1 machinery to initiate MLL fusion- and AEP-dependent transcription." *Nature Communications* 6, p. 8869.

Okuda, H., M. Kawaguchi, A. Kanai, H. Matsui, T. Kawamura, T. Inaba, *et al.* (Apr. 2014). "MLL fusion proteins link transcriptional coactivators to previously active CpG-rich promoters." *Nucleic Acids Research* 42.7, pp. 4241–4256.

Okuda, H., S. Takahashi, A. Takaori-Kondo, and A. Yokoyama (Oct. 2016). "TBP loading by AF4 through SL1 is the major rate-limiting step in MLL fusion-dependent transcription." *Cell Cycle* 15.20, pp. 2712–2722.

Oliner, J. D., J. M. Andresen, S. K. Hansen, S. Zhou, and R. Tjian (Nov. 1996). "SREBP transcriptional activity is mediated through an interaction with the CREB-binding protein." *Genes & Development* 10.22, pp. 2903–2911.

Orlando, D. A., M. W. Chen, V. E. Brown, S. Solanki, Y. J. Choi, E. R. Olson, *et al.* (Nov. 2014). "Quantitative ChIP-Seq normalization reveals global modulation of the epigenome." *CELREP* 9.3, pp. 1163–1170.

Ørom, U. A., T. Derrien, M. Beringer, K. Gumireddy, A. Gardini, G. Bussotti, *et al.* (Oct. 2010). "Long noncoding RNAs with enhancer-like function in human cells." *Cell* 143.1, pp. 46–58.

Papaemmanuil, E., M. Gerstung, L. Bullinger, V. I. Gaidzik, P. Paschka, N. D. Roberts, *et al.* (June 2016). "Genomic Classification and Prognosis in Acute Myeloid Leukemia." *New England Journal of Medicine* 374.23, pp. 2209–2221.

Papathanasiou, P., R. Tunningley, D. R. Pattabiraman, P. Ye, T. J. Gonda, B. Whittle, *et al.* (Dec. 2010). "A recessive screen for genes regulating hematopoietic stem cells." *Blood* 116.26, pp. 5849–5858.

Paralkar, V. R., C. C. Taborda, P. Huang, Y. Yao, A. V. Kossenkov, R. Prasad, *et al.* (Apr. 2016). "Unlinking an lncRNA from Its Associated cis Element." *Molecular cell* 62.1, pp. 104–110.

Parker, D., M. Rivera, T. Zor, A. Henrion-Caude, I. Radhakrishnan, A. Kumar, *et al.* (Aug. 1999). "Role of secondary structure in discrimination between constitutive and inducible activators." *Molecular and Cellular Biology* 19.8, pp. 5601–5607.

Parry, P., M. Djabali, M. Bower, J. Khristich, M. Waterman, B. Gibbons, *et al.* (May 1993). "Structure and expression of the human trithorax-like gene 1 involved in acute leukemias." *Proceedings of the National Academy of Sciences of the United States of America* 90.10, pp. 4738–4742.

Pattabiraman, D. R. and T. J. Gonda (Feb. 2013). "Role and potential for therapeutic targeting of MYB in leukemia." *Leukemia* 27.2, pp. 269–277.

Pattabiraman, D. R., C. McGirr, K. Shakhbazov, V. Barbier, K. Krishnan, P. Mukhopadhyay, *et al.* (Apr. 2014). "Interaction of c-Myb with p300 is required for the induction of acute myeloid leukemia (AML) by human AML oncogenes." *Blood* 123.17, pp. 2682–2690.

Pattabiraman, D. R., J. Sun, D. H. Dowhan, S. Ishii, and T. J. Gonda (Sept. 2009). "Mutations in multiple domains of c-Myb disrupt interaction with CBP/p300 and abrogate myeloid transforming ability." *Molecular cancer research : MCR* 7.9, pp. 1477–1486.

Persson, M., Y. Andren, J. Mark, H. M. Horlings, F. Persson, and G. Stenman (Nov. 2009). "Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck." *Proceedings of the National Academy of Sciences* 106.44, pp. 18740–18744.

Petruk, S., J. Cai, R. Sussman, G. Sun, S. K. Kovermann, S. A. Mariani, *et al.* (Apr. 2017a). "Delayed Accumulation of H3K27me3 on Nascent DNA Is Essential for Recruitment of Transcription Factors at Early Stages of Stem Cell Differentiation." *Molecular cell* 66.2, 247–257.e5.

Petruk, S., S. A. Mariani, M. De Dominici, P. Porazzi, V. Minieri, J. Cai, *et al.* (Apr. 2017b). "Structure of Nascent Chromatin Is Essential for Hematopoietic Lineage Specification." *CELREP* 19.2, pp. 295–306.

Pfitzner, E., J. Kirfel, P. Becker, A. Rolke, and R. Schule (May 1998). "Physical interaction between retinoic acid receptor and the oncoprotein myb inhibits retinoic acid-dependent transactivation." *Proceedings of the National Academy of Sciences of the United States of America* 95.10, pp. 5539–5544.

Pineault, N., C. D. Helgason, H. J. Lawrence, and R. K. Humphries (Jan. 2002). "Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny." *Experimental Hematology* 30.1, pp. 49–57.

Prasad, R., T. Yano, C. Sorio, T. Nakamura, R. Rallapalli, Y. Gu, *et al.* (Dec. 1995). "Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia." *Proceedings of the National Academy of Sciences of the United States of America* 92.26, pp. 12160–12164.

Preker, P., J. Nielsen, S. Kammler, S. Lykke-Andersen, M. S. Christensen, C. K. Mapendano, *et al.* (Dec. 2008). "RNA Exosome Depletion Reveals Transcription Upstream of Active Human Promoters". *Science* 322.5909, p. 1851.

Pronk, C. J. H., D. J. Rossi, R. Månsson, J. L. Attema, G. L. Norddahl, C. K. F. Chan, *et al.* (Oct. 2007). "Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy." *Cell Stem Cell* 1.4, pp. 428–442.

Pui, C.-H., W. L. Carroll, S. Meshinchi, and R. J. Arceci (Feb. 2011). "Biology, risk stratification, and therapy of pediatric acute leukemias: an update." *Journal of Clinical Oncology* 29.5, pp. 551–565.

Raisner, R., S. Kharbanda, L. Jin, E. Jeng, E. Chan, M. Merchant, *et al.* (Aug. 2018). "Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation." *CELREP* 24.7, pp. 1722–1729.

Ramachandran, S. and S. Henikoff (Apr. 2016). "Transcriptional Regulators Compete with Nucleosomes Post-replication." *Cell* 165.3, pp. 580–592.

Ramaswamy, K., L. Forbes, G. Minuesa, T. Gindin, F. Brown, M. G. Kharas, *et al.* (Jan. 2018). "Peptidomimetic blockade of MYB in acute myeloid leukemia." *Nature Communications* 9.1, p. 110.

Ramsay, R. G., S. Ishii, and T. J. Gonda (Oct. 1991). "Increase in specific DNA binding by carboxyl truncation suggests a mechanism for activation of Myb." *Oncogene* 6.10, pp. 1875–1879.

Ramsay, R. G., S. Micallef, S. Lightowler, M. L. Mucenski, T. Mantamadiotis, and I. Bertoncello (June 2004). "c-myb Heterozygous mice are hypersensitive to 5-fluorouracil and ionizing radiation." *Molecular cancer research : MCR* 2.6, pp. 354–361.

Reddy, T. E., J. Gertz, F. Pauli, K. S. Kucera, K. E. Varley, K. M. Newberry, *et al.* (May 2012). "Effects of sequence variation on differential allelic transcription factor occupancy and gene expression." *Genome Research* 22.5, pp. 860–869.

Rehwinkel, J., I. Behm-Ansmant, D. Gatfield, and E. Izaurralde (Nov. 2005). "A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing." *RNA (New York, N.Y.)* 11.11, pp. 1640–1647.

Reisenauer, M. R., M. Anderson, L. Huang, Z. Zhang, Q. Zhou, B. C. Kone, *et al.* (Dec. 2009). "AF17 competes with AF9 for binding to Dot1a to up-regulate transcription of epithelial Na+ channel alpha." *The Journal of biological chemistry* 284.51, pp. 35659–35669.

Roberts, S. G. (Aug. 2000). "Mechanisms of action of transcription activation and repression domains." *Cellular and molecular life sciences : CMLS* 57.8-9, pp. 1149–1160.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth (Jan. 2010). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." *Bioinformatics* 26.1, pp. 139–140.

Rogan, D. F., D. J. Cousins, S. Santangelo, P. A. Ioannou, M. Antoniou, T. H. Lee, *et al.* (Feb. 2004). "Analysis of intergenic transcription in the human IL-4/IL-13 gene cluster." *Proceedings of the National Academy of Sciences of the United States of America* 101.8, pp. 2446–2451.

Rosenthal, M. A., M. A. Thompson, S. Ellis, R. H. Whitehead, and R. G. Ramsay (July 1996). "Colonic expression of c-myb is initiated in utero and continues throughout adult life." *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* 7.7, pp. 961–967.

Rosson, D. (Apr. 1990). "Effects of 5' and 3' truncations of the myb gene on the transforming ability of avian myeloblastosis virus (AMV)." *Virology* 175.2, pp. 562–567.

Sabari, B. R., A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, *et al.* (July 2018). "Coactivator condensation at super-enhancers links phase separation and gene control." *Science* 361.6400.

Saikumar, P., R. Murali, and E. P. Reddy (Nov. 1990). "Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions." *Proceedings of the National Academy of Sciences of the United States of America* 87.21, pp. 8452–8456.

Sakamoto, H., G. Dai, K. Tsujino, K. Hashimoto, X. Huang, T. Fujimoto, *et al.* (Aug. 2006). "Proper levels of c-Myb are discretely defined at distinct steps of hematopoietic cell development." *Blood* 108.3, pp. 896–903.

Sakura, H., C. Kanei-Ishii, T. Nagase, H. Nakagoshi, T. J. Gonda, and S. Ishii (Aug. 1989). "Delineation of three functional domains of the transcriptional activator encoded by the c-myb protooncogene." *Proceedings of the National Academy of Sciences of the United States of America* 86.15, pp. 5758–5762.

Sanborn, A. L., S. S. P. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, *et al.* (Nov. 2015). "Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes." *Proceedings of the National Academy of Sciences of the United States of America* 112.47, E6456–65.

Sandberg, M. L., S. E. Sutton, M. T. Pletcher, T. Wiltshire, L. M. Tarantino, J. B. Hogenesch, *et al.* (Feb. 2005). "c-Myb and p300 Regulate Hematopoietic

Stem Cell Proliferation and Differentiation". *Developmental Cell* 8.2, pp. 153–166.

Sandmann, T., C. Girardot, M. Brehme, W. Tongprasit, V. Stolc, and E. E. M. Furlong (Feb. 2007). "A core transcriptional network for early mesoderm development in Drosophila melanogaster." *Genes & Development* 21.4, pp. 436–449.

Sano, Y. and S. Ishii (Jan. 2001). "Increased affinity of c-Myb for CREB-binding protein (CBP) after CBP-induced acetylation." *The Journal of biological chemistry* 276.5, pp. 3674–3682.

Sanyal, A., B. R. Lajoie, G. Jain, and J. Dekker (Sept. 2012). "The long-range interaction landscape of gene promoters." *Nature* 489.7414, pp. 109–113.

Sauvageau, G., P. M. Lansdorp, C. J. Eaves, D. E. Hogge, W. H. Dragowska, D. S. Reid, *et al.* (Dec. 1994). "Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells." *Proceedings of the National Academy of Sciences of the United States of America* 91.25, pp. 12223–12227.

Schaukowitch, K., J.-Y. Joo, X. Liu, J. K. Watts, C. Martinez, and T.-K. Kim (Oct. 2014). "Enhancer RNA facilitates NELF release from immediate early genes." *Molecular cell* 56.1, pp. 29–42.

Schoenfelder, S. and P. Fraser (May 2019). "Long-range enhancer–promoter contacts in gene expression control". *Nature Reviews Genetics* 20.8, pp. 437–455.

Schoenfelder, S., M. Furlan-Magaril, B. Mifsud, F. Tavares-Cadete, R. Sugar, B.-M. Javierre, *et al.* (Apr. 2015). "The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements." *Genome Research* 25.4, pp. 582–597.

Schotte, D., E. A. M. Lange-Turenhout, D. J. P. M. Stumpel, R. W. Stam, J. G. C. A. M. Buijs-Gladdines, J. P. P. Meijerink, *et al.* (Oct. 2010). "Expression of miR-196b is not exclusively MLL-driven but is especially linked to activation of HOXA genes in pediatric acute lymphoblastic leukemia." *Haematologica* 95.10, pp. 1675–1682.

Schuijers, J., J. C. Manteiga, A. S. Weintraub, D. S. Day, A. V. Zamudio, D. Hnisz, *et al.* (Apr. 2018). "Transcriptional Dysregulation of MYC Reveals Common Enhancer-Docking Mechanism." *CELREP* 23.2, pp. 349–360.

Shapiro, L. H. (Apr. 1995). "Myb and Ets proteins cooperate to transactivate an early myeloid gene." *The Journal of biological chemistry* 270.15, pp. 8763–8771.

Sheffield, N. C. and C. Bock (Feb. 2016). "LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor". *Bioinformatics* 32.4, pp. 587–589.

Shen-Ong, G. L., R. M. J. Skurla, J. D. Owens, and J. F. Mushinski (June 1990). "Alternative splicing of RNAs transcribed from the human c-myb gene." *Molecular and Cellular Biology* 10.6, pp. 2715–2722.

Shilatifard, A., W. S. Lane, K. W. Jackson, R. C. Conaway, and J. W. Conaway (Mar. 1996). "An RNA polymerase II elongation factor encoded by the human ELL gene." *Science* 271.5257, pp. 1873–1876.

Shin, D. H., H. W. Lee, G. S. Jeon, H. Y. Lee, K. H. Lee, and S. S. Cho (Feb. 2001). "Constitutive expression of c-myb mRNA in the adult rat brain." *Brain research* 892.1, pp. 203–207.

Shin, Y., Y.-C. Chang, D. S. W. Lee, J. Berry, D. W. Sanders, P. Ronceray, *et al.* (Nov. 2018). "Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome." *Cell* 175.6, 1481–1491.e13.

Shrinivas, K., B. R. Sabari, E. L. Coffey, I. A. Klein, A. Boija, A. V. Zamudio, *et al.* (Aug. 2019). "Enhancer Features that Drive Formation of Transcriptional Condensates." *Molecular cell* 75.3, 549–561.e7.

Sigler, P. B. (May 1988). "Transcriptional activation. Acid blobs and negative noodles." *Nature* 333.6170, pp. 210–212.

Slany, R. K., C. Lavau, and M. L. Cleary (Jan. 1998). "The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX." *Molecular and Cellular Biology* 18.1, pp. 122–129.

Small, S., A. Blair, and M. Levine (Nov. 1992). "Regulation of even-skipped stripe 2 in the Drosophila embryo." *Embo Journal* 11.11, pp. 4047–4057.

Somervaille, T. C. P., C. J. Matheny, G. J. Spencer, M. Iwasaki, J. L. Rinn, D. M. Witten, *et al.* (Feb. 2009). "Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells." *Cell Stem Cell* 4.2, pp. 129–140.

Spitz, F. and E. E. M. Furlong (Sept. 2012). "Transcription factors: from enhancer binding to developmental control." *Nature Reviews Genetics* 13.9, pp. 613–626.

Sramko, M., J. Markus, J. Kabat, L. Wolff, and J. Bies (Dec. 2006). "Stress-induced inactivation of the c-Myb transcription factor through conjugation of SUMO-2/3 proteins." *The Journal of biological chemistry* 281.52, pp. 40065–40075.

Staby, L., C. O'Shea, M. Willemoes, F. Theisen, B. B. Kragelund, and K. Skriver (July 2017). "Eukaryotic transcription factors: paradigms of protein intrinsic disorder." *The Biochemical journal* 474.15, pp. 2509–2532.

Stavropoulou, V., S. Kaspar, L. Brault, M. A. Sanders, S. Juge, S. Morettini, *et al.* (July 2016). "MLL-AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT-Related Genes Linked to Poor Outcome." *Cancer Cell* 30.1, pp. 43–58.

Storey, J., A. Bass, A. Dabney, and R. D (2020). *qvalue: Q-value estimation for false discovery rate control.* R package version 2.20.0.

Sumner, R., A. Crawford, M. Mucenski, and J. Frampton (July 2000). "Initiation of adult myelopoiesis can occur in the absence of c-Myb whereas subsequent development is strictly dependent on the transcription factor." *Oncogene* 19.30, pp. 3335–3342.

Svaren, J., E. Klebanow, L. Sealy, and R. Chalkley (Mar. 1994). "Analysis of the competition between nucleosome formation and transcription factor binding." *The Journal of biological chemistry* 269.12, pp. 9335–9344.

Taberlay, P. C., T. K. Kelly, C.-C. Liu, J. S. You, D. D. De Carvalho, T. B. Miranda, *et al.* (Dec. 2011). "Polycomb-repressed genes have permissive enhancers that initiate reprogramming." *Cell* 147.6, pp. 1283–1294.

Tahirov, T. H., M. Sasaki, T. Inoue-Bungo, A. Fujikawa, K. Sato, T. Kumasaka, *et al.* (Oct. 2001). "Crystals of ternary protein–DNA complexes composed of DNA-binding domains of c-Myb or v-Myb, C/EBP$\alpha$ or C/EBP$\beta$ and tom-1A promoter fragment". *Acta Crystallographica Section D Biological Crystallography* 57.11, pp. 1655–1658.

Takahashi, H., T. J. Parmely, S. Sato, C. Tomomori-Sato, C. A. S. Banks, S. E. Kong, *et al.* (July 2011). "Human mediator subunit MED26 functions as a docking site for transcription elongation factors." *Cell* 146.1, pp. 92–104.

Tanaka, Y., T. Nomura, and S. Ishii (Aug. 1997). "Two regions in c-myb proto-oncogene product negatively regulating its DNA-binding activity." *FEBS Letters* 413.1, pp. 162–168.

Tanaka, Y., N. P. Patestos, T. Maekawa, and S. Ishii (Oct. 1999). "B-myb is required for inner cell mass formation at an early stage of development." *The Journal of biological chemistry* 274.40, pp. 28067–28070.

Tanikawa, J., T. Yasukawa, and M. Enari (1993). "Recognition of specific DNA sequences by the c-myb protooncogene product: role of three repeat units in the DNA-binding domain". *Proc Natl Acad Sci USA*.

Tavner, F. J., R. Simpson, S. Tashiro, D. Favier, N. A. Jenkins, D. J. Gilbert, *et al.* (Feb. 1998). "Molecular cloning reveals that the p160 Myb-binding protein is a novel, predominantly nucleolar protein which may play a role in transactivation by Myb." *Molecular and Cellular Biology* 18.2, pp. 989–1002.

Thomas, M. D., C. S. Kremer, K. S. Ravichandran, K. Rajewsky, and T. P. Bender (Sept. 2005). "c-Myb is critical for B cell development and maintenance of follicular B cells." *Immunity* 23.3, pp. 275–286.

Thoms, J. A. I., D. Beck, and J. E. Pimanda (Dec. 2019). "Transcriptional networks in acute myeloid leukemia." *Genes, chromosomes & cancer* 58.12, pp. 859–874.

Thorsteinsdottir, U., A. Mamo, E. Kroon, L. Jerome, J. Bijl, H. J. Lawrence, *et al.* (Jan. 2002). "Overexpression of the myeloid leukemia-associated Hoxa9 gene in bone marrow cells induces stem cell expansion." *Blood* 99.1, pp. 121–129.

Thurman, R. E., E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano, E. Haugen, *et al.* (Sept. 2012). "The accessible chromatin landscape of the human genome." *Nature* 489.7414, pp. 75–82.

Tillo, D., N. Kaplan, I. K. Moore, Y. Fondufe-Mittendorf, A. J. Gossett, Y. Field, *et al.* (Feb. 2010). "High nucleosome occupancy is encoded at human regulatory sequences." *PLoS ONE* 5.2, e9129.

Tkachuk, D. C., S. Kohler, and M. L. Cleary (Nov. 1992). "Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias." *Cell* 71.4, pp. 691–700.

Tomita, A., M. Towatari, S. Tsuzuki, F. Hayakawa, H. Kosugi, K. Tamai, *et al.* (Jan. 2000). "c-Myb acetylation at the carboxyl-terminal conserved domain by transcriptional co-activator p300." *Oncogene* 19.3, pp. 444–451.

Toscani, A., R. V. Mettus, R. Coupland, H. Simpkins, J. Litvin, J. Orth, *et al.* (Apr. 1997). "Arrest of spermatogenesis and defective breast development in mice lacking A-myb." *Nature* 386.6626, pp. 713–717.

Trauth, K., B. Mutschler, N. A. Jenkins, D. J. Gilbert, N. G. Copeland, and K. H. Klempnauer (Dec. 1994). "Mouse A-myb encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes." *The EMBO Journal* 13.24, pp. 5994–6005.

Triezenberg, S. J. (Apr. 1995). "Structure and function of transcriptional activation domains." *Current Opinion in Genetics & Development* 5.2, pp. 190–196.

Tuan, D., S. Kong, and K. Hu (Dec. 1992). "Transcription of the hypersensitive site HS2 enhancer in erythroid cells." *Proceedings of the National Academy of Sciences of the United States of America* 89.23, pp. 11219–11223.

Tzelepis, K., E. De Braekeleer, D. Aspris, I. Barbieri, M. S. Vijayabaskar, W.-H. Liu, *et al.* (Dec. 2018). "SRPK1 maintains acute myeloid leukemia through effects on isoform usage of epigenetic regulators including BRD4". *Nature Communications* 9.1, 5378–1012.e6.

UniProt Consortium (Jan. 2019). "UniProt: a worldwide hub of protein knowledge." *Nucleic Acids Research* 47.D1, pp. D506–D515.

Vainberg Slutskin, I., S. Weingarten-Gabbay, R. Nir, A. Weinberger, and E. Segal (Feb. 2018). "Unraveling the determinants of microRNA mediated regulation using a massively parallel reporter assay." *Nature Communications* 9.1, p. 529.

Vaquerizas, J. M., S. K. Kummerfeld, S. A. Teichmann, and N. M. Luscombe (Apr. 2009). "A census of human transcription factors: function, expression and evolution." *Nature Reviews Genetics* 10.4, pp. 252–263.

Vegiopoulos, A., P. Garcia, N. Emambokus, and J. Frampton (June 2006). "Coordination of erythropoiesis by the transcription factor c-Myb." *Blood* 107.12, pp. 4703–4710.

Verbeek, W., A. F. Gombart, A. M. Chumakov, C. Muller, A. D. Friedman, and H. P. Koeffler (May 1999). "C/EBPepsilon directly interacts with the DNA binding domain of c-myb and cooperatively activates transcription of myeloid promoters." *Blood* 93.10, pp. 3327–3337.

Vorbrueggen, G., J. Lovric, and K. Moelling (Nov. 1996). "Functional analysis of phosphorylation at serine 532 of human c-Myb by MAP kinase." *Biological chemistry* 377.11, pp. 721–730.

Voss, T. C., R. L. Schiltz, M.-H. Sung, P. M. Yen, J. A. Stamatoyannopoulos, S. C. Biddie, *et al.* (Aug. 2011). "Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism." *Cell* 146.4, pp. 544–554.

Wada, T., T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa (Dec. 1998). "Evidence that P-TEFb alleviates the negative effect of DSIF on RNA

polymerase II-dependent transcription in vitro." *Embo Journal* 17.24, pp. 7395–7403.

Walhout, A. J. M. (Dec. 2006). "Unraveling transcription regulatory networks by protein-DNA and protein-protein interaction mapping." *Genome Research* 16.12, pp. 1445–1454.

Wan, L., H. Wen, Y. Li, J. Lyu, Y. Xi, T. Hoshii, *et al.* (Mar. 2017). "ENL links histone acetylation to oncogenic gene expression in acute myeloid leukaemia." *Nature* 543.7644, pp. 265–269.

Wang, D., I. Garcia-Bassets, C. Benner, W. Li, X. Su, Y. Zhou, *et al.* (May 2011). "Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA." *Nature* 474.7351, pp. 390–394.

Wee, L. M., C. F. Flores-Jasso, W. E. Salomon, and P. D. Zamore (Nov. 2012). "Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties." *Cell* 151.5, pp. 1055–1067.

Wei, J., M. Wunderlich, C. Fox, S. Alvarez, J. C. Cigudosa, J. S. Wilhelm, *et al.* (June 2008). "Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia." *Cancer Cell* 13.6, pp. 483–495.

Weintraub, A. S., C. H. Li, A. V. Zamudio, A. A. Sigova, N. M. Hannett, D. S. Day, *et al.* (Dec. 2017). "YY1 Is a Structural Regulator of Enhancer-Promoter Loops." *Cell* 171.7, 1573–1588.e28.

Weston, K. and J. M. Bishop (July 1989). "Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb." *Cell* 58.1, pp. 85–93.

Whyte, W. A., D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, *et al.* (Apr. 2013). "Master transcription factors and mediator establish super-enhancers at key cell identity genes." *Cell* 153.2, pp. 307–319.

Wilson, N. K., S. D. Foster, X. Wang, K. Knezevic, J. Schutte, P. Kaimakis, *et al.* (Oct. 2010). "Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators." *Cell Stem Cell* 7.4, pp. 532–544.

Wu, S.-Y. and C.-M. Chiang (May 2007). "The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation." *The Journal of biological chemistry* 282.18, pp. 13141–13145.

Wunderlich, Z. and L. A. Mirny (Oct. 2009). "Different gene regulation strategies revealed by analysis of binding motifs." *Trends in Genetics* 25.10, pp. 434–440.

Xi, H., H. P. Shulha, J. M. Lin, T. R. Vales, Y. Fu, D. M. Bodine, *et al.* (Aug. 2007). "Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome." *PLoS genetics* 3.8, e136.

Xu, Y., J. P. Milazzo, T. D. D. Somerville, Y. Tarumoto, Y.-H. Huang, E. L. Ostrander, *et al.* (Jan. 2018). "A TFIID-SAGA Perturbation that Targets MYB and Suppresses Acute Myeloid Leukemia." *Cancer Cell* 33.1, 13–28.e8.

Yamaguchi, Y., T. Takagi, T. Wada, K. Yano, A. Furuya, S. Sugimoto, *et al.* (Apr. 1999). "NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation." *Cell* 97.1, pp. 41–51.

Yan, J., S.-A. A. Chen, A. Local, T. Liu, Y. Qiu, A.-Y. Lee, *et al.* (Feb. 2017). "Histone H3 Lysine 4 methyltransferases MLL3 and MLL4 Modulate Long-range Chromatin Interactions at Enhancers". *bioRxiv*, p. 110239.

Yi, R., Y. Qin, I. G. Macara, and B. R. Cullen (Dec. 2003). "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." *Genes & Development* 17.24, pp. 3011–3016.

Yin, J.-w. and G. Wang (Mar. 2014). "The Mediator complex: a master coordinator of transcription and cell lineage development." *Development* 141.5, pp. 977–987.

Yokoyama, A. and M. L. Cleary (July 2008). "Menin critically links MLL proteins with LEDGF on cancer-associated target genes." *Cancer Cell* 14.1, pp. 36–46.

Yokoyama, A., M. Lin, A. Naresh, I. Kitabayashi, and M. L. Cleary (Feb. 2010). "A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription." *Cancer Cell* 17.2, pp. 198–212.

Yokoyama, A., T. C. P. Somervaille, K. S. Smith, O. Rozenblatt-Rosen, M. Meyerson, and M. L. Cleary (Oct. 2005). "The Menin Tumor Suppressor Protein Is an Essential Oncogenic Cofactor for MLL-Associated Leukemogenesis". *Cell* 123.2, pp. 207–218.

Yuan, W., G. Condorelli, M. Caruso, A. Felsani, and A. Giordano (Apr. 1996). "Human p300 protein is a coactivator for the transcription factor MyoD." *The Journal of biological chemistry* 271.15, pp. 9009–9013.

Yuh, C. H., A. Ransick, P. Martinez, R. J. Britten, and E. H. Davidson (Aug. 1994). "Complexity and organization of DNA-protein interactions in the 5'-regulatory region of an endoderm-specific marker gene in the sea urchin embryo." *Mechanisms of development* 47.2, pp. 165–186.

Zatkova, A., S. Merk, M. Wendehack, M. Bilban, E. M. Muzik, A. Muradyan, *et al.* (June 2009). "AML/MDS with 11q/MLL amplification show characteristic gene expression signature and interplay of DNA copy number changes." *Genes, chromosomes & cancer* 48.6, pp. 510–520.

Zatkova, A., R. Ullmann, J. M. Rouillard, B. J. Lamb, R. Kuick, S. M. Hanash, *et al.* (Apr. 2004). "Distinct sequences on 11q13.5 and 11q23-24 are frequently coamplified with MLL in complexly organized 11q amplicons in AML/MDS patients." *Genes, chromosomes & cancer* 39.4, pp. 263–276.

Zeng, L. and M. M. Zhou (Feb. 2002). "Bromodomain: an acetyl-lysine binding domain." *FEBS Letters* 513.1, pp. 124–128.

Zhang, H., F. A. Kolb, L. Jaskiewicz, E. Westhof, and W. Filipowicz (July 2004). "Single processing center models for human Dicer and bacterial RNase III." *Cell* 118.1, pp. 57–68.

Ziemin-van der Poel, S., N. R. McCabe, H. J. Gill, R. Espinosa, Y. Patel, A. Harden, *et al.* (Dec. 1991). "Identification of a gene, MLL, that spans the

breakpoint in 11q23 translocations associated with human leukemias." *Proceedings of the National Academy of Sciences* 88.23, pp. 10735–10739.

Zor, T., R. N. De Guzman, H. J. Dyson, and P. E. Wright (Mar. 2004). "Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb." *Journal of molecular biology* 337.3, pp. 521–534.

Zorbas, M., C. Sicurella, I. Bertoncello, D. Venter, S. Ellis, M. L. Mucenski, *et al.* (Oct. 1999). "c-Myb is critical for murine colon development." *Oncogene* 18.42, pp. 5821–5830.

Zuber, J., A. R. Rappaport, W. Luo, E. Wang, C. Chen, A. V. Vaseva, *et al.* (Aug. 2011). "An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance." *Genes & Development* 25.15, pp. 1628–1640.