# Regulation and role of TBX2 and TBX3 in tumour progression




Sizhu Lu

Supervised by Professor Colin Goding

Green Templeton College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity Term 2020

# Disclaimer

I declare that this submission is my own written work. To the best of my knowledge, it contains no material that has been written by another or previously submitted in the same or a different form as academic work, to this or any other university.

The results presented herein derive from my own experiments, with the exception of the original USF1/2 ChIP-seq data in Figure 4.4 and Figure 5.6 were provided by Dr Pakavarin Louphrasitthiphol.

Sizhu Lu
Trinity term 2020

# Acknowledgements

First I would like to thank my supervisor Professor Colin Goding for the opportunity to conduct my DPhil in LICR and for his guidance, ideas and kind encouragement, especially when things were not going as planned. I never cease to be amazed at the passion for science and dedication to work found in Colin and my colleagues. I am hugely grateful to everyone in the lab for the supportive and inspiring environment to work in, and for all the help, feedback, suggestions and constructive criticism I received.

For this research project, I would like to thank Jenna Bleloch for her partnership in the drug screen and cheering smile (despite the dreadful amount of work). And my thanks also go to Val, Elena and Dan in TDI for the fun robotic experience. I would also like to thank Mafalda Almeida, Professor Neil Brockdorff, Dr Neil Blackledge and Professor Rob Klose for providing plasmids for PRC1. I am grateful to the Ludwig Institute for Cancer Research for the wonderful facilities and support, particularly Richard Lisle for IF support. I am deeply indebted to Paka for his help with ChIP-seq and bioinformatics; to Rom for his critical advice in Co-IP as well as the (annoying) green dots, and introducing me to ICSA; to Yurena for her help especially when I first started in the lab (I will never forget protein goes from the dark to the bright in a WB transfer). And all my friends in the Goding lab, Linxin, Sarah, Diogo, Laura, Katie and Paola; special thanks go to Sarah, who brings to the lab a refreshing new vibe (and laughs!), to Diogo for all his help and encouragement, as well as scientific discussion and amusing chatters.

I owe a debt of gratitude to my friends in Oxford, especially Wenjun, Carol, Xiao and Chen-Yi for their support, company and all the happy memories, and Adrienne for sharing with me invaluable experience and inspiration. Thanks also to Green Templeton College and my college advisor Professor Shoumo Bhattacharya.

Finally, I would like to thank my parents for their unreserved support and love, and my dearest friend June, JFF and Celia. Miles away, year after year, yet we never grew apart.

# Abstract

TBX2 and TBX3 belong to the T-box family of transcription factors and are important in embryonic development and tumourigenesis. Overexpression of TBX2/3 has been observed in several types of cancer, promoting tumour proliferation, senescence bypass and invasiveness.

Given their important contribution to tumour progression, it might be of therapeutic benefit to pharmacologically inhibit their expression for cancer treatment. To identify novel regulators of TBX2/3, we carried out a high throughput screen for FDA-approved drugs that could affect TBX2/3 expression and/or subcellular localisation using melanoma cells ectopically expressing TBX2/3. We identified hit drugs from the initial imaging-base screen and confirmed the results of selected compounds targeting calcineurin.

In addition to the regulation of TBX2/3 protein, to further understand their role in tumour progression, we undertook genome-wide ChIP-seq and RNA-seq experiments to identify potential target genes and explore functional implications. As there is a lack of ChIP-grade TBX2/3 antibodies, we tagged the endogenous proteins with 3xHA using the CRISPR/Cas9 system. The generated cell lines produced high-quality ChIP-seq datasets of TBX2 and TBX3 and revealed the repertoire of binding sites genome-wide. Surprisingly, TBX2 binds to not only T-elements as expected, but also to a subset of E-box sequences bound by USF2. Integrative analysis of ChIP-seq and RNA-seq datasets identified E2F1 as a novel target gene of TBX2 that promotes cell growth.

Last, but not least, we also identified novel cofactors of TBX2, with preliminary data suggesting that TBX2 represses gene expression by interacting with polycomb repressive complex 1, and may be converted to a transactivator when interacting with basis helix-loop-helix transcription factor USF2.

# Contents

**List of Figures** **ix**

**List of Tables** **xi**

**1 Introduction** **1**
1.1 Melanoma . . . 1
1.1.1 Cancer . . . 1
1.1.2 Melanoma progression . . . 10
1.1.3 Melanoma therapy . . . 13
1.2 Gene regulation by transcription factors . . . 17
1.2.1 Transcriptional control of genes . . . 17
1.2.2 Transcription factors . . . 19
1.3 The T-box transcription factor family . . . 20
1.3.1 TBX2 and TBX3 gene, mRNA splicing and protein . . . 22
1.3.2 TBX2 and TBX3 in cancer . . . 25
1.3.3 TBX2 and TBX3 in development . . . 33
1.3.4 TBX2/3 in other human diseases . . . 37
1.3.5 TBX2 and TBX3 others . . . 38
1.4 Objectives . . . 39

**2 Materials and methods** **40**
2.1 Mammalian cell methods . . . 40
2.1.1 Routine cell culture . . . 40
2.1.2 Immunofluorescence (IF) . . . 40
2.2 Nucleic acid methods . . . 41
2.2.1 Clone sgRNA into Cas9 plasmid . . . 41
2.2.2 Linear short homology arms synthesis . . . 44
2.2.3 qRT-PCR . . . 45
2.3 Protein methods . . . 48
2.3.1 Western blotting . . . 48
2.4 Cell biology methods . . . 50

2.4.1 eFlut tagging endogenous Tbx2 . . . 50
2.4.2 Nuclear tethering assay . . . 52
2.4.3 Crystal violet staining . . . 53
2.4.4 Cell cycle analysis by FACS . . . 54
2.5 High-throughput sequencing and screen . . . 54
2.5.1 ChIP-seq sample preparation . . . 54
2.5.2 RNA-seq sample preparation . . . 58
2.5.3 FDA-approved drug library screen . . . 59
2.6 Bioinformatics . . . 61
2.6.1 RNA-seq analysis . . . 61
2.6.2 ChIP-seq analysis . . . 62
2.7 Materials . . . 63
2.7.1 Oligonucleotide list . . . 63
2.7.2 siRNA list . . . 64
2.7.3 Buffer recipes . . . 65
2.7.4 Reagents and Kits . . . 68

**3 Tagging endogenous Tbx2 using the CRISPR/Cas9 system 71**
3.1 Introduction . . . 71
3.1.1 CRISPR/Cas9 for gene editing . . . 72
3.1.2 eFlut tagging system . . . 74
3.2 Optimisation of tagging methods . . . 75
3.2.1 Cut-to-modification distance . . . 76
3.2.2 Blocking mutation to prevent re-cleavage . . . 77
3.3 Validation of eFlut-tagged Cell Lines . . . 78
3.3.1 Validate TBX2-3xHA protein fusion . . . 79
3.3.2 Validation of TBX2 subcellular localisation . . . 80
3.3.3 Validation of genomic DNA modification . . . 81
3.3.4 Validation of TBX2 function . . . 84

**4 Identification of TBX2 target genes 85**
4.1 Introduction . . . 85
4.1.1 Known direct target genes of TBX2 . . . 85
4.1.2 Identifying novel TBX2 target genes . . . 89
4.2 Tbx2 Binding Sites - ChIP-seq . . . 90
4.2.1 Tbx2 ChIP-seq . . . 90
4.2.2 Tbx2 ChIP-seq results . . . 91
4.3 siTbx2 gene expression profile - RNA-seq . . . 98
4.3.1 siTbx2 RNA-seq experiment . . . 98
4.3.2 siTbx2 RNA-seq results . . . 100

4.4 Identification of novel Tbx2 targets . . . 104
4.4.1 Integration of DNA binding and gene expression profile . . . 105
4.4.2 Direct target genes of Tbx2 . . . 108
4.4.3 Functional assays . . . 109
4.5 Discussion . . . 112
4.5.1 Tbx2 ChIP-seq discussion . . . 112
4.5.2 siTbx2 RNA-seq discussion . . . 114

**5 Binding sites of TBX3** **116**
5.1 Introduction . . . 116
5.1.1 Known direct target genes of TBX3 . . . 116
5.1.2 Identification of novel TBX3 target genes . . . 119
5.2 ChIP-seq experiment . . . 120
5.2.1 Inducible TBX3-3xHA cell line . . . 120
5.2.2 TBX3 ChIP experiment . . . 122
5.3 ChIP-seq results . . . 123
5.3.1 ChIP experiment validation . . . 123
5.3.2 ChIP-seq quality check . . . 124
5.3.3 ChIP-seq analysis . . . 125
5.4 Discussion . . . 129
5.4.1 Comparison of TBX2 and TBX3 binding sites . . . 129

**6 Interacting factors of Tbx2** **133**
6.1 Introduction . . . 133
6.1.1 Known TBX2 interactors . . . 133
6.1.2 Known TBX3 interactors . . . 136
6.1.3 Identifying novel TBX2/3 interactors . . . 138
6.2 Protein-protein interaction of TBX2 . . . 138
6.2.1 Mass spectrometry data . . . 138
6.2.2 Interaction between Tbx2 and PRC1 . . . 140
6.3 Other cofactors of TBX2 . . . 143
6.4 Discussion . . . 145

**7 Regulation of TBX2 and TBX3** **147**
7.1 Introduction . . . 147
7.1.1 Transcriptional regulation of TBX2/3 . . . 148
7.1.2 Post-transcriptional regulation of TBX2/3 . . . 150
7.1.3 Post-translational modification of TBX2/3 . . . 151
7.1.4 Other known regulators of TBX2/3 . . . 152
7.2 Screen for TBX2 and TBX3 regulators . . . 154

7.2.1 Drug screen experiment . . . . . 155
7.2.2 Drug screen results – overview . . . . . 157
7.2.3 Drug screen validation . . . . . 160
7.3 Epigenetic control of TBX2 expression . . . . . 163
7.4 Discussion . . . . . 164

**8 Discussion** **166**
8.1 Mechanism of bifunctional transcription factors . . . . . 166
8.1.1 TBX2/3 can be a transactivator . . . . . 167
8.1.2 TBX2/3 function depends on cofactors . . . . . 168
8.2 Targeting TBX2/3 in cancer treatment . . . . . 169
8.2.1 Cancer hallmarks and therapies . . . . . 169
8.2.2 Inhibiting TBX2/3 for cancer treatment . . . . . 171
8.3 Limitations and future perspectives . . . . . 174
8.3.1 Lack of appropriate tools . . . . . 174
8.3.2 Impact of the Covid-19 pandemic . . . . . 175
8.3.3 Future perspectives . . . . . 176

**Works Cited** **179**

# List of Figures


1.1 The hallmarks of cancer . . . 3
1.2 Histopathological subtypes and cause of melanoma . . . 12
1.3 BRAF-MEK signalling pathway in melanoma . . . 15
1.4 Regulation of transcription in eukaryotic cells . . . 18
1.5 Phylogenetic tree of the T-box transcription factors . . . 21
1.6 TBX2/3 protein structure . . . 23

3.1 eFlut tagging strategy . . . 74
3.2 Optimisation of the eFlut experiment . . . 77
3.3 Validate TBX2-3xHA protein fusion and subcellular localisation . . . 80
3.4 Validate genomic DNA modification . . . 82

4.1 Tbx2 ChIP-seq experiment . . . 92
4.2 Tbx2 ChIP-seq quality control . . . 94
4.3 Tbx2 ChIP-seq analysis . . . 96
4.4 Comparison of Tbx2 and USF1/2 binding sites . . . 99
4.5 siTbx2 experiment and RNA-seq analysis pipeline . . . 100
4.6 Tbx2 RNA-seq quality control . . . 101
4.7 siTbx2 RNA-seq analysis — GO and GSEA . . . 103
4.8 siTbx2 RNA-seq analysis — KEGG . . . 105
4.9 Integrative analysis of ChIP-seq and RNA-seq . . . 107
4.10 Tbx2 potential direct targets . . . 110
4.11 Functional assays of Tbx2 silencing . . . 111

5.1 Inducible expression of TBX3 in 501mel cells . . . 120
5.2 ChIP experiment results . . . 123
5.3 TBX3 ChIP-seq quality control . . . 125
5.4 TBX3 ChIP-seq peak examples . . . 126
5.5 TBX3 ChIP-seq analysis . . . 128
5.6 Expression levels of TBX2 and TBX3 in melanoma cell lines . . . 130
5.7 Comparison of Tbx2 and USF1/2 binding sites . . . 132

6.1 LacO/LacI tethering assay illustration . . . . . 141
6.2 LacO/LacI tethering assay result of Tbx2 and Pcgf1 . . . . . 143
6.3 LacO/LacI tethering assay of Tbx2 and Usf1/2 . . . . . 144

7.1 Assay development for the drug screen . . . . . 156
7.2 Overview of the drug screen results . . . . . 158
7.3 Preliminary validation of CaNis . . . . . 161
7.4 KDM inhibitor reduces TBX2 expression . . . . . 163

# List of Tables

4.1 Co-occurring motifs in Tbx2 peak sequences . . . . . 97
4.2 Distance between TBE and E-box . . . . . 97
4.3 Statistically significant DEGs from two siRNAs against Tbx2 . . . . 102

6.1 BioID result for Tbx2 . . . . . 139

7.1 Known target and/or function of hit drugs . . . . . 159

# List of Abbreviations

| | |
|---|---|
| BCL9L | B-cell lymphoma 9-like |
| bHLH | Basic helix-loop-helix |
| BSA | Bovine serum albumin |
| bZIP | Basic leucine zipper |
| CaNis | Calcineurin inhibitors |
| CHD | Congenital heart disease |
| CHD1 | Chromodomain helicase DNA-binding protein 1 |
| ChIP | Chromatin immunoprecipitation |
| CIA | Collagen-induced arthritis |
| Co-IP | Co-immunoprecipitation |
| CRC | Colorectal cancer |
| CRISPaint | CRISPR-assisted insertion tagging |
| CRISPR | Clustered, regularly interspaced, short palindromic repeat |
| CsA | Cyclosporin A |
| CSD | Chronic sun-induced damage |
| CTLs | Cytotoxic T lymphocytes |
| DAI | DNA affinity immunoblotting |
| DBD | DNA-binding domain |
| DCIS | Ductal carcinoma *in situ* |
| DEGs | Differentially expressed genes |

| | |
|---|---|
| DFS | Disease-free survival |
| dn | Dominant negative |
| DSB | Double-strand break |
| EDTA | Ethylenediaminetetraacetic acid |
| eFlut | Endogenous Fluorescent tagging |
| EGFR | Epidermal growth factor receptor |
| EMSA | Electrophoretic mobility shift assay |
| EMT | Epithelial-mesenchymal transition |
| ESCs | Embryonic stem cells |
| FBS | Fetal bovine serum |
| FI | Fluorescence intensity |
| GBM | Glioblastoma multifome |
| GC | Gastric cancer |
| gDNA | Genomic DNA |
| GO | Gene onotology |
| GSEA | Gene set enrichment analysis |
| HA | (Influenza) Hemagglutinin |
| HDAC | Histone deacetylase |
| HDR | Homology directed repair |
| HER2 | Human epidermal growth factor receptor 2 |
| HMEC | Human mammary epithelia cells |
| HNSCC | Head and neck squamous cell cancer |
| HRP | Horseradish peroxidase |
| HUVECs | Human umbilical vein endothelial cells |
| IF | Immunofluorescence |

| | |
|---|---|
| IFE | Interfollicular epidermal |
| IHC | Immunohistochemistry |
| iPSCs | Induced pluripotent stem cells |
| ITRs | Inverted terminal repeat sequences |
| lncRNA | Long non-coding RNA |
| M.W. | Molecular weight |
| MAPK | Mitogen-activated protein kinase |
| MBP | Maltose-binding protein |
| MC1R | Melanocortin 1 receptor |
| MDSCs | Myeloid-derived suppressor cells |
| MEFs | Mouse embryonic fibroblasts |
| mESCs | Mouse embryonic stem cells |
| MITF | Microphthalmia-associated transcription factor |
| MMP | Matrix metalloproteinase |
| MS | Mass spectrometry |
| MUPs | Melanomas of unknown primary |
| NHEJ | Non-homologous end joining |
| NK | Natural killer (cells) |
| NSCLC | Non-small cell lung cancer |
| OIS | Oncogene-induced senescence |
| ORR | Objective response rate |
| OS | Overall survival |
| PAM | Protospacer adjacent motifs |
| PDAC | Pancreatic ductal adenocarcinoma |
| PFA | Paraformaldehyde |

| | |
|---|---|
| PFS | Progression-free survival |
| PI | Propidium iodide |
| PI3K | Phosphatidylinositol 3-kinase |
| PML | Promyelocytic leukemia protein |
| PRC | Polycomb repressive complexes |
| PTC | papillary thyroid carcinoma |
| PTMs | Post-translational modifications |
| RARE | Retinoic acid response element |
| RFS | Recurrence-free survival |
| RGP | Radial growth phase |
| ROS | Reactive oxygen species |
| RT | Reverse transcription |
| rtTA | Reverse tetracycline transactivator |
| RUNX3 | Runt-related transcription factor 3 |
| SAN | Sinoatrial node |
| SCC | Squamous cell carcinoma |
| SDS | Sodium dodecyl sulfate |
| sgRNA | Single-guide RNA |
| STK19 | Serine/threonine-protein kinase 19 |
| TAC | Tacrolimus |
| TBE | T-box binding element |
| Tet-on | Tetracycline-controlled transcriptional activation |
| TIS | Therapy-induced senescence |
| TRE | Tetracycline response element |
| TSS | Transcription start site |

| | |
|---|---|
| UMS | Ulnar-mammary syndrome |
| UTR | Untranslated region |
| UVC | Ultraviolet C |
| VGP | Vertical growth phase |
| WB | Western blotting |
| WREs | Wnt-responsive-elements |
| β-ME | β-mercaptoethanol |

# 1

# Introduction

## 1.1 Melanoma

### 1.1.1 Cancer

The rising incidence of cancer worldwide is costing millions of lives and causing an increasingly heavy burden to healthcare systems. In 2018, an estimated 17 million new cases of cancer occurred worldwide, and an estimated 9.6 million people died from cancer in the same year (Cancer Research UK, accessed Apr/2020). In the UK, there are 367,167 new cases of cancer diagnosed from 2015 to 2017, causing 164,901 deaths (Cancer Research UK, accessed Apr/2020), and the number of people living with cancer is rising by 3% every year (Macmillan cancer statistics February 2019). In 2017, cancer accounts for 28% of deaths in UK (Cancer Research UK, accessed Apr/2020).

While early stage cancer may be curable through surgery, metastatic disease remains difficult to treat despite years of research and generations of different therapeutic strategies. Treatment failure can be largely attributed to the complexity of the tumourigenesis, including both genetic and phenotypic heterogeneity of tumour cells and the immunosuppressive tumour microenvironment. Better understanding of the cellular and molecular mechanisms underpinning the malignancy can pave the way for new cancer treatments.

### Hallmarks of cancer

The molecular mechanisms underlying cancer progression are incompletely understood, with the pathology of cancer varying depending on the tissue of origin, complicating a general discussion of this disease. Hallmarks of cancer were proposed in two reviews (Hanahan and R. A. Weinberg 2000; Hanahan and R. A. Weinberg 2011), and described several common traits for this multifaceted disease (Figure 1.1).

**Sustaining proliferative signalling** Arguably the most fundamental hallmark of cancer is the ability to sustain proliferative signalling. While the growth-promoting signals in normal tissue are fine-tuned to maintain a homeostasis of cell number and functional tissue architecture, these signals are deregulated in cancer to allow excessive proliferation. Cancer cells might do so by responding to the growth factors secreted by themselves, signalling to neighbouring normal stromal cells to provide a tumour-supportive microenvironment, as well as elevating cell surface receptor number or changing receptor structure to facilitate the ligand binding (Janku et al. 2018). Alternatively, somatic mutations in downstream effectors causing constitutive activation of signalling pathways can promote cell proliferation without the need of growth factor stimuli or receptor activation. Signalling pathways that are often hyperactivated in cancer include the RAF-mitogen-activated protein kinase (MAPK) and phosphatidylinositol 3-kinase (PI3K)–AKT signalling, caused by activating mutation of oncogenic BRAF, NRAS or PI3K, or loss-of-function

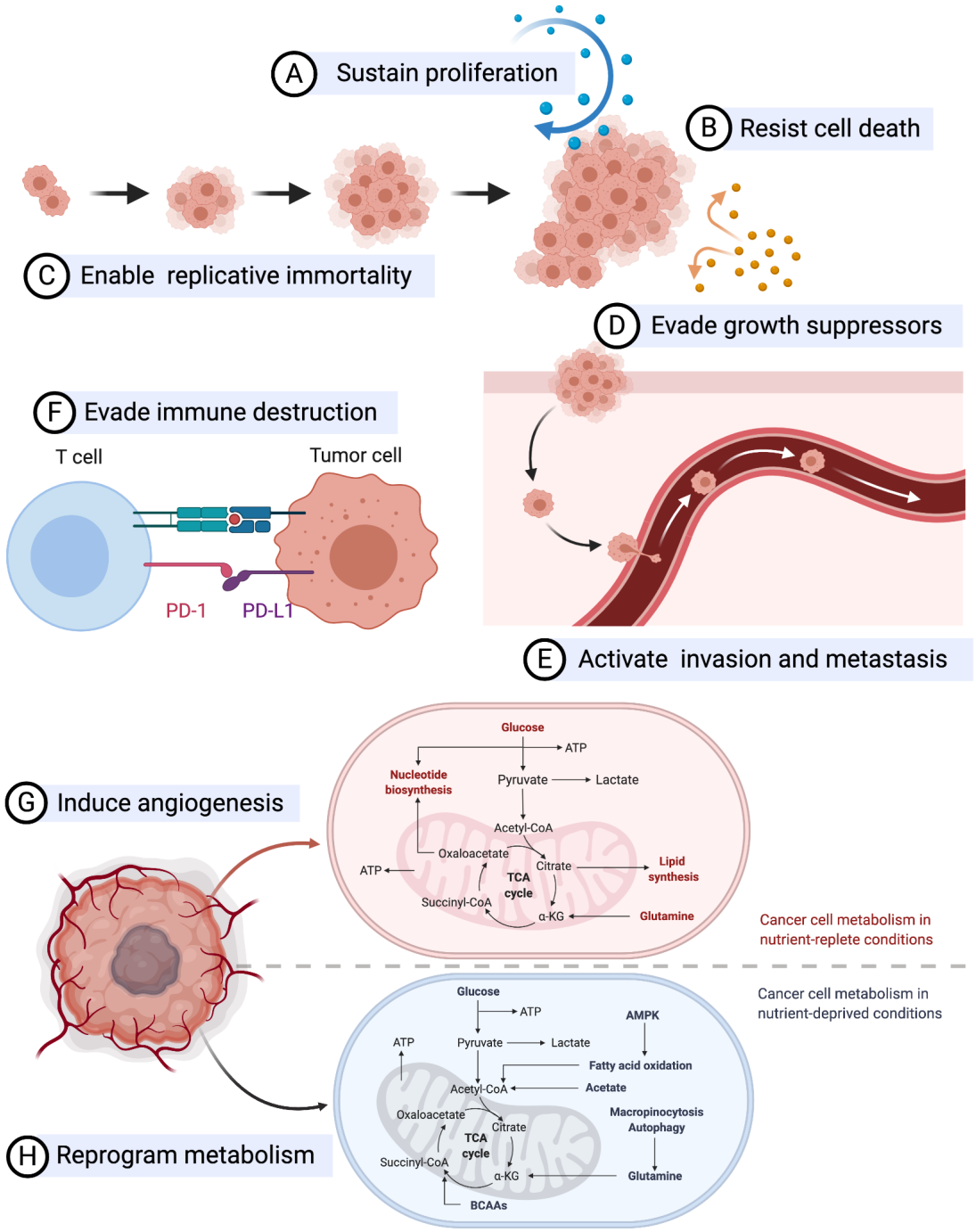



**Figure 1.1: The hallmarks of cancer** Adapted from Hanahan and R. A. Weinberg 2011, created with BioRender.com

mutations of the tumour suppressor PTEN, a negative regulator of PI3K signalling (Karoulia et al. 2017).

**Evading growth suppressors** After proliferative signalling is activated, cancer cells must then be able to circumvent fail-safe mechanisms triggered by inappropriate

cell growth. For example, activating mutations in oncogenes alone can give rise to senescence, thus tumour progression must be accompanied by other events to circumvent this barrier, in particular loss or inactivation of tumour suppressor genes (Courtois-Cox et al. 2008).

Senescence is defined as a stable cell cycle arrest, induced at the end of the cellular lifespan (replicative senescence), or in response to different stresses. Early oncogenic mutations can lead to a pre-malignant lesion where DNA damage accumulation leads to p53 up-regulation and a p21-dependent senescence programme. This oncogene-induced senescence (OIS) has a crucial role in preventing tumour progression. Pathways involved in OIS induction include K-Ras activation, $BRAF^{V600E}$ mutation and inactivation of PTEN. OIS can be bypassed and tumour progression can be accelerated when senescence is disabled, for example, when loss of *TP53* or *CDKN2A* occurs (Sharpless and Sherr 2015). In addition, conventional chemotherapy and ionizing radiation, which cause extensive DNA damage, can also induce senescence (therapy-induced senescence, TIS) (Acosta and Gil 2012). As senescence depends on the Rb1 and p53 related pathways, additional mutations inactivating the tumour suppressor network might lead to therapy resistance.

Apart from the senescence, contact inhibition is another mechanism to obviate overpopulation and maintain tissue homeostasis. Tumour suppressors E-cadherin, Merlin and LKB1, which regulate contact inhibition by strengthening cell-cell adhesion and reduce proliferative signals by occupying growth factor receptors, are often lost or suppressed (Mendonsa et al. 2018; Curto and McClatchey 2008; Momcilovic and Shackelford 2015). In addition, the normal anti-proliferative TGF-β signalling can be hijacked in cancer to promote an epithelial-mesenchymal transition (EMT) that is frequently associated with invasion (Colak and Dijke 2017).

**Resisting cell death** Hyperproliferation often results in DNA damage and nutrient supply-demand imbalance, which might trigger programmed cell death. Cancer cells have evolved ways to avoid cell death via a range of mechanisms.

Pro-apoptotic factors Bax and Bak disrupt the mitochondrial membrane to release pro-apoptotic signals, subsequently activating a caspase cascade to trigger apoptosis. Bax and Bak are usually bound by and repressed by Bcl-2 regulator proteins. The apoptotic machinery might be limited in cancer cells through loss of the TP53 stress sensor, up-regulation of pro-survival Bcl-2 or down-regulation of Bax and Bak (K. J. Campbell and Tait 2018). Autophagy, which can be induced by cellular stresses like nutrient deficiency, enables cells to break down and recycle macromolecules for biosynthesis and energy production when resources are limited. While loss of Beclin-1, a protein necessary for autophagy induction, increases susceptibility to tumour formation in mice, autophagy induced by cytotoxic drugs or nutrient limitation can be protective to cancer cells under stress conditions (Qu et al. 2003). Similarly, necrotic cell death on the one hand limits tumour proliferation, but on the other hand releases pro-inflammatory signals and regulatory factors that can in turn promote angiogenesis, invasiveness and cancer growth (Z.-g. Liu and Jiao 2020).

**Enabling replicative immortality** In contrast to cancer cells that have unlimited proliferative potential, the majority of normal cells are only granted few division cycles (around 40-60 times for human cells) before they enter senescence, a typically irreversible cell cycle arrest, or undergo programmed cell death (Hayflick 1965; Shay and Wright 2000). This phenomenon is termed the Hayflick limit. The replicative potential of cells largely depends on the length of telomeres, which protect chromosome ends and shorten during cell proliferation. Telomerase, which extends telomeric DNA and thus counteracts the telomere attrition that occurs during cell propagation, is expressed and active in most cancers, but almost absent in non-immortalised cells. While the absence of telomerase increases chromosome alterations and renders pre-malignant lesions susceptible to senescence, this might also create a high mutation burden environment that could eventually give rise to cells with telomerase expression and oncogene activation through clonal selection. Thus, the telomere shortening has opposing effects in tumourigenesis (Maciejowski

and Lange 2017).

**Inducing angiogenesis** As a deregulated growing cell mass, a solid tumour requires a sufficient of nutrient and oxygen supply as well as timely removal of metabolic waste, for which neovasculature formation is needed. Angiogenesis is controlled by inducers like vascular endothelial growth factor-A (VEGF-A) and inhibitors like thrombospondin-1 (TSP-1), and can be induced by oncogenes or TGF-β signalling, or switched on by tumour-infiltrating immune cells or hypoxia. Tumour neovasculature is also supported by pericytes, which maintain endothelial cells in normal vasculature (Lugano et al. 2020).

**Activating invasion and metastasis** As malignant tumours grow, some cells might disseminate from the primary site and colonise distant tissues. In the invasion-metastasis cascade, cancer cells typically undergo local invasion, intravasation into nearby blood and lymphatic vessels and then after circulation, extravasation into parenchyma of distant tissues to form microstastases, and finally outgrowth into macroscopic tumours (Fares et al. 2020). Key factors in these migratory programmes include E-cadherin that forms adherens connecting adjacent cells, and N-cadherin that is normally expressed in migrating or mesenchymal cells (W. Yu et al. 2019).

Cells might migrate using different mechanism, for example as mesenchymal or amoeboid cells, in either collective or single cell pattern. In many cases, the mesenchymal mode of invasion is regulated by epithelial-mesenchymal transition (EMT) (Krakhmal et al. 2015). EMT is an important programme during embryogenesis and has been implicated in cancer invasion. Included among the EMT-regulating transcription factors are Snail, Slug, Twsit and Zeb1/2, which can directly suppress expression of E-cadherin leading to loss of adherens junctions, and activate expression of extracellular matrix-degrading enzymes to clear the path for migrating cells. Several stromal cells in the tumour microenvironment, such as mesenchymal stem cells and tumour-associated macrophages, can be activated by signals secreted by tumour cells and reciprocally stimulate the invasive behaviour

of cancer cells (Fares et al. 2020).

After dissemination of the primary tumour site, circulating tumour cells then lodge at distant tissues. In several types of cancer, such as breast cancer and melanoma, disseminated cancer cells can remain dormant for years, due to lack of nutrient or angiogenesis-inducing signal before finally adapting to and colonising the distant tissue (Hen and Barkan 2019).

**Reprogramming energy metabolism** For tumour cells to sustain abnormal persistent growth and survive in an often nutrient-deprived environment, reprogramming of cellular metabolism is paramount and consistent with multiple alterations in metabolism observed in cancer. For example, cancer cells exhibit deregulated uptake of glucose and glutamine, two important building blocks for biosynthesis of macromolecules, mediated by increased expression of the glucose transporter GLUT1, and the glutamine transporters ASCT2 and SN2. Increased consumption of glucose and amino acids provide metabolic resources for tumour growth. As a result, however, tumour cells are often exposed to nutrient-deprived conditions. To address the need for anabolic precursors, tumours may activate opportunistic ways of nutrients acquisition. For example, cancer cells experiencing depleted supplies can recover free amino acids by degrading macromolecules captured through macropinocytosis, entosis and phagocytosis, as well as by obtaining free fatty acids directly from plasma or neighbouring stromal cells such as adipocytes and fibroblasts. Even not under extreme conditions, cancer cells can utilise autophagy to liberate and reuse free amino and fatty acids (Pavlova and Thompson 2016).

Apart from changes in nutrient uptake and resources, cancer cells also alter how nutrients are used. For example, in cancer cells, much of the pyruvate generated from glycolysis is directed away from TCA cycle to create lactate, despite possession of functional mitochondria and presence of oxygen, which is known as the Warburg Effect. Although explanations for this phenomenon remains elusive, several proposed mechanisms converge on the advantage of maintaining large pools of glycolytic

intermediates for various biosynthetic processes (DeBerardinis and Chandel 2020). In addition, cancer cells have an increased demand for nitrogen and thus more dependence on glutamine and arginine as nitrogen donor, and some cancer types display addiction to glutamine, despite it is a nonessential amino acid that can be synthesised from glutamate. For example, in melanoma, glutamine limitation triggers cell cycle arrest (Falletta et al. 2017). Furthermore, an imbalance of nutrient or oxygen demand and supply in the tumour microenvironment can trigger cancer invasion, which is typically mediated by eIF2α phosphorylation-dependent translation reprogramming (Falletta et al. 2017; García-Jiménez and Colin R. Goding 2018).

**Evading immune destruction** Under normal conditions immune surveillance can effectively detect and eliminate infected or neoplastic cells whilst preserving self-tolerance. Thus, for tumours to form they must evade immune destruction. Negative regulators of T cell activation act as "checkpoint molecules" to prevent hyperactivation of immune response, and some of them are implicated in aiding cancer cells to avoid immune elimination (Havel et al. 2019). Arguably the most prominent are cytotoxic T lymphocyte antigen 4 (CTLA4) and programmed cell death (PD1), which are cell surface receptors belonging to immunoglobin superfamily. CTLA4 and PD1 are expressed on the surface of T cells and bind to B7 and PDL1/2 ligands respectively, which are expressed on antigen-presenting cells, and thus inhibit T cell activation and proliferation. Overexpression of B7 and/or PDL1/2 in cancer cells suppresses $CD8^+$ T cell cytotoxic response and thus protects cancer cells from immune destruction. Blockade of these checkpoints using neutralising antibodies against CTLA4, PD1 or PD-L1/2 enhances anti-tumour immunity and confers therapeutic benefit (Waldman et al. 2020). Immune oncology, thus opens up a new era for cancer treatment.

### Enabling characteristics

Underlying the above traits governing cancer transformation are crucial characteristics that enable the acquisition of these hallmarks, included among which are genome instability and an inflammatory microenvironment.

**Genome instability and mutation** Normal cells have a machinery to maintain the integrity of the genome and are capable of detecting and resolving mutations occurring during cell replication. On the other hand, for tumours to evolve, a succession of genome alterations and/or epigenetic modifications is required to give them selective advantages under abnormal and stressful conditions. In this sense, genome instability is an important attribute of cancer progression. Many tumour suppressors, like TP53, are involved in maintaining DNA integrity, and they are often inactivated or lost in cancer cells through genomic mutations, epigenetic repression or chromosomal segment amplification/deletion (Olivier et al. 2010). The subsequent genomic instability may eventually give rise to more mutations facilitating clonal expansion of cells that are best adapted to different stages of tumour progression or resistant to therapies.

**Tumour-promoting inflammation** Tumours are infiltrated by immune cells, resembling inflammatory conditions. While inflammation can indicate immune responses attempting to kill tumour cells, paradoxically it can also boost tumourigenesis. Inflammation can release factors promoting growth, survival, angiogenesis and invasion, as well as mutagenic agents like reactive oxygen species (ROS). Tumour-promoting immune cells include M2 (alternative-activated) macrophages (Lin et al. 2019) and myeloid-derived suppressor cells (MDSC) (V. Kumar et al. 2016), which can suppress tumour-antagonising $CD8^+$ cytotoxic T lymphocytes (CTLs) or natural killer (NK) cells.

Though mechanisms underpinning progression of different types of cancer can vary, the hallmarks and relevant enabling characteristics provide a macroscopic view

of multifaceted cancer. In my research project, I used cutaneous melanoma cells as a model to study key proteins implicated in tumour progression, how they are regulated and how they regulate other genes to facilitate the pathological process. Hence, I will go into more details of melanoma progression and current therapies.

### 1.1.2 Melanoma progression

Cutaneous melanoma is a type of skin cancer originating from melanocytes, which locate within the basilar epidermis and produce melanin for neighbouring keratinocytes. Apart from the epidermis, melanocytes are also found in other hair follicles, the uveal tract of the eyes, meninges and the anogenital tract. Accordingly, melanoma subtypes originating froono these melanocytes differ in their cells of origin, precursor lesion and mutations patterns (Costin and Hearing 2007).

In this subsection, I discuss tumour progression stages of and relevant pathogenic mutations underpinning cutaneous melanoma, the melanoma subtype used as a model in my research project.

**Epidermal melanocytes** In response to UV radiation, in particular UVB (280–320 nm), keratinocytes produce α-melanocyte stimulating hormone (α-MSH), which binds to the melanocortin 1 receptor (MC1R) on melanocytes (Brenner and Hearing 2008). MC1R in turn signals to activate the master regulator of pigmentation, microphthalmia-associated transcription factor (MITF), and the subsequent production of melanin (Williams et al. 2011; García-Borrón et al. 2005). Melanin produced by epidermal melanocytes absorbs and scatters UV radiation and thus protects keratinocytes from UV radiation-induced DNA damage and ROS (Kaidbey et al. 1979).

**Melanocytic naevi** Naevi are considered a benign neoplasm of melanocytes, as they are not likely to progress into melanoma. After acquiring an initiating mutation, with $BRAF^{V600E}$ being the most common, melanocytes will proliferate

and then enter a senescent status (Yeh et al. 2013; Michaloglou et al. 2005). At this stage, a small portion of melanocytes are still able to proliferate, however, this might be counteracted by apoptosis induced by oncogenic stress and elimination by infiltrating immune cells (Shain and Bastian 2016). Thus, the size of naevi remain the same for years.

**Dysplastic naevi** Before turning malignant, naevi undergo an intermediate stage, dysplastic naevi, which have grown bigger in size and have variable pigmentation and irregular borders. Compared to benign naevi, which only have a single driver mutation such as $BRAF^{V600E}$, the intermediate neoplasm can acquire additional mutations, including *TERT*, encoding for a major component of telomerase (Shain, Yeh, et al. 2015) that contributes to longer lifespan. Unlike benign naevi, multiple dysplastic naevi correlate with increased risk of developing melanoma (Shors et al. 2006).

**Melanoma *in situ*** The irregular proliferation of melanocytes with large nuclei locating entirely within the epidermis is termed melanoma *in situ* (Figure 1.2). While this can be also observed at the edges of invasive melanoma, free-standing melanoma *in situ* is considered the earliest stage of melanoma (Balch et al. 2009). Precursors of melanoma *in situ* differ depending on the role of UV radiation as a pathogenic factor. Chronic sun-induced damage (CSD) melanoma *in situ* has a lentiginous growth pattern, and typically arises *de novo* without a premalignant lesion. On the other hand, melanoma *in situ* arising from naevi has pagetoid features, and is associated with non-CSD melanoma (Shain and Bastian 2016). Apart from the initiating mutations activating the MAPK pathway such as *BRAF*, *NF1* and *NRAS*, and *TERT* promoter mutations in dysplastic naevi, melanoma *in situ* also harbour heterozygous mutations in the *CDKN2A* locus, which encodes two cell cycle inhibitors $p16^{INK4A}$ and $p14^{ARF}$ (Shain, Yeh, et al. 2015).

**Invasive melanoma** Melanoma *in situ* can persist for years before turning invasive — meaning melanoma cells leave the epidermis and enters subjacent dermis.

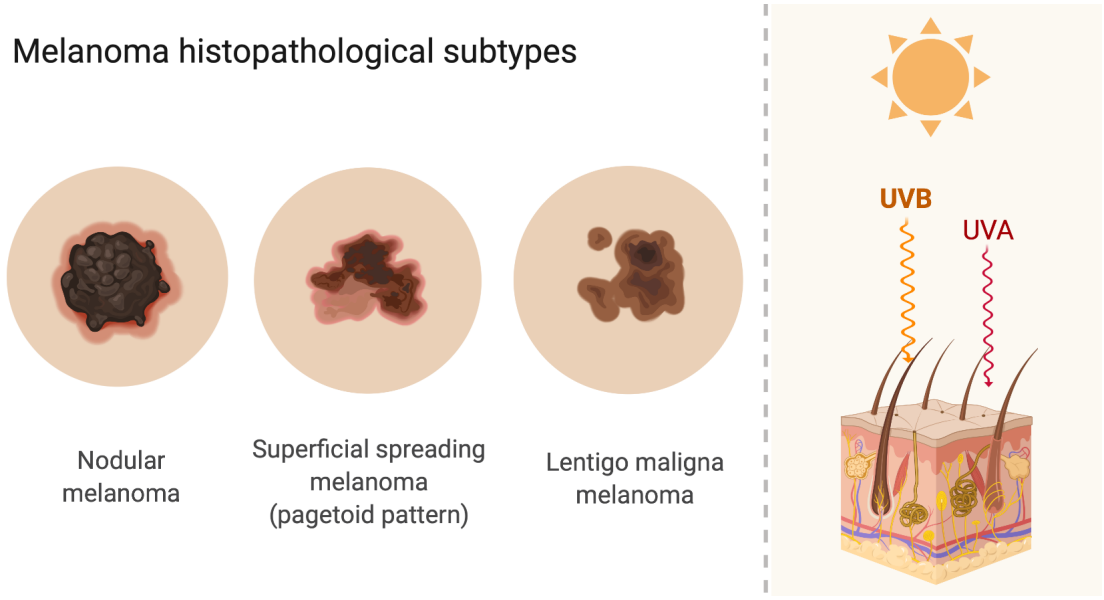



**Figure 1.2: Histopathological subtypes and cause of melanoma** Left panel shows patterns of nodular melanoma and melanoma *in situ* with pagetoid or lentigo features. Right panel shows ultraviolet (UV) radiation, namely UVA (315–400 nm) and UVB (280–315 nm), is a cause of melanoma. Adapted from Shain and Bastian 2016 and Sample and Y.-Y. He 2018, created with BioRender.com

Most invasive melanomas arise from a melanoma *in situ*, with the exception of nodular melanomas, which can arise *de novo* from normal-looking skin, typically progress fast and are overrepresented in lethal melanoma (Menzies et al. 2013). Bi-allelic inactivation of *CDKN2A* is evident exclusively in invasive melanomas (Shain, Yeh, et al. 2015). $p16^{INK4A}$, one of the *CDKN2A* locus products, is a crucial barrier to the transition to invasiveness, as its loss correlates with invasive melanoma (Talve et al. 1997). In addition, the chromatin remodelling SWI/SNF complex is also implicated as a tumour suppressor in melanoma by maintaining genomic integrity, and disabling mutations affecting its members *ARID2* and *ARID1A* occur at the invasive stage (X. Wang et al. 2014). Furthermore, mutation frequencies of *TP53* and *PTEN* are higher in metastases in comparison to primary melanoma, suggesting that mutations of *TP53* and *PTEN* emerge at later stages of invasive melanoma (Shain, Yeh, et al. 2015).

**Metastatic melanoma** Typically melanoma metastases first appear at lymph

nodes near the primary site, followed by distant metastases at visceral sites. Though this observation might suggest a serial progression of melanoma metastasis, counter-intuitively, multiple lines of evidence indicate distant metastases arise in parallel with primary melanoma. For instance, circulating tumour cells are found in patients without any metastases (Reid et al. 2013); phylogenetic analysis of matched metastatic and primary melanomas also confirmed the parallel dissemination in most patients (Sanborn et al. 2015). Strikingly, even melanocytes from benign neoplasms are capable of dissemination and limited colonisation, giving rise to benign metastases (Carson et al. 1996). This might explain the origin of melanomas of unknown primary (MUPs) — apparent metastases without detectable primary lesions (Anbari et al. 1997). Supporting the hypothesis that MUPs arise from disseminated melanocytic naevi, which are considered to be benign, MUPs have a better prognosis compared to bona fide metastatic melanomas (Anbari et al. 1997). In addition, genetic and phenotypic heterogeneity between metastastic lesions and subclones within a metastatic site might harbour resistant cells and thus promote fast relapse after targeted therapies.

### 1.1.3 Melanoma therapy

#### Surgery and chemotherapy

Surgical removal is a gold standard treatment for cutaneous melanoma patients with localised disease (patients with negative lymph nodes) or locoregional disease (patients with positive regional lymph nodes) (Schadendorf et al. 2015). In the latter scenario, the remaining negative lymph nodes in the area are sometimes removed to prevent further metastases (L. E. Davis et al. 2019).

The first and only chemotherapy approved for melanoma treatment is Dacarbazine, which was approved by FDA in 1975 (L. E. Davis et al. 2019). Until recently, Dacarbazine served as the standard treatment for inoperable metastatic melanoma

(Ugurel et al. 2013), though the survival benefit from this drug is quite modest with an objective response rate (ORR) of 5-6% and a progression-free survival (PFS) of 1.6-2.7 months (Chapman et al. 2011; McArthur et al. 2014).

### Targeted therapy

Following the discovery of a high $BRAF^{V600E}$ mutation rate in melanoma patients, the era of BRAF-targeting therapies came. The first-in-class selective $BRAF^{V600E}$ inhibitor, Vemurafenib, was approved by FDA for treatment of metastatic melanoma in 2011, followed by Dabrafenib, which was approved in 2013. Compared with Dacarbazine, Vemurafenib and Dabrafenib monotherapies largely increased ORR to 48% and 50% respectively, and improved PFS to 5 months (Chapman et al. 2011; McArthur et al. 2014).

As BRAF signalling depends on the activation of downstream MEK1/2 (Figure1.3), inhibiting MEK1/2 might synergise with BRAF inhibitors. MEK inhibitors, Trametinib, Cobimetinib and Binimetinib, were developed and tested for melanoma treatment. While Trametinib alone only demonstrated modest therapeutic benefit (Keith T Flaherty et al. 2012a), a combination of Dabrafenib and Trametinib achieved an ORR of 76% and a PFS of 9,4 months (Keith T. Flaherty et al. 2012b). Similar synergy was observed between Vemurafenib and Cobimetinib (Ribas et al. 2014).

### Immunotherapy

Interleukin-2 (IL2) was the first approved immune therapy for metastatic melanoma. However, the therapeutic benefit is modest with an overall response rate of 18% and median overall survival of 9.6 months (Atkins et al. 1999; Davar et al. 2017).

More recently, with the understanding of mechanisms and key players underpinning tumour immune evasion, immunotherapy strategies have focused on overcoming the suppression on immune system imposed by the tumour microenvironment.

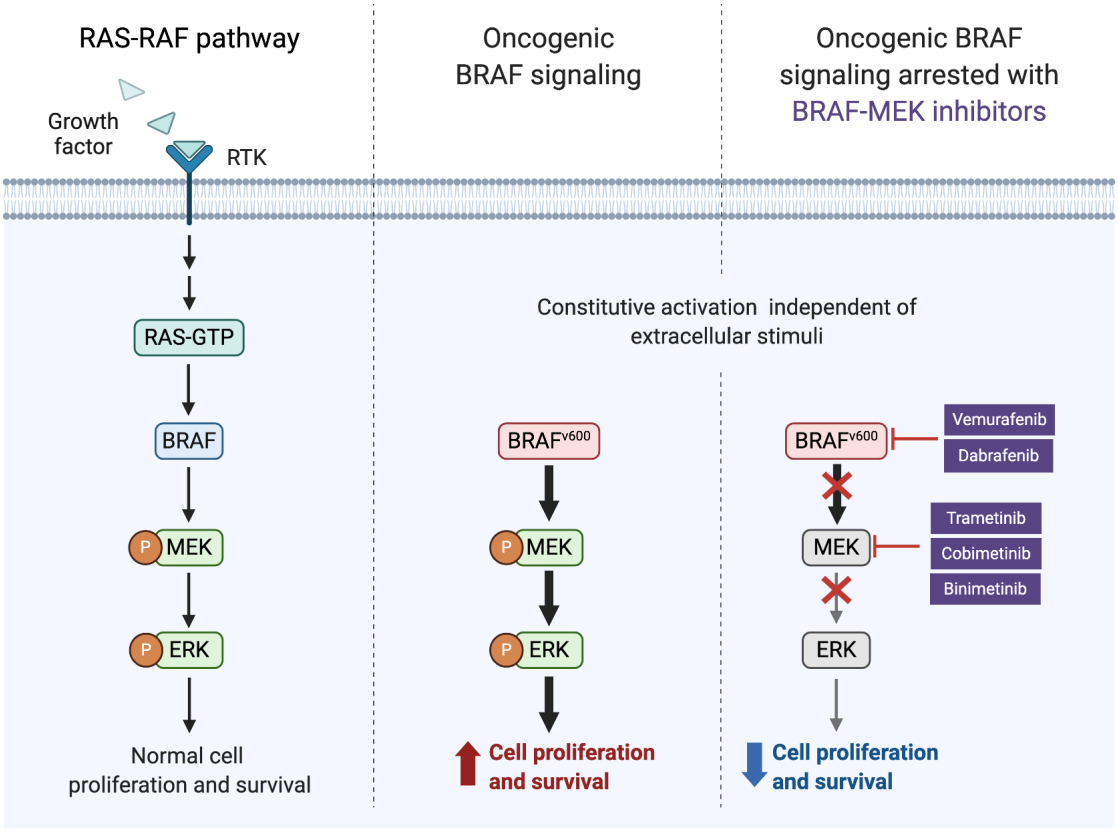



**Figure 1.3: BRAF-MEK signalling pathway in melanoma** Adapted from Schadendorf et al. 2015, created with BioRender.com

Since 2011, immune checkpoint inhibitors have gained regulatory approval for melanoma therapy, including anti-CTLA4 Ipilimumab, anti-PD1 Nivolumab and Pembrolizumab. Clinical trial using Ipilimumab as a single agent or in combination with Dacarbazine showed ORRs of 11–15%, a median PFS of 3 months (Hodi et al. 2010; Robert et al. 2011). Similar results were obtained in clinical trials of Pembrolizumab and Nivolumab, demonstrating ORRs of 20-40% (Hamid et al. 2013; Topalian et al. 2014). The combination of anti-CTLA4 and anti-PD1 antibodies improved clinical efficacy but also increased adverse effects compared to monotherapy alone (Larkin et al. 2015; Wolchok et al. 2013; Luke et al. 2017).

Most recently, talimogene laherparepvec (T-VEC) was approved as the first oncolytic virus therapy for advanced melanoma in 2015. T-VEC is a modified herpes virus that selectively proliferates in tumour cells to evoke immune response, and

produces granulocyte-macrophage colony-stimulating factor (GM-CSF) to induce antigen-presentation by macrophages and thus further enhance T-cell recognition of virus-infected tumour cells (Dranoff 2002). Compared to administration of GM-CSF, use of T-VEC increased durable response rate from 2% to 16% (Andtbacka et al. 2015). In addition, compared to other systemic therapies T-VEC showed modest toxicities (Luke et al. 2017).

**Emerging new strategies**

Apart from the above therapies that have obtained regulatory approval, there are also other therapeutic strategies proposed by new pre-clinical studies. For example, in melanoma the prevailing NRAS mutation (Q61R/K/L) traps NRAS in a constitutive active conformation and recruits RAF for dimerisation and activation. Given the importance of oncogenic $NRAS^{Q61R/K/L}$ mutation in melanomagenesis, there has been little success in developing selective inhibitors, partly because RAS GTPases generally lack deep hydrophobic pockets crucial for small compound binding (O'Bryan 2018). Nevertheless, a small molecule specifically inhibiting $KRAS^{G12C}$ mutant, Sotorasib (AMG 510), has been developed. In preclinical analyses, Sotorasib treatment reduced tumour growth of patient-derived xenografts (Canon et al. 2019). Even more encouraging, phase I clinical trial revealed that 32% of patients with non-small cell lung cancer responded to Sotorasib treatment (Hong et al. 2020).

In addition, recently a report showcased a new pharmacological approach to treat NRAS-driven melanoma. The authors set out to identify upstream kinases that regulate NRAS activity and target the kinase to treat NRAS-mutant melanoma. Human kinome siRNA library screening identified serine/threonine-protein kinase 19 (STK19) as NRAS regulator. TCGA data showed that STK19 has an alteration rate of 25% in melanoma cases, the most of which are gain-of-function mutations, suggesting a role for STK19 in melanoma progression. STK19 overexpression

and STK19$^{D89N}$ enhance melanomagenesis induced by oncogenic NRAS$^{Q61R}$. The authors developed a potent STK19 inhibitor, ZT-12-037-01 (1a), which inhibits SK-MEL-2 (with NRAS$^{Q61R}$ mutation) cell proliferation and induces apoptosis *in vitro* and *in vivo*, thus blocks melanoma development and growth (C. Yin et al. 2019).

## 1.2 Gene regulation by transcription factors

### 1.2.1 Transcriptional control of genes

The human genome is thought to contain around 20,000 protein-coding genes and many non-coding genes (Piovesan et al. 2019). Transcription describes the process of synthesising RNA based on a DNA template, catalysed by RNA polymerase (Figure 1.4), and can be divided into the following steps (Renkawitz 2006; T. I. Lee and Young 2013).

1. **Formation of a preinitiation complex** Transcription factors bind to enhancer elements and recruit coactivators, which bind to RNA polymerase II, which in turn binds to general transcription factors at the core promoter. The DNA loop formed between the enhancer and transcription start site (TSS) is stabilised by cofactors such as the Mediator complex and cohesin.
2. **Initiation and pausing** RNA polymerase II initiates transcription, and then pauses after transcribing a short distance. This process is regulated by pause control factors DSIF and NELF that are physically associated with RNA polymerase II.
3. **Pause release and elongation** Transcription factors and cofactors recruit elongation factors such as P-TEFb, which phosphorylates the pause control factors and the paused polymerase, allowing elongation to proceed.
4. **Termination** A fully mature RNA transcript is formed and then released from the RNA polymerase and the DNA template.

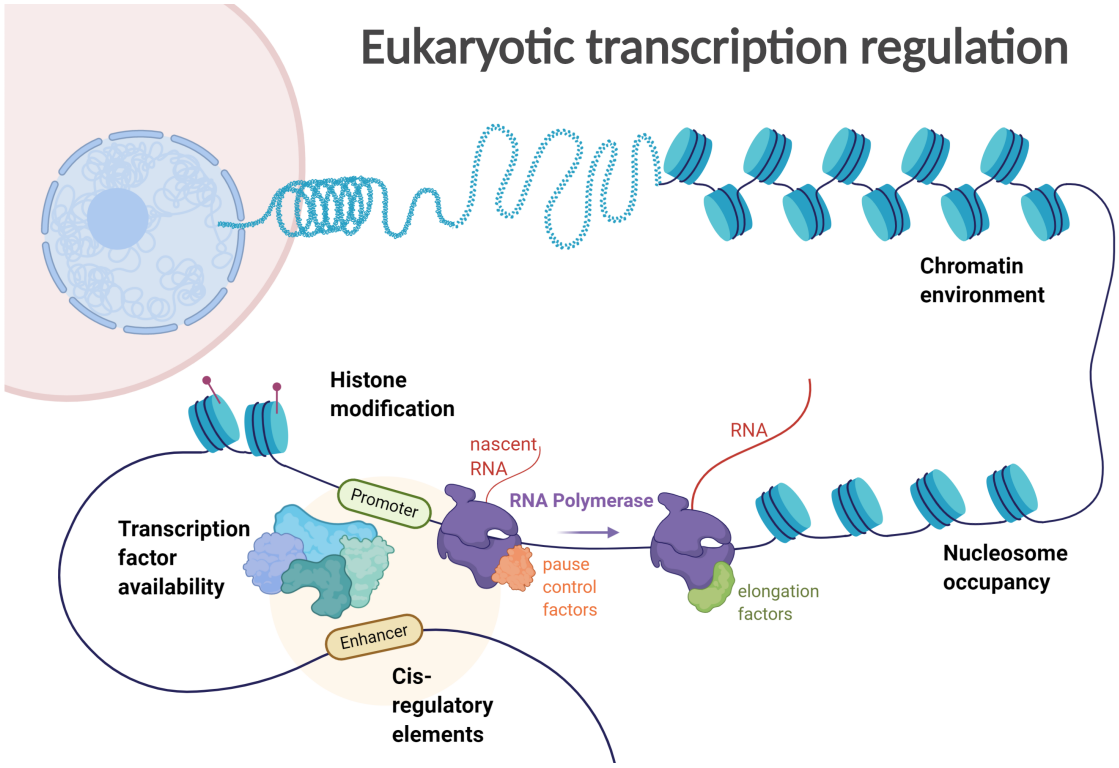



**Figure 1.4: Regulation of transcription in eukaryotic cells** Adapted from T. I. Lee and Young 2013, created with BioRender.com

In addition to transcription factors, another level of transcription regulation involves chromatin and its regulators (T. I. Lee and Young 2013). The fundamental unit of chromatin is nucleosomes, the formation of which is paramount to eukaryotic genome packaging. In a canonical nucleosome, a 147-bp of DNA segment is wrapped around the histone octamer for a total of 1.65 turns (F. Zhu et al. 2018). To activate a gene, the transcription apparatus requires access to DNA, for which mobilisation or modification of nucleosomes is crucial. For instance, SWI/SNF family chromatin remodelling complexes can be recruited to active genes and facilitate access of the transcription apparatus to DNA by mobilising nucleosomes (Clapier and Cairns 2009). In addition, histone-modifying enzymes are recruited to transcription regulatory sites and alter stereotypical nucleosome organisation to provide interaction surfaces for other proteins (Bannister and Kouzarides 2011). Conversely, histone deacetylation and HP1-induced heterochromatin formation

result in a more compact chromatin structure rendering DNA less accessible to transcription machinery (Renkawitz 2006).

### 1.2.2 Transcription factors

#### Definition and characteristics

The term "transcription factors" used to stand for any protein involved in transcription or gene regulation, but now transcription factors are defined by two properties, their ability to bind DNA in a sequence-specific manner and regulate transcription (Vaquerizas et al. 2009).

Transcription factors bind to DNA via their DNA-binding domains (DBDs), and most eukaryotic transcription factors also recruit cofactors via effector domains. There are around 100 known classes of eukaryotic DBDs based on their structural properties, including zinc finger domain, homeodomain, basic helix-loop-helix (bHLH) and basic leucine zipper (bZIP) (Finn et al. 2016). While the majority of transcription factors only have a single DBD, there are a few with multiple DBDs, either homotypic or heterotypic (Lambert et al. 2018). On the other hand, dedicated effector domains in transcription factors are responsible for recruiting specific cofactors. For example, the KRAB domain is found in around 350 human C2H2-zinc finger proteins and recruits TRIM28, a scaffold protein that recruits mediators of heterochromatin formation (Ecco et al. 2017).

#### Cooperation and effectors

Few eukaryotic transcription factors can directly recruit RNA polymerase, however, most regulate transcription via interaction with cofactors (Lambert et al. 2018).

The cooperative binding between transcription factors can be either a result of direct protein-protein interaction or mediated by DNA. Interaction between transcription factors can modify their DNA-binding specificity (Slattery et al.

2011), or can stabilise the DNA-bound complex (Chen et al. 2014). Another type of transcription factor cooperativity is mediated by DNA. In some cases, the DNA between the proteins is deformed, facilitating transcription factor interaction; alternatively, the binding of the first transcription factor changes the shape of DNA in a way that promotes the binding of a second transcription factor (Jolma, Y. Yin, et al. 2015). In the latter scenario, transcription factors may bind to the opposite side of DNA or relatively far from each other (Lambert et al. 2018).

In addition to cooperative binding, transcription factors can also collaborate by impacting chromatin state and transcription via different mechanisms. This synergistic regulation can be exemplified by pioneer transcription factors like FoxA, which accesses silent chromatin and displaces linker histone H1, rendering enhancer nucleosomes more accessible to other transcription factors to bind and stimulate transcription (Iwafuchi-Doi et al. 2016).

## 1.3 The T-box transcription factor family

Transcription factors of the T-box family have crucial roles in embryonic development. A study of the short-tailed mouse identified a mutation affecting tail development in 1927 and led to the discovery of the *Brachyury* gene (from the Greek for short tail) in the 1990s, which is the founding member of the family also known as *T* (Papaioannou 2014; Wansleben et al. 2014). Phylogenetic analysis shows that there are five subfamilies of T-box factor: T, Tbx1, Tbx2, Tbx6 and Tbr1 (Figure 1.5), consisting of 17 paralogues in human, mouse and rat (Khan et al. 2019).

The T-box transcription factors are characterised by a highly conserved DNA-binding domain known as the T-box domain. The target sequence for T-box domain in T was identified as a partially palindromic sequence (T(G/C)ACACCTAGGTG TGAAATT), made up of two half sites to which the protein binds as a monomer Kispert and Herrmann 1993 Khan et al. 2019, while other T-box proteins can bind

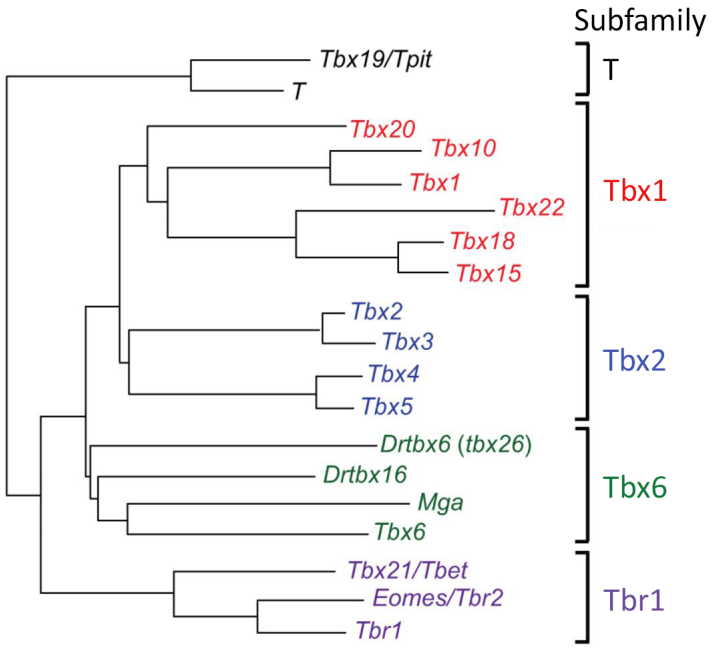



**Figure 1.5: Phylogenetic tree of the T-box transcription factors in vertebrates** Different colours indicate five subfamilies of T-box transcription factors found in vertebrates. All genes represented here except for *Drtbx16* and *Drtbx6* are present in human and mouse. Adapted from Papaioannou 2014

.

the consensus half site (AGGTGTGAAA) known as the T-box binding element (TBE), or the T-element (Papaioannou 2014).

The closely related TBX2 and TBX3 belong to the TBX2 subfamily, and they are the only mammalian T-box factors known to function mainly as transcriptional repressors (Juan Lu et al. 2010).

In this introduction chapter, the discussion will be focused on the gene and protein structure of TBX2 and TBX3, as well as their roles in cancer, development and other human diseases. Known direct target genes of TBX2 and TBX3 are summarised in section 4.1 and section 5.1 respectively, how TBX2 and TBX3 repress their target genes by recruiting different interacting factors will be discussed in section 6.1, and up-to-date knowledge about regulation of TBX2 and TBX3 is included in section 7.1.

### 1.3.1 TBX2 and TBX3 gene, mRNA splicing and protein

#### TBX2/3 gene

The human TBX2 gene has been mapped to chromosome 17q23, comprises 7 exons spanning 3.4 kb, encodes a product of 711 amino acids (Abrahams, Parker, et al. 2010). The human TBX3 gene locates on the reverse strand of chromosome 12q23-24.1 and consists of 7 exons spanning 4.7 kb (Khan et al. 2019).

#### TBX3 isoforms

Alternative splicing of TBX3 pre-mRNA giving rise to two predominant isoforms, TBX3 and TBX3+2a. TBX3 is 723 amino acids in length, while TBX3+2a has an extra exon, which results from alternative splicing of the second intron, and encodes a protein with 20 more amino acids in the T-box domain (Figure 1.6.A).

Given the variation in DNA-binding domain, one might wonder whether the two isoforms differ in their target gene repertoire and biological function. Compelling evidence showed TBX3 and TBX3+2a have similar roles and bind to common target genes (Willem M H Hoogaars et al. 2008; Krstic, Macmillan, et al. 2016), and for this reason I will use 'TBX3' to represent both isoforms in the rest of my introduction chapter unless a differential effect is emphasised.

However, there is also contradicting information. For example, Fan et al. 2004 reported that while overexpression of TBX3 immortalises mouse embryonic fibroblasts (MEFs), TBX3+2a accelerates senescence. A recent study showed that TBX3 and TBX3+2a overexpression led to differential changes in angiogenesis-related genes in breast cancer cell line, and *in vitro* assay confirmed their different ability to induce endothelial cell vascular channel formation by conditioned media (Krstic, Hassan, et al. 2019). TCGA breast cancer cohort data show that the TBX3 to TBX3+2a ratio in tumour is 8.1 fold higher than that in normal tissue, and the high ratio correlates with high-tumour grade and aggressive subtype.

Another study conducted in human mesangial cells showed while TBX3 locates in the nuclear compartment, TBX3+2a showed cytoplasmic localisation (Wensing and Campos 2014).

### TBX2/3 protein functional domains

The DNA binding domains of human TBX2 and TBX3 share 95% homology and the dominant repression domains in the C-terminus share 66% homology, though their second repression domains locate at different positions and share no homology (Khan et al. 2019).

The human TBX2 protein contains two repression domains, one in the N-terminus (1-53) and one in the C-terminus (529-573) (Figure 1.6.A). Interestingly, Paxton et al. 2002 identified a weak activation domain within the T-box domain that is capable of activating a T-element containing promoter.

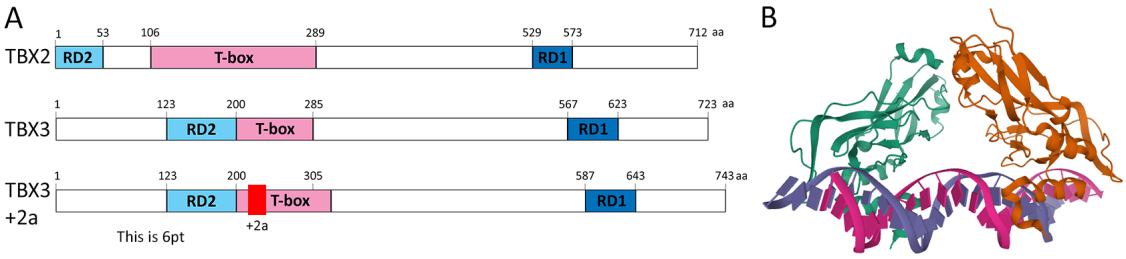



**Figure 1.6: TBX2/3 protein structure** (A) Functional domains of TBX2 and TBX3 proteins, adapted from Krstic, Kolendowski, et al. 2019; Khan et al. 2019 (B) Crystal structure of the T-box domain from human TBX3 bound with DNA (PDB accession code 1H6F)

TBX3 isoforms also have two known transcriptional repression domains. The N-terminal repression domain RD2 and the nuclear localisation signal (NLS) "RREKRK" locate within the T-box DNA-binding domain (Carlson, Ota, C. E. Campbell, et al. 2001), while the dominant repression domain RD1 is in the C-terminus (Figure 1.6.A) (Carlson, Ota, C. E. Campbell, et al. 2001). Notably, between human and mouse the T-box domain shares 100% homology, whereas the RD1 and activation domain are 98.4% and 77.5% conserved respectively

(Khan et al. 2019).

These functional domains are essential for TBX2's and TBX3's ability to repress transcription, as missense mutations or deletion of these regions affect their DNA-binding or repression ability.

For example, point mutations (R122E, R123E and A272E) within the T-box domain or deletion of the repression domain impair the ability of Tbx2 to repress $p19^{Arf}$ (Lingbeek et al. 2002). In addition, interestingly, Tbx2's transcriptional repression function is dependent on the interaction with retinoblastoma protein (Rb1). A Tbx2 L294AL296A mutant, which cannot associate with Rb1, fails to repress the p21 promoter. Even more strikingly, in Rb1-negative cells, Tbx2 activates the p21 promoter (Vance, Shaw, et al. 2010).

Similarly, while TBX3 immortalises primary MEFs, its mutant that lacks RD1 (amino acids 555-605) fails to allow the senescence escape (Carlson, Ota, C. E. Campbell, et al. 2001). Tarryn Willmer, Hare, et al. 2016 reported in chondrosarcoma cells, R130 in the DNA-binding domain is required for TBX3 to bind and repress the *CDKN1A* promoter. Tbx3 N249D and Tbx3 +2a L143P DNA-binding mutants fail to repress the *Nppa* promoter (Willem M H Hoogaars et al. 2008). Point mutants of TBX3, L143P and Y149S, cannot decrease the mouse $p19^{ARF}$ and human $p14^{ARF}$ levels and restore proliferation (Brummelkamp et al. 2002). The Tbx3 Y149S mutant fails to enhance liver tumour cell growth like its wild-type counterpart (Renard et al. 2007). A Tbx3 mutant lacking the C-terminal 300 amino acids failed to repress $p19^{Arf}$ (Carlson, Ota, Song, et al. 2002), whereas, deletion of Tbx3 N-terminal (exons 1-3) resulted in an abnormal protein that is located in the cytoplasm (Frank et al. 2013), confirming the 292-297 "RREKRK" NLS signal identified in Carlson, Ota, C. E. Campbell, et al. 2001.

**Crystal structure of the T-box domain bound with DNA**

Coll et al. 2002 published the crystal structure of T-box domain of TBX3 (101-291 preceded by N-terminal residues Met99 and Lys100). The human TBX3 monomers bind independently to the palindromic DNA duplex, and the T-box domain contacts DNA in both the major and minor grooves (Figure 1.6.B), with R130 being the only residue contacting the major groove and F279 and F283 contacting the minor groove. Within the TBX3 recognition site, AGGTGTGAAAT, only base pairs 3, 5, 8 and 9 are recognised by direct contacts. The crystal structure showcases the importance of the T-box domain in recognising the DNA sequence — TBX3 point mutations L143P and Y149S, which are found in ulnar-mammary syndrome (UMS) patients, destabilise the T-box domain and impair DNA-binding ability.

### 1.3.2 TBX2 and TBX3 in cancer

Early studies showed that TBX2 and TBX3 facilitate the immortalisation of cells and help bypass of senescence induced by oncogene activation. For example, Jacobs et al. 2000 used senescence-predisposed MEFs to screen for immortalising oncogenes and found TBX2 as a strong candidate. TBX2 immortalises MEFs and human fibroblasts and drives their proliferation by down-regulating *Cdkn2a*. TBX2 overexpression bypasses senescence induced by Myc or HRAS oncogene activation in MEFs by attenuating the induction of $p19^{Arf}$. Another study showed silencing of TBX2 in fibroblasts reduced cell proliferation and induced senescence-associated β-Gal expression (Martin et al. 2012).

A similar transforming ability was also observed using TBX3. *In vitro* colony formation assays showed that Tbx3 cooperates with oncogenic Myc or $H\text{-}Ras^{V12}$ to transform primary MEFs. Overexpression of Myc in MEFs increased the G1 fraction of cells especially under serum-limiting condition, while co-expression of Tbx3 rescued the cell cycle arrest, which is partly due to Tbx3's ability to repress

$p19^{Arf}$ (Carlson, Ota, Song, et al. 2002). Transduction of TBX2 or TBX3 in mouse embryonic striatum cells restored their proliferation and inhibited $p19^{ARF}$ expression (Brummelkamp et al. 2002).

### TBX2/3 contribute to tumour growth

Increasing evidence suggests that TBX2 and TBX3 are widely overexpressed in cancer and are important players in tumourigenesis. The cancer types where TBX2/3 are overexpressed and promote tumour growth are listed below.

**Breast cancer** Jacobs et al. 2000 first found that TBX2 contributes to senescence bypass and is overexpressed in a set of human breast cancer cells. Redmond et al. 2010 confirmed the overexpression of TBX2 in a panel of breast cancer cell lines. They also showed silencing of endogenous TBX2 or expression of a dominant negative TBX2 (dnTBX2) drastically reduced breast cancer cells proliferation.

Similar results were obtained for TBX3. Among cell lines derived from the same breast cancer patient, both TBX3 isoforms are more abundant in the invasive cell lines compared to the rest. Overexpression of TBX3 isoforms in ductal carcinoma *in situ* (DCIS)-like cells increases cell proliferation and decreases apoptosis, and promotes migration and invasion. On the other hand, however, TBX3 knockdown in invasive mammary carcinoma cells only moderately reduced cell proliferation (Krstic, Macmillan, et al. 2016). The dispensable role of TBX3 on cell growth was also observed in mammary epithelial cells (MECs) — though down-regulation of endogenous Tbx3 impairs growth of MECs, over-expression of Tbx3 did not drastically promote MECs growth (Platonova et al. 2007).

**Melanoma** TBX2 and TBX3 are important in melanoma progression, facilitating senescence bypass, promoting cell growth and driving invasion. Suzanne Carreira et al. 2000 reported that TBX2 is overexpressed in melanoma cells compared to melanocytes, and that silencing of TBX2 caused G1/S cell cycle arrest and reduced melanoma cell number. Agreeing with this, activation of a dnTbx2 protein, which

competes with endogenous Tbx2 for binding to DNA but cannot repress target genes, induced senescence and impaired cell growth in *CDKN2A*-null mouse melanoma cells, as shown by β-Gal assay, BrdU incorporation, cell number changes (Vance, Suzanne Carreira, et al. 2005). In addition, dnTbx2 induces G1 cell cycle arrest and heterochromatin foci in human melanoma cells, which are both features of senescence induced by oncogenic stress. Similar to dnTbx2, silencing of Tbx2 also leads to increased senescence and p21 expression.

**Prostate cancer** Immunohistochemical (IHC) staining showed that 40 out of 53 prostate cancer patients express TBX2 in tumour tissue, while 22 also have TBX2 in adjacent healthy tissue (W.-L. Du et al. 2017). However, the IHC staining showed that TBX2 locates in the cytoplasm. As it binds DNA, TBX2 as a transcription factor should be in nucleus to be functional. Another prostate cancer study reported that TBX2 expression levels in cell lines and xenografts correlate with metastatic potential (Nandana et al. 2017). In human prostate cancer specimens, TBX2 levels increase as the disease progresses. Contrary to the previous study, TBX2 was mainly nuclear in the IHC analysis.

**Rhabdomyosarcomas** Depletion of endogenous TBX2 or expression of dnTBX2 up-regulates $p21^{CIP1}$ and $p14^{ARF}$, as well as other muscle-specific genes in rhabdomyosarcoma cells. Expression of dnTBX2 inhibits proliferation, cell migration, anchorage-independent growth and tumour formation of rhabdomyosarcoma cells (Bo Zhu et al. 2014).

**Chondrosarcoma** TBX3 promotes proliferation and anchorage-independent growth of chondrosarcoma cells, which is mediated by repressing key cell cycle regulators like $p21^{Cip1}$ and $p19^{Arf}$ (Willmer et al. 2016).

**Bladder carcinoma** In rat bladder carcinoma cell line, Tbx3 down-regulation resulted in reduced cell growth and increased apoptosis (ITO et al. 2005).

**Liver tumour** Ectopic expression of Tbx3 enhances liver tumour cell growth *in vitro* and *in vivo* (Renard et al. 2007).

**Papillary thyroid carcinoma (PTC)** Silencing of TBX3 inhibited S-phase entry and thus reduced cell growth in PTC cell lines. The growth-promoting function of TBX3 was also confirmed *in vivo*, as stable KD of TBX3 resulted in smaller tumour size (X. Li et al. 2018).

**Gastric cancer** TBX2 is overexpressed in 90 out of 161 gastric cancer patients, and the overexpression correlates with tumour invasion and metastasis, leading to poor patient survival (X. Liu et al. 2019). TBX2 promotes proliferation and invasion of gastric cancer cells *in vitro*, by promoting expression of N-cadherin, matrix metalloproteinase (MMP)2 and MMP9, cyclin E as well as phosphorylation of ERK, but inhibiting E-cadherin and p21 expression. Similarly, TBX3 is overexpressed in 46 out of 98 human gastric cancer samples as shown in IHC. TBX3 promotes gastric cancer cells growth, and drives cells invasion *in vitro* and accelerates G1/S cell cycle progression (Miao et al. 2016).

**Neuroblastoma** The concept of the core regulatory circuitry (CRC) was established in neuroblastoma. CRC consists of super enhancer-marked transcription factors that are dysregulated and drive proliferation through MYCN-dependent recurrent gains of DNA copy number. TBX2 was identified to be a constituent of a CRC in neuroblastoma. Knockdown of TBX2 in neuroblastoma cell lines reduced cell proliferation by causing G1 cell cycle arrest (Decaesteker et al. 2018).

**Lung cancer** Silencing of TBX2 in lung cancer cell lines reduced cell growth and altered cell cycle (H. Yu et al. 2020).

**Hypopharyngeal carcinoma** IHC shows that TBX3 is overexpressed in hypopharyngeal cancer compared to normal tissue (Y. Huang et al. 2020). TBX3 contributed to hypopharyngeal cancer cells growth, while TBX3 silencing altered cell cycle (S-phase arrest), decreased cell viability and increased apoptosis.

**Head and neck squamous cell cancer (HNSCC)** Sampling of 33 HNSCC patients found that TBX3 mRNA and protein levels are increased in HNSCC tissues with respect to their normal tissue counterparts (Burgucu et al. 2012).

**Pancreatic cancer** TBX3 is overexpressed in pancreatic ductal adenocarcinoma (PDAC) samples as compared to normal tissue and healthy donor patients (Perkhofer et al. 2016).

### TBX2/3 promote EMT

Many studies indicate an important role for TBX3 in migration and invasion of different cancer types. For instance, clinical data in colorectal cancer (CRC) suggests that overexpression of TBX3 is associated with the EMT phenotype, as well as a poorer overall survival (OS) and disease-free survival (DFS) (Shan et al. 2014). This observation is confirmed in many other *in vitro* studies. For example, TBX3 promotes the transition from ductal carcinoma *in situ* (DCIS) to invasive breast cancer (Krstic, Kolendowski, et al. 2019). Overexpression of TBX3 drives metastasis of HCT116 colorectal tumour cells (Zimmerli et al. 2020). TBX3 promotes hepatocellular carcinoma (HCC) metastasis through repressing E-cadherin (Dong et al. 2018; X. Feng et al. 2018). TBX3 gene is up-regulated in EMT-like squamous cell carcinoma (SCC) cell lines, while siTBX3 modestly decreased cell invasion and promoted apoptosis (Humtsoe et al. 2011). TBX3 overexpression promotes migration, invasion and sphere formation of human PDAC cell lines by increasing expression of angiogenesis-associated genes *FGF2* and *VEGF-A* (Perkhofer et al. 2016). TBX3 promotes migration of chondrosarcoma cells, rhabdomyosarcoma cells and liposarcoma cells (SW372), as suggested by scratch wound and transwell mobility assays (Willmer et al. 2016). Moreover, overexpressing TBX3 can switch radial growth phase (RGP) melanoma cells to a less proliferative, more migratory, vertical growth phase (VGP)-like status (Peres and Prince 2013).

H. Yu et al. 2020 suggested a similar role for TBX2. Silencing of TBX2 in lung cancer cell lines reduced cell migration by up-regulating E-cadherin and down-regulating N-cadherin, Slug and Vimentin. In addition, B. Wang et al. 2012 showed TBX2 induces EMT, and promotes migration and invasion in breast epithelial cells

*in vitro*. Conversely, knockdown of TBX2 in metastatic human breast cancer cell lines restores an epithelial phenotype and abrogates tumour cell invasion *in vitro*. Importantly, silencing of TBX2 reduces pulmonary metastasis of human metastatic breast carcinoma cells in *in vivo* xenograft model.

### TBX2/3 as prognostic or predictive markers for cancer

The multiple roles of TBX2 and TBX3 in cancer have led to increasing interest into the potential of these genes to act as cancer biomarkers.

For example, a recent gastric cancer (GC) study showed that 15 out of 20 patients have higher level of TBX2 protein in GC tissue compared to adjacent normal tissue, and that patients with relatively low level of TBX2 benefit more from postoperative adjuvant chemotherapy (Jun Lu et al. 2020). TBX2 is overexpressed in glioblastoma in comparison to normal brain, and high TBX2 confers a worse overall survival in TCGA dataset (Pan et al. 2020). High level of TBX2 in lung cancer patients indicates poor survival (H. Yu et al. 2020).

X. Feng et al. 2018 and Z. Li et al. 2018 reported that TBX3 is overexpressed in hepatocellular carcinoma (HCC) compared to matched normal tissue and that high expression of TBX3 in HCC patients confers a worse overall survival (OS) and recurrence-free survival (RFS). In addition, a study of 201 renal carcinoma patients reported that TBX3 is overexpressed in renal carcinoma tissues compared to adjacent tissue or normal renal tissue, and that the prognosis for patients with high TBX3 expression is worse, with a shorter median survival time and lower 3- and 5-year survival rates (Y. Wang 2018). Similarly, among human hepatoblastoma patients, high TBX3 expression also confers a worse survival probability (Renard et al. 2007).

While the above studies indicate TBX2 and TBX3 overexpression correlates with a worse prognosis, however, in glioblastoma patients, an unmethylated TBX3 promoter correlates with better response to treatment as well as higher survival (Etcheverry et al. 2010). In addition, methylation of the TBX2 and TBX3 genes

correlates with progression to muscle-invasive disease in bladder cancer (Kandimalla et al. 2012). As hypermethylation leading to tumour suppressor gene silencing is often observed in cancer, the above studies implied a positive role for TBX2/3 in tumour progression, though further investigation is needed to confirm the expression level of TBX2/3 is affected by the promoter methylation status.

### TBX2/3 and tumour drug-resistance

Consistent with their importance in cancer progression, TBX2 and TBX3 are also involved in tumour drug-resistance.

For example, in prostate cancer, the loss of chromodomain helicase DNA-binding protein 1 (CHD1) was identified to be a cause of resistance to anti-androgen Enzalutamide. TBX2 expression increased following the chromatin changes introduced by CHD1 loss, while TBX2 deletion re-sensitised shCHD1 cells to Enzalutamide, suggesting that TBX2 contributes to anti-androgen resistance in prostate cancer (Zeda Zhang et al. 2020). In addition, high TBX2 expression associates with platinum resistance and a worse overall survival probability among ovarian serous carcinoma patients, while silencing of TBX2 sensitises human ovarian serous carcinoma cells to cisplatin and carboplatin treatment (Tasaka et al. 2018). Furthermore, ectopic expression of TBX2 in adrenal carcinoma cells promotes anchorage-independence, and confers resistance to UV- and doxorubicin-induced apoptosis by inhibiting caspases activity and reducing caspase 1 and -4 expression, which is independent on p53 (Ismail and Bateman 2009). Overexpression of TBX2 also leads to resistance to apoptosis caused by ER stress-inducing thapsigargin and tunicamycin, which inhibit sarco/endoplasmic reticulum $Ca^{2+}$-ATPases (SERCA) and GlcNAc phosphotransferase (GPT) respectively.

Similarly, silencing of TBX3 sensitised U2OS cells to doxorubicin that were resistant due to overexpression of β-catenin (Renard et al. 2007).

### TBX2/3 and cancer stemness

Tumours are heterogeneous and comprise distinct subpopulations of cells. Stem cell-like cancer cells are proposed to be responsible for self-renewal, seeding and maintenance of tumours as well as giving rise to therapy-resistant cells. TBX2 and TBX3 are important in embryonic development, and they have also been implicated in cancer stemness.

For instance, in primary human PDAC cells, TBX3 co-localises with pluripotency-defining factors OCT3/4, and is highly expressed in the $CD133^+$ subpopulation that is enriched for cancer stem cells (Perkhofer et al. 2016). In breast cancer cells, estrogen stimulation increases cancer stem-like cell subpopulations, as marked by CD44+/CD24-/ESA+ surface markers, in an estrogen receptor positive ($ER\alpha^+$) cell line. TBX3 is required for breast CSC expansion in response to estrogen stimulation, and can promote sphere-forming ability of immortalised human mammary epithelia cells (HMECs) and $ER\alpha^+$ breast cancer MCF7 cells (Fillmore et al. 2010).

### TBX2/3 as a tumour suppressor

Contrary to the prevailing evidence showing TBX2 and TBX3 contribute to tumour development, there are a few studies suggesting a tumour suppressor role for TBX3.

Willmer et al. 2016 showed that TBX3 inhibits proliferation, anchorage-independent growth, *in vivo* tumour-forming ability and migration of fibrosarcoma cells. Interestingly, in the same paper they showed that TBX3 does the opposite in chondrosarcoma cells. In addition, Oh et al. 2019 showed that TBX3 overexpression inhibited proliferation and anchorage-independent growth of rhabdomyosarcoma cells, and the former is partially rescued by TBX2 expression. TBX3 overexpression also impairs cell migration as shown by wound healing assay. Collectively these results suggest a tumour suppressing function for TBX3.

In addition, mutations in T-box factors have also been found in cancer patients.

Among all T-box family members, TBX3 has the most mutations in breast cancer, including 15 frameshift mutations and 8 missense mutations. While wild-type TBX3 drastically represses the *CDKN1A* promoter activity and p21 expression, TBX3 mutants that have deficient or weak DNA-binding ability are less capable to do so, and other mutants like TBX3 H187Y are less stable resulting in shorter protein half-life. These mutations are intriguing as TBX3 is normally overexpressed in cancers and is known to repress tumour suppressors (Fischer and Pflugfelder 2015).

### 1.3.3 TBX2 and TBX3 in development

TBX2 and TBX3 are essential in development of the limbs, mammary glands, lungs and heart, and TBX3 mutations have been identified in human diseases, including the developmental disorder ulnar-mammary syndrome (UMS). In addition, TBX3 is also implicated in driving pluripotency (Papaioannou 2014; Wansleben et al. 2014). I will briefly introduce the function of TBX2 and TBX3 in development here, and expand on their target genes in and section 4.1 and section 5.1, which are most relevant to my research study.

#### TBX2/3 in embryonic development

TBX2 and TBX3 have crucial roles in the development of heart, mammary gland, limbs and lungs. Importantly, Davenport et al. 2003 reported that homozygous Tbx3 mutant mice that lack the T-domain showed embryonic lethality and defects in the yolk sac, and the mutant embryos had failed mammary bud induction and limb abnormalities. Later studies added to the knowledge of TBX2's and TBX3's expression pattern and timing as well as their roles during embryonic development.

**Lung** T. Lüdtke et al. 2013 showed Tbx2-deficient embryonic lungs have decreased proliferation and increased apoptosis, resulting in hypoplasia and reduced branching morphogenesis. As a result, mice that are homozygous for a null allele of

Tbx2 die shortly after birth. In line with this, T. H. Lüdtke et al. 2016 reported in the lung mesenchyme, both Tbx2 and Tbx3 act downstream of Shh signalling to drive the proliferative expansion via mediating Wnt signalling. In addition, Tbx2 is predominantly expressed in mesenchymal progenitors in the developing lung, and these $Tbx2^+$ progenitor cells contribute to fibroblasts, endothelial, mesothelial and smooth muscle cells (Wojahn et al. 2019). Altered contraction was observed in the Tbx2-deficient embryonic lungs, suggesting a role for Tbx2 in the functionality of bronchial smooth muscle cells.

**Heart** During heart formation, ectopically expressed Tbx3 isoforms inhibit chamber differentiation (Willem M H Hoogaars et al. 2008), and expression of Tbx3 in the embryonic heart tube represses the formation of cardiac chambers (Christoffels et al. 2004). Tbx3 is expressed in these regions where non-chamber myocardium is forming, and Tbx3 can be a marker for the developing central conduction system (W. M. Hoogaars et al. 2004). Moreover, Tbx3 regulates the gene programme of the sinoatrial node (SAN), which serves as the pacemaker of heart, by coordinately repressing an atrial gene expression programme and activating SAN-specific genes (W M H Hoogaars et al. 2007).

**Ureter** TBX2 and TBX3 are expressed in the developing ureter, and contribute to smooth muscle cells differentiation by maintaining Wnt and Bmp signalling (Aydoğdu et al. 2018).

**Mammary gland** Tbx3 is expressed in the mammary gland-forming regions in mouse embryos, and its overexpression led to extended formation of mammary gland epithelium along the dorsoventral axis (Cho et al. 2006). Adding to this, heterozygous Tbx3 mice resulted in aplasia and hypoplasia in adult mammary gland development, as a result of failed maintenance of mammary buds (Jerome-Majewska et al. 2005). On the other hand, while loss or heterozygosity of Tbx2 does not have an effect, Tbx2-Tbx3 double heterozygous mice have a more severe deficiency in mammary placode maintenance than Tbx2 single heterozygotes. Furthermore,

Tbx3 is expressed in a subset of mammary epithelial cells that mark a hormone-sensing lineage, including estrogen receptor-positive ($ER^+$) progenitor cells, and is required for the generation of hormone-sensing cells but not luminal $ER^-$ or basal cells (Kunasegaran et al. 2014).

**Liver development** Analysis of the developing mouse liver showed that Tbx3 expression was specific to the hepatoblast population. *Tbx3*-deficient embryos had defects in proliferation and differentiation of hepatic epithelial cells, causing abnormal liver development (Suzuki et al. 2008). In addition, T. H. .-W. Lüdtke et al. 2008 showed that Tbx3 is strongly expressed in the liver bud, and is required for hepatoblast migration. Tbx3-deficient mice exhibit severe liver hypoplasia at E14.5. In these embryos, the liver bud proliferation and hepatic differentiation were severely reduced.

**Epidermal development** During mouse pregnancy, Tbx3 is highly expressed in the interfollicular epidermal (IFE) basal cells and drives IFE basal cell proliferation. These $Tbx3^+$ basal cells also promote wound healing (Ichijo, Kobayashi, et al. 2017). The same research group also showed that in mouse embryos, Tbx3 is expressed in epidermal cells and regulates the proliferation and planar-cell division of basal cells during epidermal development (Ichijo, Iizuka, et al. 2017).

**Pigmentation** Localised expression of Tbx3 in hair follicles causes asymmetric deposition of pigment and distribution of hair follicle melanocytes in horse. This underlies a wild-type coat colour named 'Dun', in which non-coding mutations in Tbx3 leads to more intensely pigmented hairs in horses (Imsland et al. 2016).

**Limb** Tbx3 is required for normal limb initiation and regulates digit development in later stages (Emechebe et al. 2016), while in the prospective anterior hindlimb bud region, *Tbx3* is down-regulated (Zhen Zhang et al. 2011; Osterwalder et al. 2014). Farin et al. 2013 showed limb-specific expression of Tbx2 terminates limb bud outgrowth, as ectopic Tbx2 expression disrupts epithelial-mesenchymal signalling in the developing limb. On the other hand, Tbx2-deficient mice are likely to develop

polydactyly, the condition of having more than the normal number of toes and/or fingers, which can be ameliorated by reduced Fgf-signalling.

**Eye formation** Tbx3 is required for normal eye formation, and Tbx3 knockdown leads to retinal progenitor apoptosis and eye defects in Xenopus (Motahari et al. 2016).

### TBX3 and pluripotency

A previous study in 2010 identified Tbx3 as a driver of mouse induced pluripotent stem cells (iPSCs). On top of the Yamanaka factors Oct4, Sox2 and Klf4, the addition of Tbx3 improved the colonisation of gonads and germ-line transmission from iPSCs (Han et al. 2010). This study also showed that in embryonic stem cells (ESCs) shTbx3 induced differentiation. In line with this, a following report showed that knockdown of TBX3 in human keratinocytes and fibroblasts reduced reprogramming efficiency, though the generated iPSCs showed indistinguishable *in vitro* differentiation capacity (Klingenstein et al. 2015).

However, another report showed that Tbx3 is dispensable for establishment and maintenance of pluripotency — low passage $Tbx3^{-/-}$ MEFs did not show a difference in reprogramming capacity towards iPSCs (Russell et al. 2015). It is also intriguing that in mouse ESCs, while low Tbx3 levels led to reduced potential for *in vivo* chimeric embryos, complete loss of Tbx3 was compensated by other processes as core pluripotency factors like Klf2, Nanog and Sox2 remained unchanged.

In addition, TBX3 overexpression in human iPSCs facilitates the differentiation into cardiac pacemaker-like cells by promoting a pacemaker-specific gene expression programme (Zhao et al. 2020), and Tbx3 also facilitates the transition from pluripotent cells to mesoderm progenitors (Waghray et al. 2015).

Apart from iPSCs, TBX3 also has essential roles in ESCs. For example, Tbx3 is expressed in undifferentiated ESCs, and silencing of Tbx3 attenuates self-renewal ability and results in aberrant differentiation processes (R. Lu et al. 2010). Moreover,

Tbx3, together with Klf4, can support the self-renewal of mouse ES cells in the absence of the cytokine LIF, which is pivotal to maintain pluripotency of mouse ESCs and avoid differentiation, however, TBX3 expression is not required to maintain pluripotency in all contexts (Niwa et al. 2009). Though Tbx3 is important in maintaining pluripotency, a subset of mouse ESCs exist in an alternate pluripotent status in the absence of Tbx3, and these Tbx3-null mESCs maintain an ESC-like morphology and gene expression signature (Waghray et al. 2015).

### 1.3.4 TBX2/3 in other human diseases

**Ulnar mammary syndrome (UMS)** UMS is an autosomal dominant disorder affecting limb, tooth, hair, apocrine gland and genital development, which is caused by haploinsufficiency of TBX3. Bamshad et al. 1997 identified two TBX3 mutations from UMS patients family: ΔT227 (deletion of T227) that shifts the reading frame and produces a premature termination codon, and Int2:G1→C (G→C transversion in the first nucleotide of intron2) that alters the consensus splice donor site and is predicted to alter normal gene splicing. A subsequent study found the C-terminal UMS-related mutations locate in the repression domain that is required for TBX3 transcriptional repression activity (Carlson, Ota, C. E. Campbell, et al. 2001). To date, 18 TBX3 mutations have been identified to cause UMS (Khan et al. 2019).

**Obesity** Quarta et al. 2019 first showed that loss of Tbx3 function in hypothalamic neurons, which control energy balance through releasing specific neuropeptides, causes weight gain and other metabolic disturbances. They also showed the *Drosophila melanogaster* Tbx3 homologue, *omb*, is produced in the central nerve system and prevents obesity. Later, another report showed that TBX3 is up-regulated in the serum of diabetic patients compared to normal participants (Z. Yu et al. 2020). High glucose treatment increased TBX3 protein levels in human umbilical vein endothelial cells (HUVECs), whereas silencing of TBX3 partially

alleviated the high glucose-induced proliferation of HUVECs. Moreover, TBX3 also contributes to glucose induced cell migration and angiogenesis.

**Rheumatoid arthritis** The genetic studies of collagen-induced arthritis (CIA), a mouse model to study rheumatoid arthritis, linked the genetic fragment containing *Tbx3* and *Tbx5* to severe arthritis (Sardar et al. 2019). Further study showed intracellular Tbx3 level strongly decreased in active B lymphocytes, and that serum Tbx3 level correlates with severity of CIA.

**Heart disease** R.-r. Zhang et al. 2019 sequenced blood samples of congenital heart disease (CHD) patients and healthy participants, and found that a regulatory variant c.-1028G>C in the *TBX2* promoter is associated with significantly lower (30%) CHD susceptibility. The activity of c.-1028G>C mutant containing promoter is reduced as shown in luciferase assays, which might lead to reduced TBX2 expression. On the other hand, TBX2 R608W, T249I and R616Q mutants, and TBX3 A192T, M65L and A562V mutants were identified in patients with conotruncal heart defects. These mutants have altered transcriptional activation ability as shown in luciferase assays (Xie et al. 2018).

**Alzheimer's disease** Reinhardt et al. 2019 reported that TBX2, but not TBX3 or TBX21, is up-regulated in cortical brain tissue of Alzheimer's disease patients.

### 1.3.5 TBX2 and TBX3 others

#### RNA-binding capacity of TBX3

Although TBX3 is known as a transcription repressor, there was, however, an interesting studying showing that TBX3 also regulates target gene expression via directly binding RNA and subsequently regulating pre-mRNA splicing (P. Kumar et al. 2014). TBX3 interacts with RNA-binding proteins and binds to the T-element in RNA. For interactions with RNA-binding protein, the T-box DNA-binding domain (DBD) of TBX3 is required, and point mutations in DBD (L143P and

N277D), which are found in UMS patients, abolish the interaction and subsequent splicing regulatory function.

## 1.4 Objectives

To summarise, TBX2 and TBX3, members of the T-box transcription factors family, play crucial roles in embryonic development and are overexpressed in several types of cancer. Their overexpression contributes to various stages of tumourigenesis, from facilitating senescence bypass, promoting invasion and metastasis, to conferring resistance to therapies. Though it is well recognised that TBX2 and TBX3 are oncogenes implicated in cancer development, their downstream effectors, upstream regulators remain obscure. It is therefore imperative that we expand our knowledge of their integration into the broader tumourigenesis signalling network.

Hence, the objective of my project is to explore the involvement and mechanism of action of TBX2 and TBX3 during cancer development using cutaneous melanoma as model. The following three main subjects regarding melanoma progression will be investigated:

1. What are the direct target genes of TBX2 and TBX3?
2. Which cofactors are involved and how they cooperate with TBX2 and TBX3?
3. What signalling pathways and proteins regulate the expression of TBX2 and TBX3?

# 2

# Materials and methods

## 2.1 Mammalian cell methods

### 2.1.1 Routine cell culture

Mammalian cells were grown in RPMI media supplemented with 10% fetal bovine serum (FBS) without antibiotics. When reaching confluence, cells were splitted 1:6 to 1:10 depending on cell line using trypsin.

### 2.1.2 Immunofluorescence (IF)

1. Adherent cells were grown on glass cover slips in 12-well plates until 80% confluence
2. Growth media was removed and cells were washed once with 1x PBS
3. PBS was removed and 4% paraformaldehyde (PFA) was added, and cells were

fixed at room temperature for 10 min

4. PFA was removed and cells were washed twice with PBS
5. Cells were permeabilised and blocked in blocking solution (5% BSA in PBS + 0.1% Triton X-100) at room temperature for 20 min
6. Primary antibodies were diluted in blocking solution and added to wells, incubated at room temperature for 1 h
7. Primary antibodies were removed, cells were washed twice with PBS
8. Secondary antibodies (conjugated to fluorescent labels) were diluted 1:1,000 in blocking solution; DAPI (0.5 µg/ml) was added; the mix was incubated at room temperature for 30 min
9. Antibody DAPI mix was removed, cells were washed three times with PBS
10. Excess PBS was removed; cover slip was mounted on polysine microscopy slides with 2 µl of VECTASHIELD mounting medium (Vectorlabs, H-1000-10), and cover slip edges were sealed using nail polish
11. Samples were imaged using Zeiss 710 microscope at 20x objective

## 2.2 Nucleic acid methods

### 2.2.1 Clone sgRNA into Cas9 plasmid

To direct Cas9 to cut the C-terminus of Tbx2, single-guide RNA (sgRNA) was designed to target the following sequence near the stop codon: GCCGGGAGTCG CCCAAGTGAGGG. To clone sgRNA into px330 plasmid, top strand (5'→3') CA CCGGCCGGGAGTCGCCCAAGTGA and bottom strand (5'→3') AAACtcacttgg gcgactcccggcC were annealed and then cloned into BbsI site of px330.

#### Oligo annealing

1. oligos were reconstituted to a final concentration of 100 µM

2. 1µl of top strand oligo and 1µl of bottom strand oligo were pooled to 48 µl of annealing buffer (100 mM NaCl, 50mM HEPES, pH 7.4 in ultrapure water)
3. oligos were annealed using the following program in a PCR machine:
   (a) 90 ℃ for 4 min
   (b) 70 ℃ for 10 min
   (c) decrease temperature by 1 ℃ every minute for 50 min
   (d) hold at 21 ℃

**Vector digestion**

1. px330 was digested for 2 h at 37 ℃ with BbsI (NEB) in the following buffer:
   - 5 µl of Buffer 2.1 (NEB)
   - 5 µg of px330 vector
   - 3 µl of BbsI (NEB)
   - water *add to* 50 µl
2. Digested px330 was isolated on a 0.7% agarose gel and extracted using QIAquick Gel Extraction Kit (QIAGEN, 28706)

**Ligation**

Annealed oligos and digested px330 vector were ligated at room temperature for 10 min (including ligation control) in the following mix:

- 1 µl of annealed oligos
- 1 µl of T4 DNA ligase buffer
- 1 µl of T4 DNA ligase
- 120 ng of digested px330 vector
- water *add to* 10 µl

**Transformation**

1. DH5α *E. Coli* cells were thawed on ice
2. 3 µl of ligation mix was added to 30 µl of DH5α, then incubated on ice for 10 min
3. The mix was heatshocked in an Eppendorf tube at 42 ℃ for 45s
4. Then cooled on ice for 2 min
5. 200 µl of SOC media was added to the mix and incubated at 37 ℃ for 45 min
6. Bacterial suspension was plated on pre-warmed LB agar plates, incubated overnight at 37 ℃

* a negative control (digested px330 vector without inserts) was included

**Colony PCR**

| **Reagents** | **Amount per reaction** |
|---|---|
| 2x MyTaq red mix | 5 µl |
| water | 4 µl |
| Forward Primer (SL056) 10 µM | 0.3 µl |
| Reverse Primer (sgRNA bottom strand) | 0.3 µl |
| DH5α | scratch with a tip |

1. Thermal Cycler Programme:

| **Steps** | | **Temperature** (℃) | **Time** |
|---|---|---|---|
| Initial Denaturation | | 95 ℃ | 2 min |
| 30x | Denaturation | 95 ℃ | 15 sec |
| | Annealing | primer pair specific | 15 sec |
| | Extension | 72 ℃ | 15 sec |
| Final Extension | | 72 ℃ | 2 min |
| Hold | | 4 ℃ | hold |

2. The PCR mix was run on a 1.5% agarose gel to check.

#### Sequencing

Selected clones were mini-prepped and sent for sequencing using primer SL056

### 2.2.2 Linear short homology arms synthesis

#### Design of donor sequence C-terminal fusion

To design the primers for PCR amplification of donor template with homology arm, the 40 bp immediately preceding the stop codon and the 40 bp immediately after were selected. Then 20 bp from the N-terminus of your chosen tag (e.g. 3xHA or any fluorescent protein) was added to the forward primer, and 20 bp from the C-terminus of the selection marker - an antibiotic resistant gene - was added to the reverse primer. To eliminate the sgRNA site, the PAM site was removed from the donor sequence and a silent mutation was introduced to the last base before the PAM.

#### PCR amplification of donor sequence

The primers designed above and respective template plasmids were used to make donor sequence for homology directed repair (HDR). Phusion polymerase (NEB, M0530S) was used with HF buffer.

1. Thermal Cycler Programme:

| **Steps** | | **Temperature** (℃) | **Time** |
|---|---|---|---|
| Initial Denaturation | | 95 ℃ | 2 min |
| 32x | Denaturation | 95 ℃ | 15 sec |
| | Annealing | primer pair specific | 15 sec |
| | Extension | 72 ℃ | 60 sec |
| Final Extension | | 72 ℃ | 2 min |
| Hold | | 4 ℃ | hold |

2. The PCR mix was purified using QIAgen PCR purification kit (QIAGEN, 28106).
3. Purified PCR products were sequenced before being used for tagging.

For tagging C-terminus of Tbx2 with 3xHA, forward primer SL095 and reverse primer SL069 were used together with Cas9 plasmid px330-SL007.

### 2.2.3 qRT-PCR

Primer stock: 100 mM in nuclease-free water

To check primers and optimise cDNA dilution: cDNA serial dilution 1/10, 1/20, 1/40, 1/80 and 1/160 were used in each experiment for each primer pair. Indicators of good primers: Ct values increase by 1 between dilutions, and there is only one peak in the melting curve. After setting threshold in the exponential region, a dilution that gives a Ct between 20 and 30 was chosen. Reagents used: refer to 2.7.4 on page 68. Unit: reagent amount used is for one 6-well plate well.

**Lysis**

1. Cells were washed twice with 1X PBS
2. 300 µl of lysis buffer was added, and plates were rotated to ensure even coverage (all on ice)
3. Cells were collected using cell scraper and transferred to labelled Eppendorf tubes (on ice)

**RNA isolation (on ice)**

1. Cell lysates were put on ice
2. Cells were homogenised using syringe and needle - up and down for 10 times
3. 300 µl of 70% ethanol was added, the Eppendorf tubes were flicked to mix for 5 times

4. The mixture was added to the column, centrifuged for 30 sec at 10,000g; the effluent was discarded
5. 700 µl of buffer RW1 was added to the column ,centrifuged for 30 sec at 10,000g, then effluent was discarded
6. 500 µl of RPE was added to the column, centrifuged for 30 sec at 10,000g, then effluent was discarded
7. 500 µl of RPE was was added to the column, centrifuged for 1 min at 10,000g, then effluent was discarded
8. Columns were transferred to fresh tubes, and centrifuged for 1 min at 11,000g to remove residual buffer
9. Columns were transferred to fresh Eppendorf tubes; 35 µl of water was added to each column, after 1 min, centrifuged for 1 min at 11,000g
10. The concentration of RNA in the Eppendorf tube was measured using a NanoDrop 2000

**genomic DNA elimination**

1. RNA and reagents from QuantiTect Reverse Transcription Kit (QIAGEN, 205313) were put on ice
2. RNA, buffer and water were added to 8-strip PCR tubes

| **Reagent** | **Per reaction** |
|---|---|
| 7x gDNA wipeout buffer | 2 µl |
| Sample RNA | 1 µg |
| RNAse-free water | added to a final volume of 14 µl |

3. The mix was incubated at 42 ℃ for 2 min on Thermal cycler
4. Then the mix was transferred back on ice

*Note: unused RNA can be stored at -80 ℃*

**cDNA synthesis**

| **RT mix Reagent** | **Per reaction** |
|---|---|
| Reverse transcriptase | 1 µl |
| RT buffer 5X | 4 µl |
| RT primer mix | 1 µl |

1. Reverse transcription (RT) mix was prepared for all reactions as shown above
2. 6 µl of RT reagent was added to the 14 µl sample mix
3. The mix was incubated in a Thermal Cycler: 42 ℃ for 20 min, then 95 ℃ for 3 min, then held at 4 ℃

*Note: a negative control without RT was added; cDNA can be stored at -20 ℃*

**qPCR**

| **Primer F/R mix** | **Per reaction** |
|---|---|
| Water | 627 µl |
| Forward primer stock | 20 µl |
| Reverse primer stock | 20 µl |

| **Master mix Component** | **Per reaction** |
|---|---|
| Brilliant II SYBR 2x | 6.5 µl |
| Primer F/R mix | 1.3 µl |
| water | 0.2 µl |

*Note: primer stock solution is 100 µM, in Master Mix is 0.5 µM, in reaction is 0.3 µM*

**For Rotorgene Thermal Cycler sample preparation**

1. Rotor disk 100 heat sealer was turned on
2. 8 µl of Master mix and 5 µl of cDNA was pipetted into each well

3. Rotor disk 100 was inserted to the metal support, and plastic cover was placed over the disk; then inserted into heat sealer and the lid was pressed down for sealing
4. The ring was taken out and transferred to the gold mount and placed in Rotorgene
5. Then the PCR reaction was run as described below:

| **Steps** | | **Temperature** (℃) | **Time** |
|---|---|---|---|
| Taq activation | | 95 ℃ | 10 min |
| 50x | Denaturation | 95 ℃ | 10 sec |
| | Annealing | primer pair specific | 15 sec |
| | Extension | 72 ℃ | 20 sec |
| Melting | | 72 to 95 ℃ | |

## 2.3 Protein methods

### 2.3.1 Western blotting

The recipes for Western Blotting (WB) buffers can be found at 2.7.3 on page 66. The formulation below was used to prepare two SDS gels freshly.

| **Component** | **Stacking Gel** | **Resolving Gel** |
|---|---|---|
| Water | 4.5 ml | 7.9 ml |
| Respective buffer | 2 ml | 5 ml |
| 30% acrylamide | 1 ml | 6.7 ml |
| Ammonium Persulfate (APS) | 30 µl | 200 µl |
| Bromophenol blue | 2 µl | none |
| Tetramethylethylenediamine (TEMED) | 30 µl | 50 µl |

**Lyse cells and collect samples**

1. Cells were washed once with 1x PBS

2. Cells were lysed with RIPA buffer (around 200 µl for a confluent 6-well)
3. Plates were rotated at 4 ℃ for 10 min
4. Cell lysates were collected and transferred to Eppendorf tubes

**(optional) BCA assay**

1. Protein concentration was measured using a BCA assay kit (Pierce, 23225)
2. Protein concentration was adjusted by adding RIPA buffer to more concentrated samples to make protein concentration even across different samples

**Sample denaturation**

1. 2x Laemmli buffer was added to the cell lysates
2. The mixture was boiled in a heat block at 95 ℃ for 15 min

**Load sample and run SDS-PAGE**

1. Gels were poured into gel casting cassettes, and well combs were inserted in stacking gel:

| Cassette | Resolving Gel | Stacking Gel |
|---|---|---|
| 1.5 mm | 8 ml | 1.5 ml |
| 1.0 mm | 5.5 ml | 1 ml |

2. Gels were put into blot module and submerged in running buffer
3. Samples and molecular weight (M.W.) ladder were loaded into gels
4. Gels were run at 100 V for 2 h until the front dye reached the edge of gel

**Transfer protein to membrane**

1. Sponges and filter papers were submerged in transfer buffer, and membrane in water
2. The transfer module was prepared in the following layers: sponge, filter paper,gel (mirrored orientation), membrane, filter paper and sponge

3. The transfer module was moved to the tank together with an ice pack
4. The tank was filled with transfer buffer, and run at 400 mA for 1 h
5. Membranes were dyed with Ponceau solution to check transfer efficiency
6. Membranes were washed in 1x TBS/T, then blocked using 5% milk for 20 min

**Antibodies Incubation**

1. Primary antibodies were diluted in blocking buffer
2. Membranes were cut into strips according to M.W.
3. Membrane strips were incubated with primary antibodies at 4 ℃ overnight
4. Membranes were washed with 1x TBX/T 3 times
5. Secondary antibodies diluted in blocking buffer were added, and incubated at room temperature for 2 h
6. Membranes were washed with 1x TBX/T 3 times

**Detection**

1. ECL Western blotting detection reagents (GE, RPN2106) were mixed 1:1
2. Excess of TBS/T from membranes were removed
3. ECL mix was added to membranes and incubated at room temperature for 1 min
4. Excess of ECL was removed
5. Membranes were placed into a cassette for developing films in dark room

## 2.4 Cell biology methods

### 2.4.1 eFlut tagging endogenous Tbx2

The details of the eFlut tagging can be found in (Stewart-Ornstein and Lahav 2016). Method for cloning the sequence encoding the sgRNA into a Cas9 plasmid

can be found in section 2.2.1 on page 41. For making the donor template, please refer to section 2.2.2 on page 44.

### Plate and transfect cells

1. 30,000 B16 cells were plated in 2 ml of growth media per well in 6-well plate.
2. Trans IT-LT1 Reagent was warmed to room temperature.
3. 200 µl of Opti-MEM serum-free media was placed in a sterile tube.
4. 1 µg of Cas9-sgRNA plasmid and 1 µg of homology donor were added to the media and mixed.
5. 10 µl of LT1 reagent was added to the diluted DNA mixture (1:5 ratio of DNA:lipid mix) and gently mixed.
6. The mix was incubated at room temperature for 15 min.
7. The DNA lipid mix was added dropwise to cells

### Recovery

48 h after transfection, cells were trypsinised and plated in a 10-cm dish. Cells were allowed overnight to recover from transfection and to express Cas9 plasmid.

### Selection

B16 cells transfected with 3xHA-NeoR donor sequence were selected in growth media containing 500 µg/ml of G418 for 6 days until negative control cells, which are transfected only with the Cas9 plasmid but not the donor sequence, are eliminated.

Viable cells were recovered in normal growth media for 3 days until individual clones are visible by eyes and it is easy to manually pick clones.

### Verification by genomic PCR

Primers flanking last exon of Tbx2 and 3' untranslated region (UTR) were used to generate a band that is distinguishable. The PCR reaction should give a large

band if the insertion was successful, a small band if not, and both bands if it is heterozygous clone or a mixed population.

For B16-Tbx2-3xHA-NeoR clones, primers SL127/128 were used in the genomic DNA PCR. Then nested PCRs were conducted - primers SL059/128 were used to amplify 3xHA-NeoR-3' UTR region, and primers SL127/132 to amplify the Tbx2-3xHA-NeoR region - followed by sequencing to confirm the locus was modified correctly.

### 2.4.2 Nuclear tethering assay

#### Transfection of bait and prey plasmids

1. 24,000 U2OS-LacO cells were plated in 24-well plates
2. In total 250 ng of the bait LacR-mCherry plasmid and the GFP-tagged prey plasmids were diluted in 80 µl of Opti-MEM
3. 2 µl of Fugene 6 was added to 80 µl of Opti-MEM and mixed well
4. Plasmids and Fugene 6 solutions were mixed and incubated at room temperature for 15 min
5. The transfection mix was added to cells
6. 24 h later, cells were split 1:6 and transferred to 8-well chambered coverglass (Nunc™ Lab-Tek™, 155411)
7. Cells were allowed to attach overnight

#### Imaging

1. Media was removed from 8-well chamber; cells were washed once with PBS
2. Cells were fixed with 4% PFA at room temperature for 10 min
3. PFA was removed and cells were washed twice with PBS
4. DAPI (1 µg/ml) diluted in PBS was added to wells, and incubated at room temperature for 5 min
5. DAPI solution was removed, cells were washed with PBS three times

6. Chamber wells were mounted with 10 µl of VECTASHIELD mounting medium (Vectorlabs, H-1000-10)
7. 8-well chamber was imaged at Zeiss 710 at 64x oil objective using 1280x740 solution; cells transfected with both plasmids were selected

**Analysis in ImageJ**

1. DAPI staining was used to mask the nuclear area; nuclear area size (ns) was measured
2. mCherry staining was used to mask the nuclear focus (dot) area; the dot area size (ds) was measured
3. Average GFP intensity in nuclear and dot area was measured ($GFP_n$ and $GFP_d$); total GFP total fluorescence in nuclear and dot area was measured ($GFP_{tn}$ and $GFP_{td}$)
4. GFP dot/nucleus ratio was calculated as

$$ratio = \frac{GFP_d}{(GFP_{tn} - GFP_{td}) \cdot (ns - ds)}$$

### 2.4.3 Crystal violet staining

1. 20,000 B16 cells were plated in a 12-well plate (0.88 ml of cell suspension) and transfected with 10 nM of siRNA (diluted in 120 µl of Opti-MEM containing 3 µl of Lipofectamine RNAiMAX)
2. After 24 h, the 1 ml of transfection mix was replaced with 0.88 ml of normal growth media
3. After 24 h, cells were again transfected with 10 nM of siRNA
4. After 48 h, the transfection mix was washed off and cells were fixed with 2% of PFA for 10 min
5. Then cells were stained with 0.1% crystal violet for 10 min

6. Then plates were air-dried and scanned using a Fuji FLA-5100 imager

### 2.4.4 Cell cycle analysis by FACS

The following protocol is for cells growing in 6-wel plates

1. Cells were washed and trypsinised
2. Cell suspension was transferred to Eppendorf tubes on ice and spun down at 800 g for 2 min at 4 ℃
3. The supernatant was discarded; cells were resuspended in 300 µl of PBS
4. 750 µl of pure ethanol was added and mixed immediately by pipetting
5. Cells were incubated with ethanol on ice for 1 h
6. Cells were spun down at 800 g for 2 min at 4 ℃
7. The supernatant was discarded; cells were resuspended in 250 µl of PI staining solution
8. Cells were incubated with PI staining solution for 40 min at 37 ℃
9. 750 µl of PBS was added and mixed
10. Cells were spun down at 800 g for 2 min at room temperature
11. The supernatant was discarded
12. Cells were resuspended in 300 µl of SS5 solution, then analysed in a FACS machine

## 2.5 High-throughput sequencing and screen

### 2.5.1 ChIP-seq sample preparation

The original ChIP method was adapted from a Nature Protocol paper (Nelson et al. 2006) by Dr Pakavarin Louphrasitthiphol in our group for transcription factors with an HA-tag (Louphrasitthiphol et al. 2020). The ChIP-seq sample

preparation method used for Tbx2 and TBX3 was described below. Buffer recipes can be found in section 2.7.3 on page 65.

Sample unit: 3x 150mm dishes of 80% confluent melanoma cells. Each dish has 20 ml of media at the time of cross-linking.

**Cross-linking (per dish)**

1. 0.54 ml of 30% formaldehyde (Sigma F8775) was added to each dish (formaldehyde final concentration 0.8%); cells were cross-linked at room temperature for 10 min
2. 1.25 ml of 2 M glycine was added to each dish (final concentration 125 mM); the cross-linking was quenched at room temperature for 10 min
3. The medium was discarded and each dish washed once with 10 ml of cold PBS
4. Adherent cells were scraped off dishes and transferred to 50 ml falcon tube (4 + 8 ml PBS); 3 dishes/tube
5. The tubes were centrifuged at 1,500 g for 10 min
6. The supernatant was removed and the pellet kept at -80 ℃

**Cell Lysis**

1. A PIC tablet was added to the ChIP lysis buffer
2. 1 ml of ChIP lysis buffer was added to the cell pellet and pipetted up and down using a pipette tip
3. Cell lysate was passed through 100 µm cell strainer, then transferred to a Covaris tube

**Shearing chromatin**

1. Sonication was carried out using a Covaris S220: 145 W peak incident power, 5% duty cycles, 200 cycles per burst

2. Sonication time: 15 min for 501mel-iTBX3-3xHA-PuroR cells, 18 min for B16-Tbx2-3xHA-NeoR cells
3. After sonication, cell lysate was centrifuged at 13,000g for 10 min at 4 ℃
4. 10 µl of supernatant was taken as 1% ChIP for sequencing and to check fragment size
5. 5 ml of supernatant was pooled in one 50ml falcon tube
6. 10 µl was taken as 20% IP input

**Incubation with antibody and blocking Dynabeads**

- 9 ml of ChIP dilution buffer was added to each falcon tube; then 12 µg of HA antibody (Roche 11666606101) was added then incubated at 4 ℃ overnight
- 60 µl of Dynabeads protein G (Invitrogen 10004D) was washed once then resuspended in 78 µl of ChIP dilution buffer; then 2 µl of BSA was added (final concentration 0.5 mg/ml); Dynabeads were blocked with BSA at 4 ℃ overnight

**Immunoprecipitation**

80 µl of blocked Dynabeads protein G was added to each falcon tube and rotated at 4 ℃ for 2 h

**Washing**

1. Samples were pooled together in one 1.5 ml Eppendorf tube placed in magnetic stand
2. 20 µl of flow through was taken as 20% unbound
3. Beads were washed with 1 ml of low salt buffer, then transferred to a new tube to avoid contamination; low salt washing was repeated once
4. Beads were washed with 1 ml of high salt buffer, then then transferred to a new tube to avoid contamination; high salt washing was repeated once

5. Beads were washed with 1 ml of LiCl buffer, then then transferred to a new tube to avoid contamination; LiCl washing was repeated once

**Elution**

1. 10 ml of fresh elution buffer was made, and microwaved for 30 sec to boil
2. 200 µl of elution buffer was added to each Eppendorf tube, including ChIP input

**Reversal of cross-linking**

1. 35.8 µl of 2M NaCl, 20 µg of RNAse A and 20 µg of proteinase K were added to the beads in elution buffer (NaCl final concentration 0.3M)
2. The mix was incubated at 55 ℃ overnight

**Recovery of ChIPed DNA**

1. QIAquick PCR purification kit was used to recover the ChIPed DNA
2. Bound DNA was elutted with 30 µl of ultra pure water; this fraction was used for ChIP-seq library preparation
3. The column was eluted again with 140 µl of water; this fraction was used for qPCR analysis

**Final preparation**

- The ChIPed DNA concentration was measured by dsDNA BR assay kit (Qubit, Q32850)
- Sonicated DNA fragment size was checked using Tapestation (this was done in WTCHG)
- For WB samples, RIPA buffer was added to the IP input, unbound and IP sample to make the final volume 40 µl; then 40 µl of 2x Laemmli buffer was added to make the WB sample

- Two biological replicates from 501mel-iTBX3-3xHA-PuroR polyclone (both formaldehyde and PFA crosslinked) as well as B16-Tbx2-3xHA-NeoR cl#02 and cl#09 were submitted for sequencing in the WTCHG
- The sequencing details are: 150 bp paired-end in a NovaSeq6000 (Illumina Truseq)

### 2.5.2 RNA-seq sample preparation

1. 1.1x$10^5$ B16, B16-Tbx2-3xHA-NeoR cl#02 and cl#09 were plated in separate 6-well plates (x$10^4$ cells/ml, 2.2 ml/well)
2. siCN and siTbx2 were diluted in 150 µl of Opti-MEM (to make a final concentration 20 nM).
3. 7.5 µl of Lipofectamine RNAiMAX (Invitrogen, 13778030) was diluted in 150 µl of Opti-MEM (to make a final concentration 0.3%).
4. Diluted siRNA and diluted Lipofectamine were mixed together and incubated at room temperature for 5min.
5. 300 µl of siRNA-Lipid mix was added dropwise to cells and mixed
6. At 24 h, transfection media was replaced with normal growth media; cells were allowed 24 h to recover.
7. At 48 h, cells were lysed to collect RNA as described in section 2.2.3 on page 45.
8. 15 µl of 200 ng/µl extracted RNA from each sample (biological triplicate for all experiments) was submitted to the Wellcome Trust genomic service, Oxford.
9. ERCC ExFold RNA Spike-In Mixes (Ambion) was added prior to Library prep using QuantSeq 3' mRNA-Seq Library Prep Kit, using 500 ng starting material to minimize the PCR amplfication step.
10. Samples prepared as biological triplicates were sequenced on HiSeq4000

(Illumina).

### 2.5.3 FDA-approved drug library screen

**Screening process**

1. *Cell seeding* 501mel cells expressing inducible FLAG-tagged TBX2 (501mel-iTBX2-FLAG) were pre-stained with CellTracker Orange, and then mixed with 501mel-iTBX3-FLAG. These two cell lines were co-seeded at a 1:1 ratio (in total 800 cells per well in 80 µl) into 30 assay plates (384-well flat bottom, black with clear bottom, polystyrene; Greiner 6007558), in complete RPMI medium (10% FBS, 1% penicillin and streptomycin) and incubated for 24 h at 37 ℃, 5% CO2.
2. *Doxycycline induction* medium was replaced with 75 µl of RPMI containing 30 ng/ml doxycycline.
3. *Preparation of Pharmakon 1600 bioactive compounds library* Pharmakon 384-well 10mM compressed library plates were transferred to 160 µM intermediate working stock plates (40µL/well, 160 µM in 1.6% DMSO).
4. *Drug treatment* 24 h after adding doxycycline, 5 µl of intermediate stock was added into each well (final concentration 10 µM in 0.1% DMSO).
5. *Time points* After 4 h, 12 h and 24 h treatment, 10 plates were fixed respectively (5 library plates per replicate). Cells were fixed with 30 µl/well of 4% PFA in PBS for 15 min at room temperature. Plates were kept at 4 ℃ until processing.
6. *Staining with antibody* Cells were permeabilised with 30 µl/well of 0.2% Triton-X100 in 5% BSA/PBS for 20 min at room temperature, probed with 20 µl/well of primary FLAG antibody (Sigma F7425, 1:500 diluted in 5% BSA/PBS) for 1 h at room temperature, then incubated with secondary Alexa488 anti-rabbit IgG (Life Technologies A21206, 1:1000 diluted in 5% BSA/PBS) and DAPI

(Life Technology D1306, 1:1000 diluted in 5% BSA/PBS) for 1 h at room temperature.

7. Plates were sealed with non-transparent film and kept at 4 °C before imaging.

*Note: 501mel-iTBX2-FLAG and 501mel-iTBX3-FLAG cell lines were made by a previous DPhil student Johannes Schmidt*

**Imaging**

- GE IN Cell Analyzer 6000 – line scanning confocal microscope
- on-line cell counting to 1000 cells (up to 16 fields of view)
- UV/DAPI: 75ms, confocal 0.99 airy unit; Blue/FITC: 2500ms, confocal 1.23 airy unit; Green/dsRed: 100ms, open aperture

**Data analysis**

- IN Cell Developer – analysis software
- Analysis process:

    1. *Segmentation* The DAPI channel (DNA) was used to mask the nuclear area, the FITC channel (FLAG) was used to mask the whole cell area; FITC fluorescence intensity in nuclei was determined as 'nuclear protein level', and that in the collar area was determined as 'cytoplasmic protein level' (collar = dilated nuclei – eroded nuclei); iTBX2 cells were selected using mean intensity of CellTracker Orange stain with nuclear region (background = 110GL, threshold set at > 200GL)
    2. *Measurements* Readouts for each cell line included cell number, length/width ratio of nuclei, mean intensity of nuclei, nuclear area size, form factor (roundness of nuclei) and standard deviation of DAPI intensity (nucleus uniformity), which indicate cell health; mean FITC intensity in the whole cell, nucleus and cytoplasm, which indicate TBX2 and TBX3 protein

level and subcellular localisation.

- Provisional data were collected in Excel; R2, Pearson, SSMD and Z-score were calculated.

## 2.6 Bioinformatics

### 2.6.1 RNA-seq analysis

#### siTbx2 RNA-seq analysis

*Alignment and Read counts* The output raw fastq files were examined for quality using Fastqc (version 0.11.7) and mapped against mm10 genome using STAR (version 2.6.1d). Reads per gene from STAR output BAM file were counted using FeatureCounts (subread/1.6.2). Counts per gene and sample information were converted to DeSeqDataSet objects and sequentially used as input for differential gene expression analysing using the R package Deseq2 (version 1.28.1).

*Differential gene expression* The Deseq function filters out lowly expressed gene, calculates normalisation size factors, estimates dispersion, applies negative binominal GLM fitting and calculates Wald statistics to identify differentially expressed genes (DEGs). Genes with an adjusted p-value less than 0.05 were considered as statistically significant DEGs.

*Data visualisation and QC* Heatmap of RNA-seq samples were generated from counted reads in DeSeqDataSet object using the R package ComplexHeatmap. Raw reads were centered and scaled around the mean and hierarchically clustered. PCA plot was generated from log transformed read counts using the R package ggfortify, to visualise inter- and intra-group variability arising from transfection, clone and replicates.

*Further Analysis* Gene set enrichment analysis (GSEA) was carried out using the R package fgsea. 1,000 permutations were conducted for each gene set. Data were

pre-ranked by a gene significance score $\pi$-value combining expression fold change and statistical significance. $\pi$-value of the $i$-th gene in a dataset was calculated as below: $\pi_i = \phi_i \cdot (-\log_{10} \rho_i)$, where $\varphi$ is the fold change and $\varrho$ is the adjusted p-value (Xiao et al. 2014).

Gene ontology (GO) was carried out on http://geneontology.org/ using statistically significant DEGs, and the top 15 pathways visualised using the R package ggplot2. KEGG pathway analysis was carried out in the R package clusterProfiler using the enrichKEGG function, and the visualisation of individual KEGG pathways was generated using the pathview package. Visualisation of the GSEA, GO and KEGG pathways was done in ggplot2.

### 2.6.2 ChIP-seq analysis

#### Tbx2 ChIP-seq analysis

*Alignment and Peak Calling* Prior to mapping to the mouse reference genome (mm10) with bowtie2 (v.2.3.5), quality of the raw sequencing data was evaluated using FastQC (v.0.11.7) and adapter contamination were removed using cutadapt (v.2.8) when necessary. Peak calling was performed using MACS2 (v.2.1.2) taking a q-value of 0.01 as threshold.

*Visualisation* Bowtie2 generated SAM files were compressed to BAM files, then indexed using SAMtools (v.1.9), and sequentially converted to bigWig files using uscstools (v.373) and uploaded to UCSC Genome Browser for visualisation. Input and ChIP bigWigs were overlayed in the same track for a more intuitive presentation for peak quality.

*Motif Detection* HOMER (v.4.8) was used to perform known motif enrichment and *de novo* motif identification, taking given peak size as input as we have paired-end data.

*Quality Control* Tag density plots were generated using seqMINER (v.1.3.4),

overlap of peaks were generated using bedtools intersect function. Peak distribution across genomic features and relative to TSS was obtained from R package ChIPseeker. Venn diagram of overlapped peaks were generated from R package ChIPpeakAnno.

#### TBX3 ChIP-seq analysis

*Alignment and Peak Calling* Raw sequencing data was mapped to the human reference genome (hg38) with bowtie2. Mapped reads were for peak calling using MACS2 taking a q-value of 0.01 as threshold.

*Visualisation* Bowtie2 generated SAM files were compressed to BAM files, then indexed using SAMtools, and sequentially converted to bigWig files using uscstools and uploaded to UCSC Genome Browser for visualisation. Input and ChIP bigWigs were overlayed in the same track for a more intuitive presentation for peak quality.

*Motif Detection* HOMER (v.4.8) was used to perform known motif enrichment and *de novo* motif identification, taking given peak size as input as I have pair-end data.

*Quality Control* Overlap of peaks were generated using bedtools intersect function. Peak distribution across genomic features and relative to TSS was obtained from R package ChIPseeker. Venn diagram of overlapped peaks were generated from R package ChIPpeakAnno.

## 2.7 Materials

### 2.7.1 Oligonucleotide list

| **Oligo ID** | **Sequence (5'→3')** |
|---|---|
| SL056 | AGGGCCTATTTCCCATGATTC |
| SL059 | ATGTTCCTGACTATGCGGGC |
| SL069 | agcccgggtggcctgcgtggcaggggagcagctgggcagccccATGGGTGGAGGCGGTTC |
| SL095 | agccagcgagccctctccccaggccgggagtcgcccaaATACCCATACGATGTTCCTGAC |
| SL127 | TGCGCTTCAGTCCTTACCAG |
| SL128 | CAGCATCCTCCACACCAGTT |
| SL132 | GGTAGCCAACGCTATGTCCT |

### 2.7.2 siRNA list

| **siRNA ID** | **Sense (5'→3')** | **Antiense (5'→3')** |
|---|---|---|
| siTBX2 #01 | GAUCGACAACAACCCGUUUtt | AAACGGGUUGUUGUCGAUCtt |
| siTBX2 #02 | GCAUGUACAUCCACCCAGAtt | UCUGGGUGGAUGUACAUGCgt |
| siTbx2 #05 | GCCAAAUACAUCCUGCUAAtt | UUAGCAGGAUGUAUUUGGCtt |
| siTbx2 #06 | GGAUUUGAUGCAAAUUUCUtt | AGAAAUUUGCAUCAAAUCCtt |
| siTbx2 #07 | ACAUCUUGAAGCUCCCAUAtt | UAUGGGAGCUUCAAGAUCUca |
| siTBX2 #13 | UGGACAUUGUAGCCGCUGAtt | UCAGCGGCUACAAUGUCCAtc |
| siTbx2 #14 | CCAAAUACAUCCUGCUAAUtt | AUUAGCAGGAUGUAUUUGGtt |

* siCN is MISSION® siRNA Universal Negative Control #1 (Sigma SIC001) from Sigma.

### 2.7.3 Buffer recipes

#### ChIP-seq buffers

All ChIP-seq buffers were filtered (0.45 µm), then stored at room temperature.

**ChIP lysis buffer**

| Component | Final concentration |
|---|---|
| Tris-HCl (pH8.0) | 50 mM |
| EDTA | 10 mM |
| Sodium butryrate | 10 mM |
| SDS | 1% (w/v in g/ml) |

Protease inhibitor cocktail (Roche by Sigma 05056489001) tablet was added to the buffer right before use (1 tablet/12.5 ml of lysis buffer).

**ChIP dilution buffer**

| Component | Final concentration |
|---|---|
| Tris (pH8.0) | 16.7 mM |
| NaCl | 167 mM |
| EDTA | 1.2 mM |
| SDS | 0.01% (w/v in g/ml) |
| Triton X-100 | 1% (v/v) |

**ChIP low salt washing buffer**

| Component | Final concentration |
|---|---|
| Tris-HCl (pH8.0) | 20 mM |
| NaCl | 150 mM |
| EDTA | 2 mM |
| SDS | 0.1% (w/v in g/ml) |
| Triton X-100 | 1% (v/v) |

**ChIP high salt washing buffer**

| Component | Final concentration |
|---|---|
| Tris-HCl (pH8.0) | 20 mM |
| NaCl | 500 mM |
| EDTA | 2 mM |
| SDS | 0.1% (w/v in g/ml) |
| Triton X-100 | 1% (v/v) |

**ChIP LiCl washing buffer**

| Component | Final concentration |
|---|---|
| Tris-HCl (pH8.0) | 10 mM |
| LiCl | 250 mM |
| EDTA | 1 mM |
| Sodium deoxycholate | 1% (w/v in g/ml) |
| NP-40 | 1% (v/v) |

**ChIP elution buffer**

| Component | Final concentration |
|---|---|
| $NaHCO_3$ | 100 mM |
| SDS | 1% (w/v in g/ml) |

ChIP elution buffer was made freshly right before use.

### WB buffers

**Gel buffers**

| Component | Stacking Gel Buffer | Resolving Gel Buffer |
|---|---|---|
| Water | 100 ml | 200 ml |
| Tris | 5.9 g (477 mM) | 36.2 g (1.465 M) |
| 20% SDS | 2 ml | 4 ml |
| Adjust pH to | pH6.8 | pH8.9 |

**RIPA buffer**

| Component | Concentration |
|---|---|
| NaCl | 150 mM |
| Tris | 50 mM |
| SDS | 0.1% (w/v in g/ml) |
| Sodium deoxycholate | 0.5% (w/v in g/ml) |
| Triton X-100 | 1% (v/v) |

**2X Laemmli Buffer**

| Component | Concentration |
|---|---|
| Tris-HCl | 125 mM |
| SDS | 4% (w/v in g/ml) |
| Bromophenol blue | 10% (v/v) |
| β-mercaptoethanol | 10% (v/v) |
| Glycerol | 20% (v/v) |

**10X Transfer Buffer**: 144 g of glycine and 38 g of Tris dissolved in MiliQ water.

**Transfer Buffer**

| Membrane | Dilution |
|---|---|
| Nitrocellulose | 10x buffer:water:isopropanol or methanol = 1:7:2 |
| PVDF | 10x buffer:water:methanol = 1:7:2 |

**Ponceau Staining Buffer**: 1 g of Ponceau S, 50 ml of acetic acid in 1 L of water. **TBS/T**: 0.1% Tween20 in 1x PBS

**Cell cycle solution**

**Stock Solutions recipe:**

- PI stock (1 mg/ml)
- Staining solution SS5 (5% FBS, 2mM EDTA, 0.01% NaN3 in PBS)
- DNAse-free RNAse A (10 mg/ml)

To make 1 ml of fresh PI Staining Solution:

- 50 µg/ml PI stock (50 µl of the 1 mg/ml stock)
- 0.1 mg/ml RNAse A (10 µl of the 10 mg/ml stock)
- 0.05% TritonX100 (0.5 µl)
- In PBS (940 µl)

### 2.7.4 Reagents and Kits

**qRT-PCR**

Kits used are RNeasy mini kit (QIAGEN, 74106), QuantiTect Reverse Transcription Kit (QIAGEN, 205313) and SYBR green mix (Promega, No.M7123).

Lysis buffer: 10 µl/ml of β-mercaptoethanol (β-ME) in 'RLT buffer'; Ethanol was added into RPE buffer before use.

**Antibodies**

| **Primary antibody** | **Catalogue Number** |
|---|---|
| Actin | A4700 |
| ERK-2 | sc-1647 |
| E2F1 | cst3742 |
| CDK2 | sc-163 |
| FLAG | F1804 |
| GAPDH | G9545 |
| H3K9me3 | ab8898 |
| HA | CST3724 |
| c-Met | cst3127 |
| PTEN | 32-5800 |
| Tbx2 (mouse) | home-made |
| TBX2 | HPA008586 |
| TBX3 | ab99302 |
| Tubulin | ab11308 |

**Chemicals and Reagents**

| **Chemicals** | **Catalogue Number** |
|---|---|
| Bovine serum albumins | Sigma, A9647 |
| DAPI | Invitrogen, D1306 |
| Doxycycline | Sigma, D9891 |
| Fetal bovine serum | Sigma, F4135 |
| 16% Formaldehyde (w/v) | Pierce, 11586711 |
| FuGENE6 | Promega, E2692 |
| IOX-1 | Abcam, ab144394 |
| JIB-04 | Sigma, SML0808 |
| Lipofectamine RNAiMAX | Invitrogen, 13778 |
| Opti-MEM | Gibco, 31985-062 |
| Propidium iodide | ThermoFisher, P3566 |
| RPMI 1640 media | Gibco, 21875-034 |
| Tacrolimus | Cayman, 10007965 |
| Cyclosporine A | Cayman, 12088 |

# 3

# Tagging endogenous Tbx2 using the CRISPR/Cas9 system

## 3.1 Introduction

One of my project objectives is to identify target genes of TBX2 and TBX3 through chromatin immunoprecipitation followed by sequencing (ChIP–seq), however, lack of ChIP-grade antibodies against endogenous TBX2 and TBX3 might compromise the quality of ChIP-seq datasets. To circumvent this problem, I sought to tag endogenous TBX2 and TBX3 with influenza hemagglutinin (HA) epitope deploying CRISPR/Cas9 as a tool.

### 3.1.1 CRISPR/Cas9 for gene editing

Clustered, regularly interspaced, short palindromic repeat (CRISPR)-Cas9 is a widely used system for genome engineering. There are three essential elements to the system: an RNA-guided endonuclease, such as Cas9 and its variants; a single-guide RNA (sgRNA) that determines the targeting specificity of Cas9 and directs Cas9 to cut the complementary DNA site if it is immediately adjacent to a short sequence termed protospacer adjacent motifs (PAM). PAM sequences vary with different Cas9 proteins. For example, the commonly used *Streptococcus pyogenes* Cas9 recognises a PAM sequence of 5'-NGG, and its two endonuclease domains cleave the opposite strands of the target DNA 3 nucleotides upstream of the PAM sequence, leaving a double-strand break (DSB) (Sander and Joung 2014). If a homologous template is provided, the DSB can be repaired by a homology-directed repair (HDR) pathway. HDR can be used to integrate a DNA fragment into the genome, but its efficiency is much lower than non-homologous end joining (NHEJ). While canonical NHEJ (cNHEJ) can faithfully re-ligate the DSB, alternate NHEJ re-ligates the resected DSB ends and might cause random mutagenesis, therefore, it is largely used to disrupt gene function (Hsu et al. 2014).

In an endeavour to tag endogenous TBX2/3, I tried two different CRISPR systems, first one being CRISPaint (CRISPR-assisted insertion tagging), which is reported to be more efficient than classic CRISPR knock-in because the DSB is repaired by cNHEJ.

The CRISPaint system uses three vectors to introduce a tag to the C-terminus of endogenously encoded proteins. The first plasmid, named 'Target Selector', encodes Cas9 and a sgRNA that targets the gene of interest. The second 'Frame Selector' plasmid expresses Cas9 and a sgRNA that cleaves a third plasmid, the 'Universal Donor', which carries the insert protein and a selection marker, the puromycin resistance gene that encodes puromycin-*N*-acetyl-transferase. The

universal donor's Cas9 target site has a poly-G sequence that can be cleaved at three adjacent nucleotide positions, allowing for correct reading frame selection (Schmid-Burgk et al. 2016).

Cas9 cleaves the target gene and the donor plasmid respectively, creating two DSB sites with blunt ends. The DSB is then repaired by cNHEJ, inserting the universal donor into the genome. The frame selector has mCherry, which can be used to check transfection efficiency. Successfully tagged cells can be enriched by adding puromycin to the growth media, which kills cells without an insertion or with the donor inserted in the wrong orientation. Hence, no homology template is needed and the efficiency is higher than traditional CRISPR knock-in methods using HDR (Schmid-Burgk et al. 2016).

However, utilising cNHEJ also means that CRISPaint might to be more error-prone as the repair does not need a homology arm and thus there is no second check to prevent off-target effects. In addition, as the whole universal donor will be integrated at the C-terminus including the poly A signal, the endogenous 3' UTR is not included in the resulting mRNA, which might modify target gene expression at a post-transcriptional level. On the other hand, my results, not included in the thesis, indicate that the tagging efficiency of TBX2/3 using CRISPaint is low and the few 'tagged cells' did not proliferate as expected.

We therefore switched to a second CRISPR knock-in strategy – eFlut (endogenous Fluorescent tagging) system. Similar to CRISPaint, eFlut has donor templates with a selection marker, making enrichment for successfully tagged cells easy. Unlike CRISPaint, which utilises cNHEJ, eFlut donor templates have a short homology arm (40 bp) and the DSB created by Cas9 will be repaired by HDR. As the homology arm is relatively short, it is easy to add them to PCR primers when amplifying the donor templates from a set of plasmids encoding various fluorescent proteins in combination with an antibiotic resistance gene.

### 3.1.2 eFlut tagging system

What makes the eFlut system superior to a classical CRISPR knock-in strategy is the simplicity of generating the repair template containing the homology sequence to loci of interest and inserts of choice. A traditional knock-in strategy utilises a plasmid containing homology arm longer than 1 kb, which thus requires a lot of cloning and can be time-consuming, whilst eFlut uses PCR primers whose 5' ends have around 40 nt of homology to the target sequence and the 3' end 20 nt anneal to a set of template plasmids (Figure 3.1.A). Though the short homology arm results in a very limited tagging efficiency (0.01% – 1%) (Stewart-Ornstein and Lahav 2016), selection with antibiotic markers will enrich for successfully modified cells even if they are not abundant.

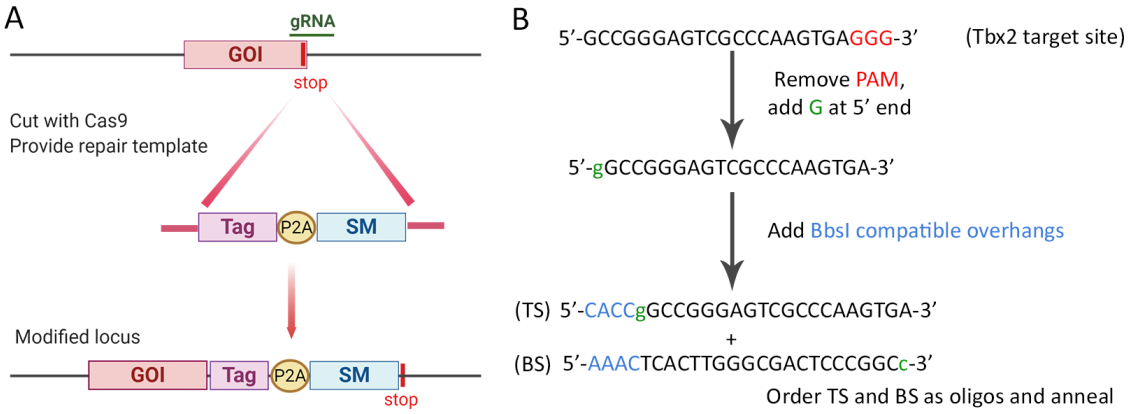



**Figure 3.1: eFlut tagging strategy** (A) Schematic illustration of the eFlut (endogenous Fluorescent tagging) system. GOI, gene of interest; SM, selection marker. (B) Example of sgRNA design targeting the C-terminus of murine Tbx2. The PAM site (red) was removed and BbsI compatible overhangs (blue) were added. An extra 'G' (green) was added to the transcription start site as required by a U6-promoter.

As shown in Figure 3.1.A, sgRNA guides Cas9 to cut near the stop codon of the protein of interest, leaving a DSB that will be resected and subsequently prime for HDR. If the DSB is repaired using the PCR products as template, a tag will be fused to the C-terminus of your protein of interest, while the antibiotic marker will be cleaved off the tag and allow for fast selection of correctly tagged cells. The expression of the selection marker relies on the endogenous promoter, and there is

no extra terminator introduced, hence the endogenous 3' UTR still regulates the expression of the protein of interest (Stewart-Ornstein and Lahav 2016).

In the case of tagging the C-terminus of TBX2/3 with 3xHA or fluorescent protein mVenus or mScarlet, the experiment design is briefly as follows:

- *sgRNA-Cas9* sgRNA was designed to target the region right after or within the stop codon of TBX2/3, depending on the availability of the PAM sequence. For the sgRNA sequence, I removed the PAM sequence from the target site and added BbsI compatible overhangs. An example of the sgRNA design process can be found in Figure 3.1.B. Then I ordered top strand and bottom strand of the designed sgRNA as oligos, annealed them, and inserted the sgRNA into the digested px330 vector.
- *Repair Template* An insert encoding a tag protein (3xHA or fluorescent protein) and aminoglycoside 3'-phosphotransferase, which confers resistance to eukaryotic cells to antibiotic G418, was amplified by PCR primers that add 40 nt of homology to the C-terminus and the 3' UTR of TBX2/3 respectively.
- *Transfection* Cas9 plasmid and linear double-strand repair template were transfected into melanoma cells with the LT1 reagent.
- *Selection* Unmodified cells were eliminated by G418 selection while successfully tagged cells were enriched.

## 3.2 Optimisation of tagging methods

CRISPR knock-in is known to be difficult because it relies on inefficient HDR for accurate repair, which comes into action mainly during S or G2- phase and thus has a low efficiency. Adding to that, the short homology arm used in eFlut further decreased the HDR rates. In this subsection, I will use tagging TBX2 with 3xHA in the IGR37 melanoma cell line as an example to showcase how eFlut works and how I optimised the system for best tagging efficiency.

My first batch of tagging experiments had some cells surviving G418 selection, however, no TBX2-tag fusion was observed despite clear evidence of Neomycin resistance gene insertion. To understand what kind of modification the locus had undergone, I PCR amplified the last exon–3' UTR of TBX2 and sequenced this genomic DNA region. To my surprise, the whole repair template sequence, including the homology sequence that is intended to guide accurate repair, was inserted into the TBX2 3' UTR leaving endogenous TBX2 stop codon intact, and causing a separation of the tag from TBX2 and alternate expression of the Neomycin resistance gene.

The sequencing results indicated that the DSB caused by Cas9 was repaired via NHEJ rather than desired HDR. Therefore, I optimised the sequence design of sgRNA and homology arm seeking to improve HDR efficiency.

### 3.2.1 Cut-to-modification distance

After Cas9 introduces a DSB in the locus of interest, the 5' end is resected creating a 3' overhang, which will invade into the homologous template and primes for DNA repair synthesis. The resected ssDNA vary in length and can proceed to a few kilobases (Symington 2016). Hence, in theory the desired insertion site, which is determined by homology arm, can be of some distance from the Cas9 cut site. However, an inverse relationship between rate of point mutation incorporation and distance from Cas9 cleavage site has been reported — HDR efficiency decreases as 'cut-to-modification' distance increases. The mutation incorporation rate drops down to nearly zero when the cut-to-modification distance is more than 35 bp, while most successful modification features homology arm starts right at the Cas9-induced DSB site (Paquet et al. 2016).

For selecting the sgRNA target sequence, I used the MIT CRISPR tool. Among the top rated sequences, I first chose one targeting the 3' UTR of TBX2 that gives the best on-target activity (Figure 3.2.A) but which cuts 57 bp away from the desired

insertion site. After realising this cut-to-insertion distance makes HDR almost impossible, I tried a second sgRNA target sequence which cleaves immediately before the TBX2 stop codon, making the distance 0 bp (Figure 3.2.B).

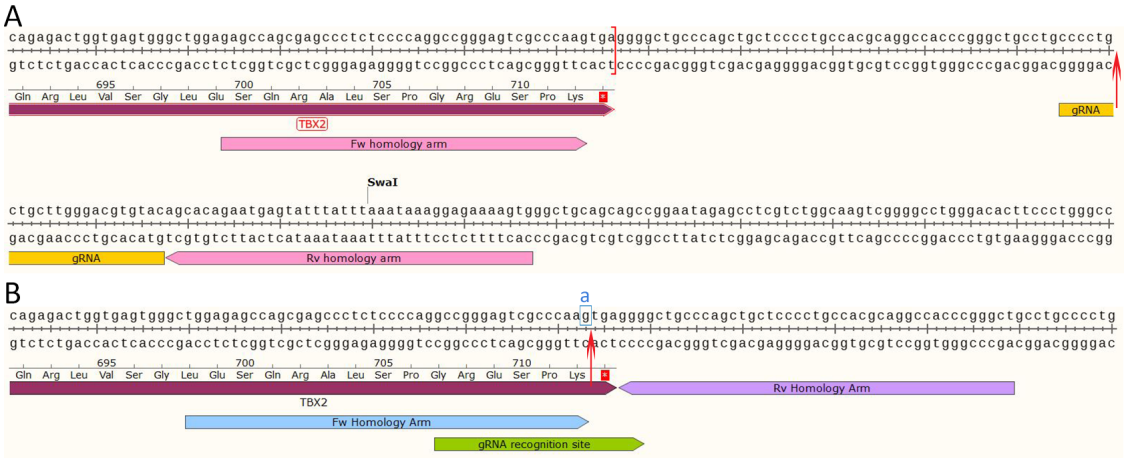



**Figure 3.2: Optimisation of eFlut experiment** (A) Initial design of sgRNA and homology arm sequence to tag the C-terminus of TBX2, where the cut-to-modification distance is 57 bp. (B) Optimised design of sgRNA and homology arm sequence to tag C-terminus of TBX2, where cut-to-modification distance is 0 bp. The red arrow points to the Cas9 cut site; The blue box shows the base where a silent mutation was introduced to block re-cleavage, The red box with * sign shows the position of the TBX2 stop codon.

### 3.2.2 Blocking mutation to prevent re-cleavage

As ectopic Cas9 expression will last a few days in cells, re-cleavage presents another obstacle for accurate modification. A study intended to improve CRISPR/Cas9 editing by HDR showed that the desired point mutation is often impaired by additional indels caused by NHEJ following Cas9 re-cutting at the same site, and that re-cleavage can be avoided by introducing a blocking mutation in the PAM site or the sgRNA target sequence (Kwart et al. 2017). Though in my case, inserting a long sequence after the cut site naturally deletes the PAM site and subsequently lowers re-cutting probability, in the homology sequence I still introduced an additional silent mutation at the most proximal base upstream of the cut site in the sgRNA target sequence to eliminate re-cleavage (Figure 3.2.B).

After minimising the cut-to-modification distance to improve HDR efficiency

and introducing blocking mutations to prevent re-cleavage, I finally succeeded in tagging TBX2 with 3xHA inhuman IGR37 cell line and mouse B16 cell line.

## 3.3 Validation of eFlut-tagged Cell Lines

The purpose of tagging endogenous TBX2/3 was to enable ChIP-seq experiment and to validate drug screen hits. ChIP-seq requires a reliable antibody to efficiently immunoprecipitate the protein-DNA complex while minimising non-specific binding. The 3xHA epitope tag was chosen because of its small size and availability of ChIP-grade HA antibody, and our lab had previously used the HA to successfully achieve an high efficiency ChIP-seq using the MITF protein (Louphrasitthiphol et al. 2020). On the other hand, we would prefer a high throughput image-based assay to conduct unbiased screen for small molecules that can alter TBX2/3 protein level for potential use in pro-senescence therapy, thus a fluorescent tag can be beneficial as it can save the cost for antibodies and would enable real-time imaging to be achieved, both in monolayer and in 3D culture.

Though I optimised the tagging method to achieve better efficiency, the outcome varies among different cell lines as some of them are difficult to transfect. In addition, insert length seems another element affecting HDR efficiency, making each tag a different case. To date, I tried 3xHA, mVenus, mScarlet and mNeon to tag TBX2 in 501mel, IGR37 and SKmel30 human melanoma cell lines, the B16 mouse melanoma line, and the MCF7 breast cancer cell line. I also tried TBX3 in 501mel, IGR39, WM266.4 and CHL1 human melanoma cell line. To date I have successfully generated and characterised two clones of B16-Tbx2-3xHA-NeoR and one clone of IGR37-3xHA-NeoR. The results are presented below.

### 3.3.1 Validate TBX2-3xHA protein fusion

After selecting with G418 for a week, unmodified cells were eliminated. From the surviving cells, I picked 11 clones of B16 cells and 5 of IGR37 and used WB to check if they express HA-tagged TBX2. The expected M.W. of endogenous TBX2 is 75 kDa, and adding a 3xHA tag will make the TBX2-3xHA fusion 79 kDa. The WB results showed that three clones of B16 (Figure 3.3.A) and 1 clone of IGR37 (Figure 3.3.B) have bands recognised by anti-HA antibody at the expected size, and the TBX2 band identified using anti-TBX2 antibody in these clones is shifted up, confirming the addition of the HA tag. Among the three clones of B16-Tbx2-3xHA-NeoR, cl#02 has two different Tbx2 bands, one at the endogenous ~ 75 kDa, the other at the tagged ~ 79 kDa size, indicating a heterozygous modification, while cl#06 has an unexpected genomic DNA sequencing result, which I will expand on later, and was excluded from other validation assays.

To confirm that HA was indeed fused to TBX2 and that the changes in TBX2 protein level can be reflected by probing for HA, I used siRNA to knockdown TBX2 and did WB to test if the HA and TBX2 antibodies give same trend. The results were positive — two different siRNAs against Tbx2 diminished the Tbx2 band in B16 parent cells as well as B16-Tbx2-3xHA-NeoR cl#02 and cl#09, whilst the HA antibody has same results as the Tbx2 antibody in the two tagged clones, but has no band in the parental cells (Figure 3.3.C) — further confirming the fusion of the HA tag to Tbx2 protein. In IGR37-TBX2-3xHA-NeoR cl#02, the HA band diminished after TBX2 was knocked-down by two different siRNAs (Figure 3.3.D), however, the antibody against human TBX2 gave such a weak signal that was not ideal for validating the protein fusion. On the other hand, the inability to reliably detect human TBX2 highlights the advantage in having an endogenously HA-tagged TBX2 protein.

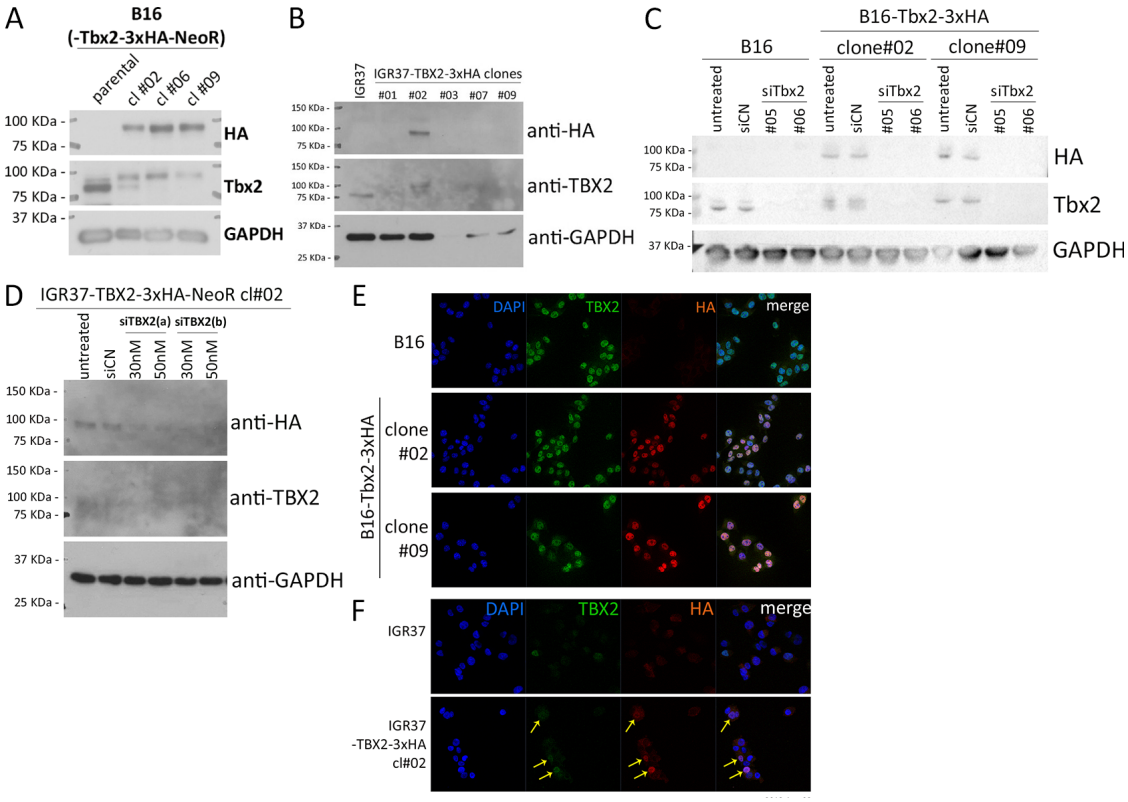



**Figure 3.3: Validate TBX2-3xHA protein fusion and subcellular localisation** (A) WB result from B16-Tbx2-3xHA-NeoR clones and parental B16 cell lysates. Unspecific Tbx2 band lower than 70 kDa not shown here. (B) WB result from IGR37-TBX2-3xHA-NeoR clones and parental IGR37 cell lysates (C) WB result from B16-Tbx2-3xHA-NeoR cl#02 and 09 and parental B16 cells transfected with scramble siRNA or siTbx2 (D) WB result from IGR37-TBX2-3xHA-NeoR cl#02 cells transfected with scramble siRNA or siTBX2 (E-F) Immunofluorescence images of B16-Tbx2-3xHA-NeoR cl#02 and 09, IGR37-TBX2-3xHA-NeoR cl#02, and their parental cell lines, co-stained with TBX2 and HA antibodies. DAPI marks nuclear area.

### 3.3.2 Validation of TBX2 subcellular localisation

I next asked whether the addition of HA tag affects TBX2 subcellular localisation. Being a transcription factor, TBX2 is mainly found to be nuclear — otherwise it will not be able to bind its target genes. Hence, it is important that we ensure that tagging does not influence its subcellular location and consequently its function.

The IF results show that in B16 parental cells, Tbx2 is in the nucleus and there is no HA signal, while in B16-Tbx2-3xHA-NeoR cl#02 and cl#09, Tbx2 stays nuclear and is co-localised with the HA staining (Figure 3.3.E). In IGR37-TBX2-3xHA-NeoR cl#02, the nuclear co-localisation of TBX2 and HA staining was also observed,

though TBX2 expression level is relatively low and heterogeneous among the cell population (Figure 3.3.F). The source of the heterogeneity is unclear, but it is known that TBX2 may be regulated during the cell cycle (Bilican and Colin R Goding 2006).

Therefore, we chose to use the two B16 clones for the ChIP-seq since high and homogeneous expression of TBX2 is preferred in order to capture abundant protein-DNA complex and ensure the enriched signal represents the whole population. Apart from a better expression level in B16, performing the experiment in two clones will enable us to rule out potential artefacts arising from clone picking and thus increase our confidence in the newly identified binding sites.

### 3.3.3 Validation of genomic DNA modification

Knowing that HA was fused to TBX2 and the subcellular localisation remained unchanged, I went on to ask if the C-terminal tag was inserted as planned and if there was any unwanted modification at the locus. To answer this, I PCR amplified the last exon-3' UTR region of TBX2 and sent the PCR products for sequencing.

The scheme of genomic DNA (gDNA) PCR is shown in Figure 3.4.A. Using primer pairs F1/R1 I amplified the region flanking the stop codon where the insertion happens, resulting in a short product in parental cells (691 bp for B16 and 728 bp for IGR37), and a longer product for the modified locus (~1.7 kb), indicating the insertion of the HA tag and selection marker. Next, I performed nested PCR on the gDNA PCR products using primer pairs F2/R3 and F3/R4, which allowed me to further enrich the modified loci and eliminate potential artefacts arising from the F1/R1 primers. In nested PCR, no product is expected in parental cells, while primer sets F2/R3 and F3/R4 should have 1.1 kb and 1.4 kb products in B16-Tbx2-3xHA cells, and 1 kb and 1.1 kb products are expected in IGR37-TBX2-3xHA cells.

The gDNA and nested PCR results of B16 are shown in Figure 3.4.B. In gDNA PCR experiment, B16 parental cells only have one endogenous band, while B16-

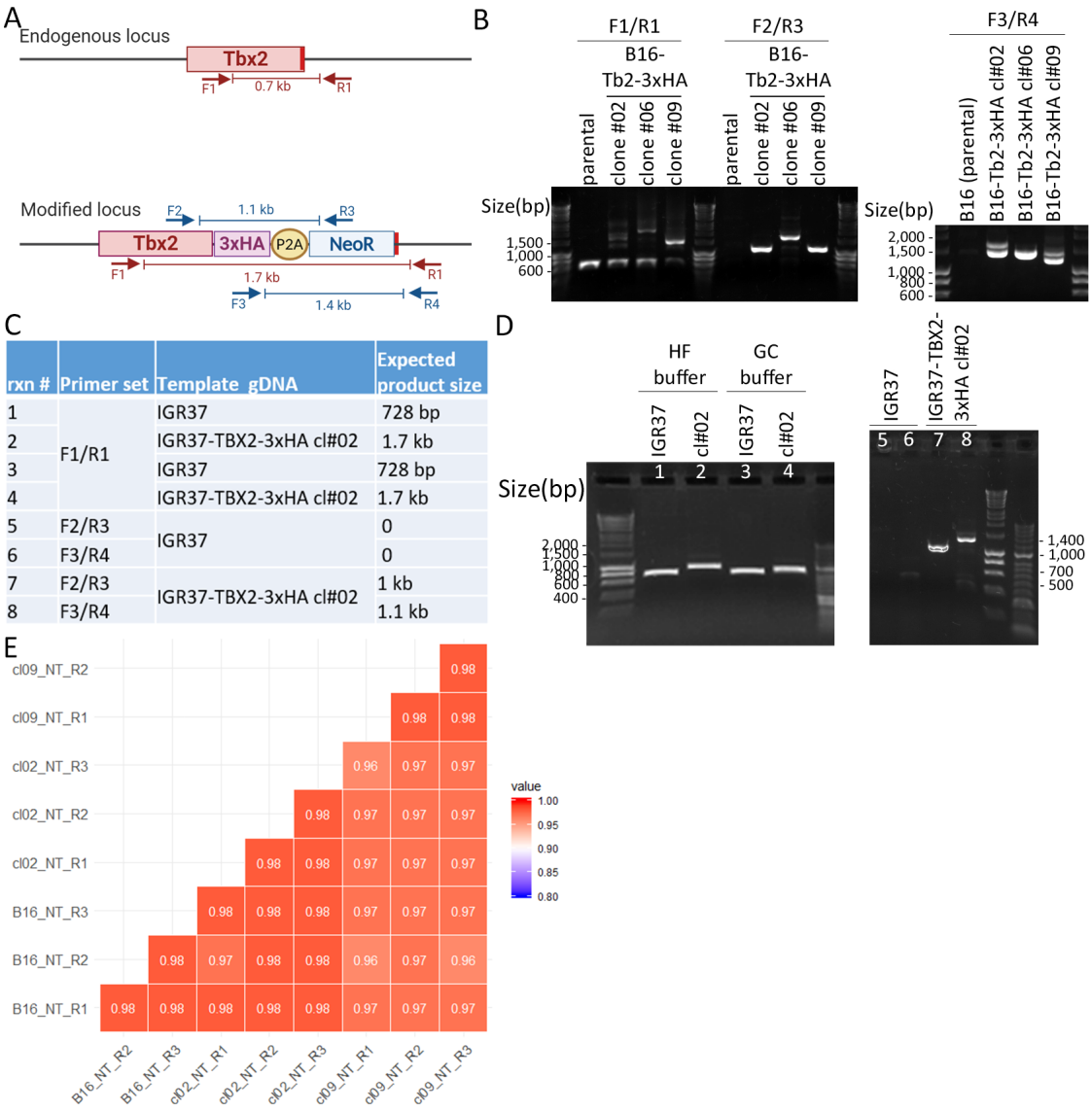



**Figure 3.4: Validate genomic DNA modification** (A) Schematic illustration of genomic DNA (gDNA) PCR and expected PCR product lengths for unmodified and modified loci, using Tbx2 tagging in B16 cells as an example. (B) gDNA PCR and the following nested PCR products, from B16-Tbx2-3xHA-NeoR clones and parental cells, were run on an agarose gel. (C) Expected PCR product lengths for parental and tagged IGR37 cells (D) gDNA PCR results using different primer sets, from IGR37-TBX2-3xHA-NeoR cl#02 and parental cells, were run on an agarose gel. (E) Correlation matrix comparing the HA-tagged clones and the parental B16 cells using data from 3' mRNA-sequencing (Spearman). Biological triplicate samples were sequenced.

Tbx2-3xHA-NeoR cl#09 has one modified band, and cl#02 and cl#06 have both bands, indicating a heterozygous modification. Notably, the longer product of cl#06 appears to be longer than 1.7 kb, which might be a result of NHEJ happening either

side of the insert. This is also reflected in nested PCR F2/R3 — cl#06 again has a longer product, suggesting the extra insert is located upstream the tag and changes Tbx2 protein sequence. On the other hand, the F2/R3 nested PCR had expected product length in cl#02 and cl#09, and the PCR products were sent for sequencing. The sequencing results showed that where the HA tag joins Tbx2 C-terminus the insertion is seamless and 100% accurate. However, the F3/R4 nested PCR had mixed products in cl#02 and cl#09, indicating different modifications on two alleles. Unfortunately F3/R4 products sequencing results from cl#02 and cl#09 did not cover the insertion site, and I was unable to know how the locus is modified precisely, though cl#06 was shown to have a seamless insertion of NeoR in the Tbx2 3'UTR. Regrettably none of these three clones are perfect, but importantly cl#02 and cl#09 achieved the desired fusion protein sequence, and WB results suggest that the partially modified 3' UTR did not drastically change Tbx2 expression (Figure 3.3.A).

The expected gDNA PCR products of the tagged and parental IGR37 cells are shown in Figure 3.4.C. Using F1/R1, IGR37-TBX2-3xHA-NeoR cl#02 did have a longer PCR product compared to parental cells, but it seems only ~1 kb and much shorter than expected 1.7 kb (Figure 3.4.D). Contemplating that this region might contain sequence that is difficult to amplify, I tried using the F2/R3 and F3/R4 primer pairs to amplify two fragments from the genomic DNA which jointly cover the whole modified locus, and obtained good results. As expected, parental IGR37 cells did not have any band, while cl#02 has bands from both primer sets (Figure 3.4.D). These PCR products were sent for sequencing, and the results show seamless and 100% accurate modification on both side of the insert.

Although the double bands in B16-Tbx2-3xHA-NeoR cl#02 gDNA PCR products might also mean it is a mixed population, the IF results indicate that every cell is positive for HA staining, indicating it is not a mix of modified and parental cells (Figure 3.3.E).

### 3.3.4 Validation of TBX2 function

HA-tagged cell lines were generated for ChIP-seq experiment, therefore it is important to ensure the tagging does not affect Tbx2 binding to DNA and interacting with cofactors. We already know from IF that tagged Tbx2 remains nuclear, but nuclear presence does not necessarily guarantee a normal function. We expect any change in Tbx2 binding or interacting capacity will affect the expression of target genes that can be detected by RNA sequencing.

We therefore performed 3' mRNA-sequencing in triplicate, and compared the gene expression profile of B16-Tbx2-3xHA-NeoR clones to the parental B16 cells. The sample correlation matrix shows that the HA tagged clones highly resemble the parental B16 cell line, with a coefficient ranging from 0.96 to 0.98 (Figure 3.4.E).

After validating the B16-Tbx2-3xHA-NeoR cell lines using WB, IF, gDNA PCR and sequencing, we are confident that the endogenously tagged cells are successfully modified and suitable for ChIP-seq and other experiments.

# 4

# Identification of TBX2 target genes

## 4.1 Introduction

TBX2, as a transcription factor, regulates its targets by binding to regulatory elements and then altering their expression. Identifying target genes of TBX2 is crucial to understanding its role in normal development and pathological processes. In this section, I will introduce the direct target genes of TBX2 that have been published, and discuss how TBX2 regulates their expression.

### 4.1.1 Known direct target genes of TBX2

**CDKN2A** $p14^{ARF}$ inhibits the p53 repressor MDM2 E3 ubiquitin ligase thus facilitating p53 activation. Jacobs et al. 2000 showed that TBX2 binds to the initiator region -19 to +54 of the human $p14^{ARF}$ (*CDKN2A*) promoter, but not the consensus

DNA binding site at position -106. Consistent to this report, Lingbeek et al. 2002 showed TBX2 can directly repress the human *CDKN2A* promoter (-19 to +54) as well as that of the murine orthologue $p19^{Arf}$ in COS-7 cells and MEFs. Electrophoretic mobility shift assay (EMSA) shows that TBX2 and TBX3 both bind a variant T-site in the $p14^{ARF}$ *in vitro*, and this variant T-site is required for the repression.

**CDKN1A** TBX2 binds to and represses the promoter of the CDK2 inhibitor $p21^{CIP1}$ (*CDKN1A*), which contains a consensus T-element close to the initiator (Vance, Suzanne Carreira, et al. 2005; Prince et al. 2004). By repressing $p21^{CIP1}$ expression, TBX2 can reduce CDK2-mediated repression of Rb1, an inhibitor of E2F transcription factors that activate S-phase genes (Sharpless and Sherr 2015). $Tbx2^{-/-}$ embryos have increased Cdkn1a and Cdkn1b expression, and ChIP-PCR experiment showed Tbx2 binds the *Cdkn1a* and *Cdkn1b* locus containing putative T-elements (T. Lüdtke et al. 2013). Ectopic expression of TBX2 in breast epithelial cells reduced p21 expression and rescued TGF-β1 induced growth inhibition (J. Li, D. Ballim, et al. 2014).

**TRP-1** Tbx2 can bind the melanocyte-specific elements located in the tyrosinase-related protein 1 (TRP-1) promoter and regulate TRP-1 expression, as shown by luciferase assay and *in vitro* EMSA (Carreira et al. 1998).

### TBX2 target genes identified in cancer

**PTEN** PTEN is an negative regulator of PI3K signalling that lies upstream from mTORC1 and serves as a tumour suppressor. Zhu et al. 2015 reported that in rhabdomyosarcoma and normal muscle cells TBX2 directly binds to and represses the *PTEN* promoter, while TBX3 does not inhibit PTEN. The binding site of TBX2, identified by luciferase assay and ChIP-qPCR, locates between position -1,345 and -710 in the *PTEN* promoter.

**CDH1** In mouse mammary epithelial cells, TBX2 binds to the most proximal *CDH1* promoter region, which contains the initiator element, and represses its

activity as shown in ChIP-PCR and luciferase assay (B. Wang et al. 2012).

**NDRG1** N-myc down-regulated gene 1 (NDRG1) is a breast tumour suppressor involved in stress responses, hormone responses, cell growth, and differentiation. Silencing of endogenous TBX2 in breast cancer cells or expression of a dnTBX2 lead to increased NDRG1 mRNA and protein levels, while overexpression of TBX2 does the opposite. Luciferase assay and ChIP-PCR showed TBX2 directly binds to the proximal *NDRG1* promoter and represses its expression (Redmond et al. 2010). The same group later showed that TBX2, together with HP1 proteins (α, β and γ), binds to the *NDGR1* promoter (-80/+4) in MCF7 cells as shown in ChIP-qPCR (Crawford et al. 2019).

**CST6** Cystatin 6 (CST6) is a cystein protease inhibitor that is down-regulated in breast cancer samples and postulated to be a tumour suppressor. Silencing of TBX2 or expression of dnTBX2 increases CST6 mRNA level, and in turn, exogenous expression of CST6 reduced the growth of TBX2-expressing breast cancer cells by inducing apoptosis. Luciferase assay shows that TBX2 represses the *CST6* proximal promoter (-101 to +30) (D'Costa et al. 2014).

**ADAM10** The ADAM (a disintegrin and metalloproteinase)-mediated cleavage of transmembrane proteins regulate important biological functions such as cell fate determination, migration and proliferation. Reinhardt et al. 2019 showed that TBX2 represses the *ADMA10* promoter (-315 to -286) activity and its expression in human neuroblastoma cells, which is mediated by two variant T-elements within the *ADAM10* core promoter sequence. Moreover. TBX2 deletion mutant that lacks the C-terminal repression domain or the K122/123E mutants that is deficient for DNA-binding fail to repress the *ADAM10* promoter.

**WNT3A** TBX2 binds to the *WNT3A* promoter and activates its transcription by recruiting RNA Pol II, as evidenced by ChIP-qPCR, and thus promotes prostate cancer progression. (Nandana et al. 2017).

### TBX2 target genes identified in development

**Grem1** Gremlin 1 (*Grem1*) is a key regulator in restricting limb bud growth and avoid digit malformation. Tbx2 and Grem1 are exclusively expressed in the hindlimb. Tbx2 binds to the regulatory element of *Grem1* as shown in ChIP-PCR, and ectopic Tbx2 expression represses Grem1 in the limb (Farin et al. 2013).

**Cx40** and **Cx43** Tbx2 represses the promoter activity of chamber-specific genes *Connexin (Cx) 40 and Cx43*, and consequently inhibits chamber formation, which relies on both the T-box DNA-binding domain and the repression domain (Christoffels et al. 2004).

**Wnt signalling effectors** In the developing murine ureter, Tbx2 directly represses the Wnt signalling antagonist Dkk2 and Shisa2, the Bmp antagonist Bmper and the chemokine Cxcl12, as shown in ChIP-PCR and microarray gene expression profiling (Aydoğdu et al. 2018). Apart from ureter, T. H. Lüdtke et al. 2016 showed in the lung mesenchyme, Tbx2 directly represses Frzb and Shisa3, the Wnt antagonists, by binding to two different TBEs, as evidenced by ChIP-PCR and microarray experiments.

### TBX2/3 target genes in viral infection

The minor capsid protein L2 is important in the life cycle of human papilomaviruses (HPVs). The long control region (LCR), which comprises an upstream enhancer and early promoter region, regulates expression of the early viral genes during HPV infection. Co-IP experiments showed that TBX2 and TBX3 interact with L2, and IF confirmed their colocalization in the cell nucleus (Farin et al. 2013). ChIP-PCR showed that TBX2 binds to the HPV16 and HPV18 LCRs. Luciferase assay shows TBX2 and TBX3 repress the LCR of HPV11, -16 and -18, with the TBX2-responsive element locating between nt 7450 and 7760 of the HPV16 LCR. $TBX2^{R122/123E}$ mutant that is deficient for DNA-binding cannot repress the

LCR but still interacts with L2.

### 4.1.2 Identifying novel TBX2 target genes

To date few validated TBX2 target genes have been identified, and our experience indicated that using antibodies to immunoprecipitate endogenous transcription factors can frequently lead to poor-quality ChIP-seq datasets. Indeed, while some studies have performed ChIP-seq on TBX2 using anti-TBX2 antibody (Decaesteker et al. 2018), the antibody used was not well validated and the quality of the ChIP-seq dataset obtained was therefore compromised.

To expand the repertoire of genes regulated by TBX2/3, we carried out ChIP-seq of Tbx2 and TBX3. A previous DPhil student Johannes tried to undertake a ChIP-seq with 501mel melanoma cells expressing Flag-tagged TBX2 and TBX3. However, the potential target genes found in the ChIP-seq could not be validated. This might be because: the ChIP-seq peaks are not very high, indicating the experiment was not ideal; the experiment was done with overexpressed TBX2/3, which might have saturated any cofactors and thereby affected the results. To circumvent these issues, I made new cell lines with TBX2/3 fused to a 3xHA Tag, which from our experience has less antibody cross-reaction and therefore is potentially superior to FLAG tag in ChIP-seq experiments. To date, I have established a B16 cell line with endogenously tagged Tbx2, and obtained two good clones as described in Chapter 2.

Next, I tested if the genes directly bound by Tbx2 are regulated by Tbx2 at the transcriptional level by 3' mRNA-seq following Tbx2 knockdown. I was then able to conduct an integrative analysis of Tbx2 binding profile and gene expression profile to explore its target genes in melanoma progression.

## 4.2 Tbx2 Binding Sites - ChIP-seq

T, the first discovered T-box family member, binds a partially palindromic sequence (T(G/C)ACACCTAGGTGTGAAATT), consisting of two half sites to which the protein binds as a monomer (Kispert and Herrmann 1993; Khan et al. 2019), whilst binding sites of other members were identified as the consensus half site (AGGTGTGAAA), known as the T-box binding element (TBE) or T-element (Papaioannou 2014). Indeed, it has been shown by a ChIP-seq experiment that TBX2 also binds TBE (AGGTGTGA) (JASPAR matrix ID: MA0688.1). This experiment was done with overexpressed TBX2 in a 293FT cell line (Jolma, Yan, et al. 2013). Whether endogenous TBX2 in melanoma cells binds similar sites and regulates the same set of genes and how the binding pattern affects the role of TBX2 in tumour progression remains unknown.

To address this question, I performed chromatin immunoprecipitation coupled to Illumina DNA sequencing (ChIP-Seq) using two clones of B16 cells that have endogenous Tbx2 tagged with 3xHA — B16-Tbx2-3xHA-NeoR cl#02 and cl#09 — the generation of which was described in the previous Chapter. Biological replicates were done for each clone together with input control, with each replicate consisting of 15 150-mm dishes of cells.

### 4.2.1 Tbx2 ChIP-seq

#### Tbx2 ChIP-seq sample preparation

Experimental details are presented in section 2.5.1 on page 54. Briefly, DNA-protein complexes were cross-linked using 0.8% formaldehyde, and chromatin was sheared by sonication into DNA fragments of proximate 200–300bp. The insoluble fraction was then removed by high-speed centrifugation. Complexes containing Tbx2 in the soluble fraction were immunoprecipitated using anti-HA antibody and magnetic Dynabeads. Protein and RNA were digested using enzymes, and

DNA fragments purified and sequenced by Illumina Truseq in the Wellcome Trust Centre for Human Genetics.

#### Tbx2 ChIP-seq analysis

A detailed analysis pipeline is presented in section 2.6.2 on page 62. Briefly, raw sequencing data were mapped to mm10 using bowtie2, peaks called using MACS2, and motifs detected using HOMER.

### 4.2.2 Tbx2 ChIP-seq results

#### ChIP experiment validation

First, we need to ensure that using HA antibody and magnetic beads successfully immunoprecipitated Tbx2-3xHA and the sonication sheared DNA to the desired fragment size.

The WB results showed that in both replicates there were bands at the expected size in the IP samples, indicative of a good IP efficiency in B16-Tbx2-3xHA-NeoR clones, and that there is no obvious background noise in B16 parental cells, as shown clearly using anti-HA antibody (Figure 4.1.A). The Tbx2 antibody gave a relatively weak signal, nevertheless in tagged clones the band size is correct and the intensity pattern agrees with the results obtained using anti-HA antibody, whilst in parental cells there are clear bands only in the insoluble fraction and the input which show an unmodified protein weight 75 kDa (Figure 4.1.A).

Tapestation results show that the fragment length for the input and ChIPed DNA peaks at around 250 and 210 bp respectively (Figure 4.1.B), which is ideal for the paired-end 150 bp flow cell type we chose, as it would give good coverage for DNA fragments under 300 bp. The peak appears to be tailing towards a longer size (Figure 4.1.B), indicative of excessive crosslinking. This is not concerning, because long DNA fragments do not affect peak summit location, though they

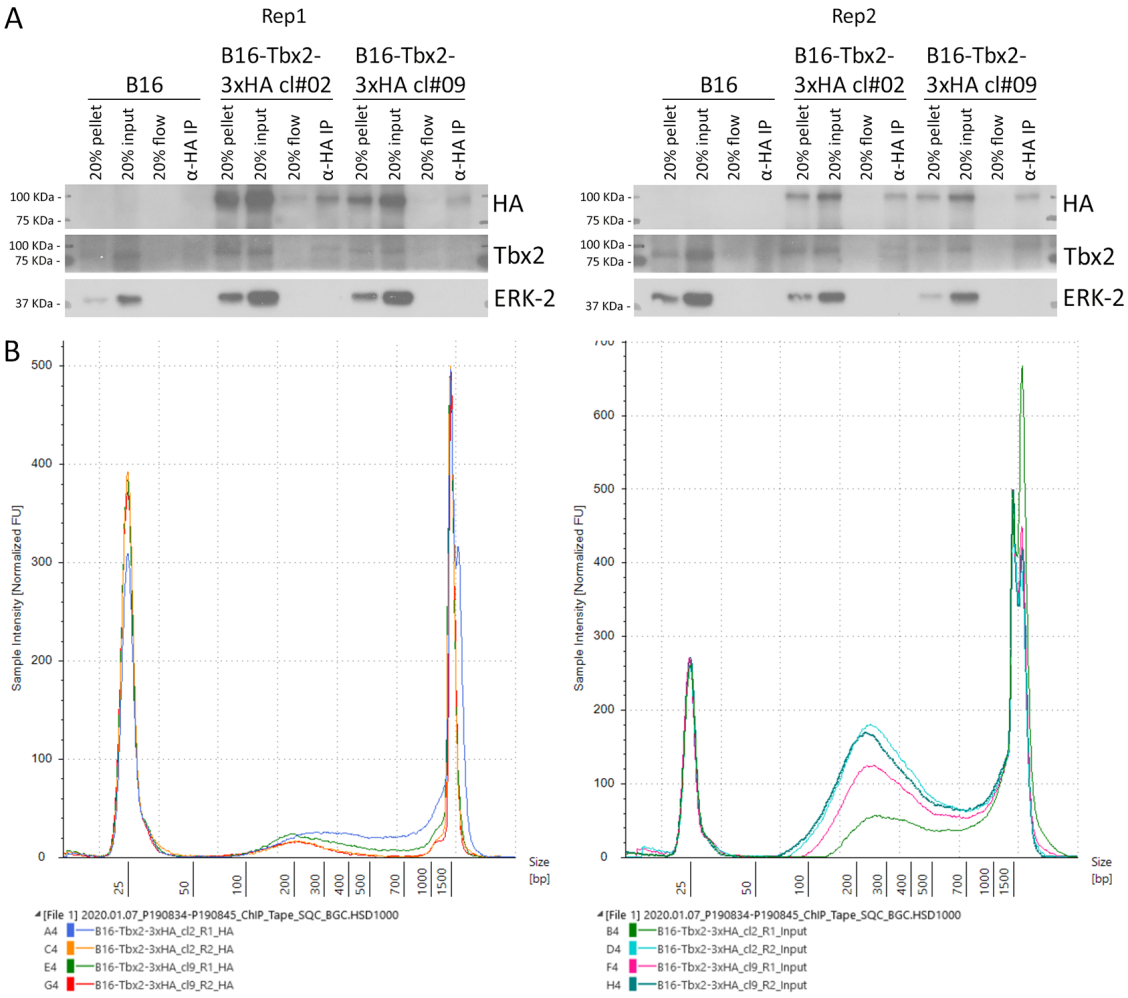



**Figure 4.1: Tbx2 ChIP-seq experiment** (A) WB results from HA immunoprecipitation of B16-Tbx2-3xHA-NeoR clones and parental B16 cells using indicated antibodies. Anti-ERK2 was used as a loading control for the input samples. (B) DNA fragments after sonication were run in TapeStation for size checking.

might lead to broader peaks. In addition, from my experience, over-crosslinking is likely to be necessary for Tbx2 ChIP-seq, because previous TBX3 ChIP-seq experiments using milder cross-linking (not included in this thesis) failed to produce good ChIP peaks despite successful IP, suggesting the immunoprecipitated TBX3 protein is not bound by DNA. In this experiment, I first trypsinised the cells, then resuspended and at the same time crosslinked them in 0.4% PFA, which is the standard condition optimised by Dr Pakavarin Louphrasitthiphol in our group for the MiT/TFE Family of Transcription Factors. Realising this might not be optimal for T-box transcription factors, I tried crosslinking using 0.8% formaldehyde while

the cells are still attached to the dish, and found this method superior than the 0.4% PFA crosslinking procedure. The fact that excessive crosslinking is needed might indicate that TBX2 and TBX3 only bind DNA when interacting with other DNA-binding factors, for example transcription factors or histone modifying enzymes, or they dissociate from DNA so quickly that the trypsinization process was long enough for TBX2/3 to come off their binding sites.

**ChIP-seq quality check**

After obtaining the sequencing data, I first checked the quality of the raw reads, and found some adaptor contamination at the 3' end. Therefore, I trimmed the adaptor sequence using Cutadapt and obtained sequences with a contamination rate lower than 0.1%. I then mapped the trimmed sequence from the input and ChIP samples to the mm10 reference genome using Bowtie2, and subsequently peaks were called using MACS2.

To assess the quality and reproducibility of the peaks, I used ChIPpeakAnno to overlap peaks called from each replicate. We can see that in both clone#02 and clone#09, while replicate 2 called more peaks, it includes the majority of replicate 1 peaks, indicating a group of high-quality, reproducible peaks (Figure 4.2.A). Note that previous work in our lab showed that increasing ChIP efficiency, for example by increasing the expression of a protein, can lead to more peaks being called as well as higher peak heights (Louphrasitthiphol et al. 2020). The increased ChIP efficiency in different replicates may therefore account for the imperfect overlap. Overall just over 2600 peaks were common between both replicates irrespective of which clone was used. Reproducibility is also reflected in peak score correlation — overlapping peaks have a similar peak score between replicates for both clones (Figure 4.2.B).

Finally, to test if there are artefact peaks, probably arising from accessible regions of the DNA, I generated Tag density plots using seqMINER. The plots are centred around peak summits. The four ChIP samples share the same tag

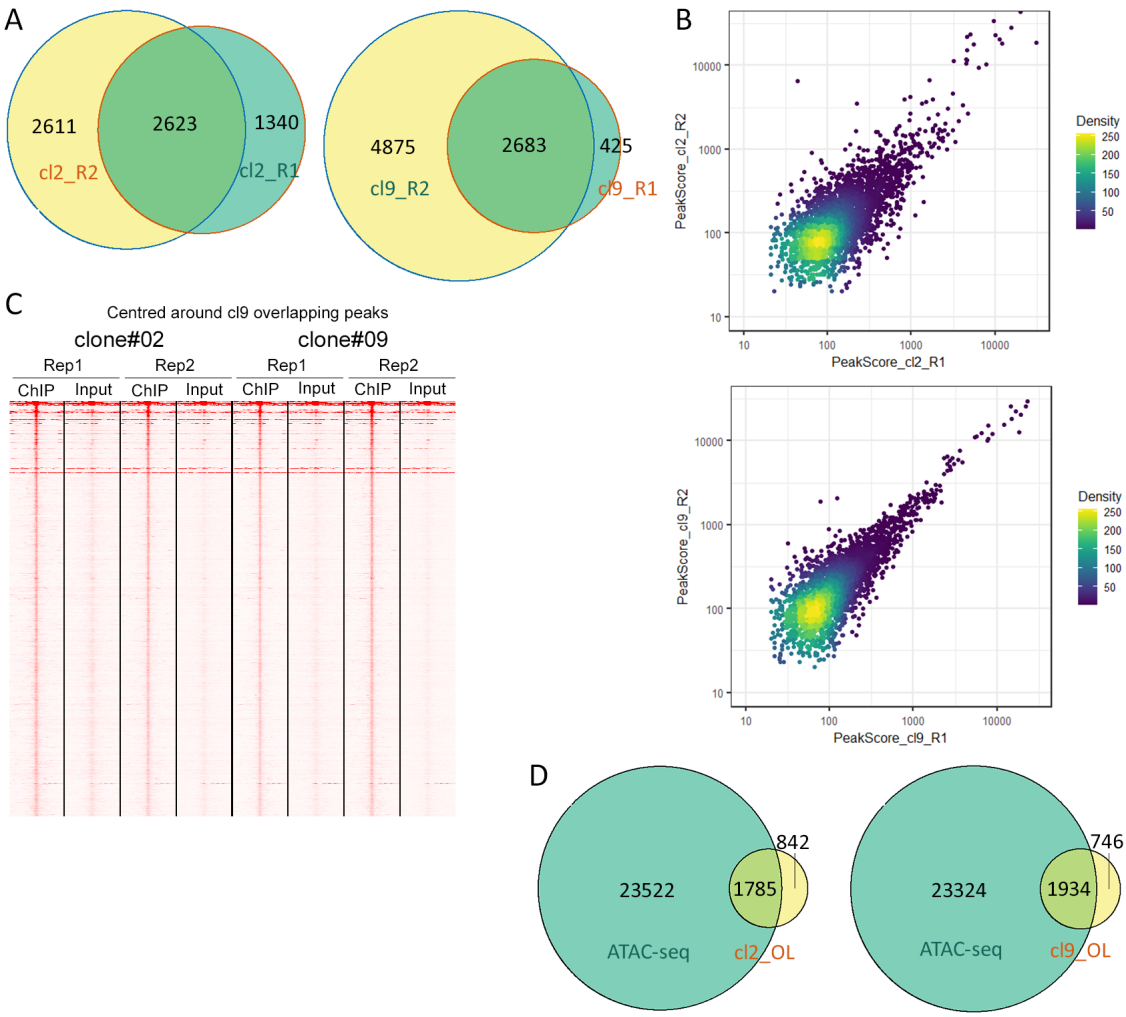



**Figure 4.2: Tbx2 ChIP-seq quality control** (A) Venn diagram and (B) Peak score correlation of overlapping peaks between biological replicates of B16-Tbx2-3xHA-NeoR clones. (C) Read density plot of ChIP and Input tags from all four samples, centred by cl#09's overlap peaks. (D) Venn diagram of overlapping peaks between Tbx2 peaks that are common between biological replicates of each clone and ATAC-seq peaks in B16 cells. (A) and (D) were generated in R package ChIPpeakAnno, (B) in R packaged ggplot2, (C) in Java seqMINER.

distribution pattern (Figure 4.2.C). In addition, the tag density contrast between ChIP and input samples is clear, especially among high peaks, indicating most peaks are genuine and not a result of any accessibility difference. Notably, there are dense horizontal lines, both in the input and ChIP samples, indicating open chromatin regions. Analysis of published ATAC-seq data generated in B16 cells (GSE83850, from Benci et al. 2016) shows that most Tbx2 peaks overlap with ATAC-seq peaks

(Figure 4.2.D), confirming the observation of open chromatin region in tag density.

**ChIP-seq analysis**

*Binding site distribution* Many transcription factors can act as either activators or repressors of transcription, the location of binding site relative to a target gene might be a key determinant of activity, in addition to differential cofactor interaction. Analysis of peak distribution using ChIPseeker showed that accumulatively 37.93% of peaks locate in promoter region — less than 3 kb from transcription start site (TSS), 28.24% locate in intron, while 25.85% are in distal intergenic region (Figure 4.3.A), and a closer look at the promoter part shows that read counts peaks at TSS (Figure 4.3.B).

*Known TBX2 binding sites* As mentioned in the introduction to this chapter, several TBX2 binding sites have been reported in other tumour studies, including the *Ndrg1* promoter in breast cancer cells (Redmond et al. 2010), and the *Pten* promoter in rhabdomyosarcoma cells (Zhu et al. 2015). In addition, though E-cadherin was reported, by our groups and others, to be mainly regulated by TBX3, there was one study showing a TBX2 binding site in the *Cdh1* promoter region in mammary epithelial cells (B. Wang et al. 2012). My results confirmed binding sites very close to the reported ones — two peaks proximal to the *Pten* transcription start site (TSS), one in the promoter, the other spanning 5' UTR (Figure 4.3.D); three peaks in the *Ndrg1* gene body, one in the promoter adjacent to TSS, and two in the first intron (Figure 4.3.E); in contrast to the reported binding sites in the *Cdh1* promoter, my ChIP-seq data called a high peak in an intron (Figure 4.3.F).

*Motif Enrichment* T-box transcription factors have defined binding elements, nevertheless it is important to test if Tbx2 binds the consensus motif in melanoma cells. As expected, HOMER *de novo* motif finding identified the TBE core sequence, GGTGTGAA, is enriched in 24% of peaks with a p-value of $1e^{-77}$ (Figure 4.3.C). This sequence was enriched in previous ChIP-seq studies of TBX2 (MA0688.1/Jaspar), as

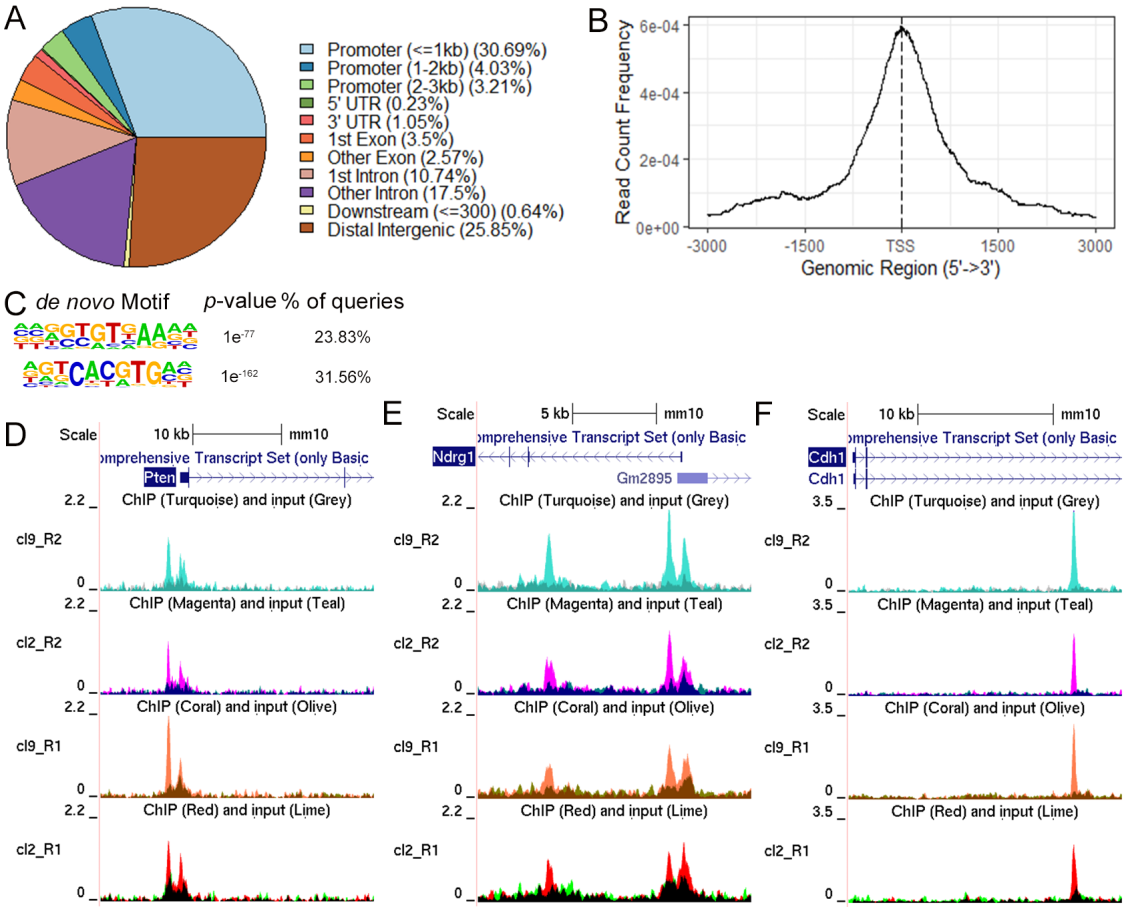



**Figure 4.3: Tbx2 ChIP-seq Analysis** (A) Genomic distribution of Tbx2 ChIP-seq peak summits (B) Peak read count in transcription start site (TSS) +/- 3 kb region (C) HOMER *de novo* motif identification results, shown are the motif sequence and respective p-value and enrichment percentage. (D–F) UCSC genome browser images showing representative peaks at (D) *Pten*, (E) *Ndrg1* and (F) *Cdh1* loci. (A) and (B) are generated in R package ChIPseeker, (C) in HOMER, using cl#09's overlapping peaks. In (D–F), ChIP sample tracks are coloured as red, magenta, coral and turquoise, Input coloured as green, teal, olive and grey respectively

well as other T-box family members: Eomes(Tbr2) (MA0800.1/Jaspar, H9-Eomes-ChIP-Seq(GSE26097)/Homer), TBR1 (MA0802.1/Jaspar), TBX20 (MA0689.1/Jaspar), TBX21(Tbet) (MA0690.1/Jaspar, CD8-Tbet-ChIP-Seq(GSE33802)/Homer), MGA (MA0801.1/Jaspar), TBX4 (MA0806.1/Jaspar). To our surprise, however, the top enriched motif has a core sequence TCACGTG (Figure 4.3.C), which is the recognition site for the MiT/TFE Family of Transcription Factors including MITF (Louphrasitthiphol et al. 2020), TFEB (MA0692.1/Jaspar), TFEC (MA0871.1/Jaspar) and TFE3 (MA0831.1/Jaspar), and other E-box binding factors

like Usf1 (GM12878-Usf1-ChIP-Seq(GSE32465)/Homer), Usf2 (C2C12-Usf2-ChIP-Seq(GSE36030)/Homer) and Arnt1(Bmal1) (MA0603.1/Jaspar, Liver-Bmal1-ChIP-Seq(GSE39860)/Homer). This might unveil a previously unknown binding motif for Tbx2, or a cooperative relationship between Tbx2 and E-box binding factors.

If Tbx2 cooperates with basic helix-loop-helix (bHLH) transcription factors, which bind the E-box sequence, one possibility is that they bind to adjacent TBEs and E-box motifs in the genome. To test this hypothesis, I used the COPS (Co-Occurrence Pattern Search) tool to identify co-occurrence and spatial arrangement of enriched motifs. COPS identifies co-occurring motifs in peak sequences, detects if they are overlapping, and compares that to background sequence; COPS also determines how far apart are the co-occurring motifs, scanning a 10-bp window (Ha et al. 2012). Statistical significance for motif co-occurrence is calculated as the *Z-score*, and a Z-score of 2.0 is often used as a cut-off. Highly overlapping motifs might result from similarity or shared sequences between motifs and thus cause ambiguous results. Hence, a high Z-score and a low binding site overlap are good indicators of true positive results.

| **Pattern** | **Log likelihood score** | **Z-score** | **Overlap %** |
|---|---|---|---|
| TBE:E-box | 0.112480267 | 4.4695649 | 0.12704918 |

**Table 4.1:** Co-occurring motifs in Tbx2 peak sequences

| **Distance between motifs** | 19~29 bp | -1~9 bp |
|---|---|---|
| **p-value** | 0.027006069 | 1.98E-06 |

**Table 4.2:** Distance between TBE and E-box

When TBE was used as the anchor sequence, the E-box was the only motif identified as co-occurring among the top 10 enriched motifs, with a Z-score of 4.5 and overlap percentage of 0.12 (Table 4.1). As for the distance between co-occurring TBEs and E-boxes, it is most likely to be in the -1~9 bp window (Table 4.2).

Our group has generated DNA-binding site profiles of bHLH transcription factors

MITF, USF1 and USF2 in human melanoma cells. We observed that USF1/2 and Tbx2 bind to the same region in certain genes, such as the promoter regions containing TSS of *MET*, *TELO2*, *E2F1*, *CCNE2*, *PTEN* and *NDRG1*, as well as intron regions of *CDK2* and *CDH1* (Figure 4.4.A–H). In addition, though a genome-wide comparison between USF1/2 ChIP-seq data and my Tbx2 ChIP-seq data, which were generated in mouse cells, is not optimal, the liftover of USF1/2 binding sites to mouse genome assembly enabled us to compare the conserved sequences. The results are shown in Figure 4.4.F, where 22.71% and 20.85% of Tbx2 binding sites overlap with those of USF1 and USF2 respectively.

Taken together, these results suggest that Tbx2 might bind to an E-box via cooperation with bHLH transcription factors USF1 and USF2. This hypothesis is followed up later by testing protein-protein interaction in section 6.3.

## 4.3 siTbx2 gene expression profile - RNA-seq

### 4.3.1 siTbx2 RNA-seq experiment

#### RNA-seq sample preparation

To determine whether the binding detected in the ChIP-seq experiment translates into transcriptional regulation, I silenced Tbx2 in B16 cells and performed 3' mRNA sequencing.

For the RNA-seq experiment, I generated triplicate samples from B16, B16-Tbx2-3xHA-NeoR cl#02 and cl#09. Silencing experiments were performed using two different siRNAs against Tbx2, and a negative control siRNA. WB results showed the two siTbx2 efficiently knocked-down Tbx2 in all experiments (Figure 4.5.A).

Experimental details can be found in section 2.5.2 on page 58. Briefly, cells were cultured in RPMI growth media without antibiotic and transfected using Lipofectamine RNAiMAX diluted in Opti-MEM. After 24 h, the transfection mix

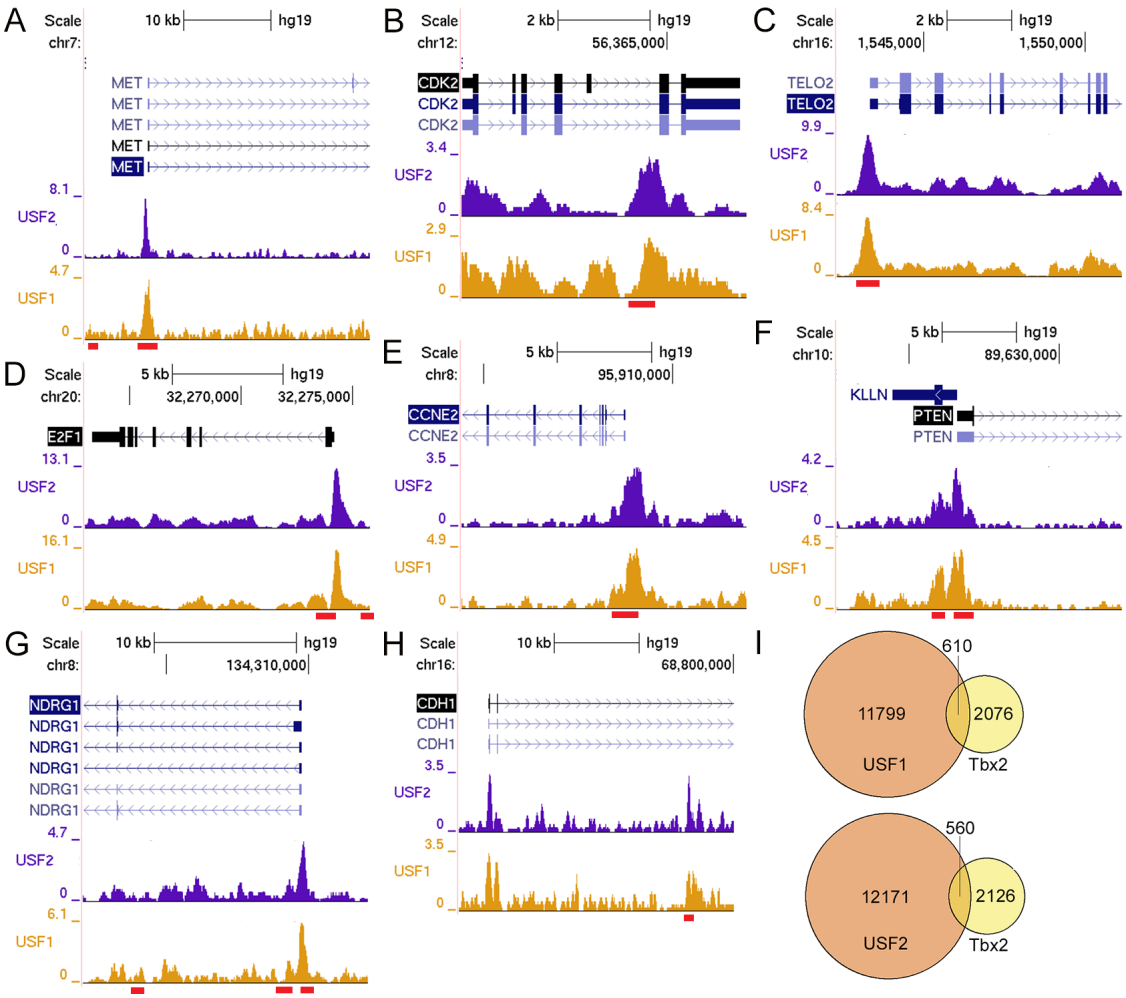



**Figure 4.4: Comparison of Tbx2 and USF1/2 binding sites** (A–H) UCSC hg19 genome browser images showing total tags of USF1/2 ChIP-seq at (A) *MET*, (B) *CDK2*, (C) *TELO2*, (D) *E2F1*, (E) *CCNE2*, (F) *PTEN*, (G) *NDRG1* and (H) *CDH1* loci; red boxes highlight the relative location of Tbx2 ChIP peaks. (I) Venn diagram of overlapping peaks between Tbx2 peaks that are common between biological replicates of clone#09 and USF1 or USF2 peaks. USF1/2 ChIP-seq experiment was conducted in 501mel cells, and the binding sites were liftovered to mm10 using UCSC tools; the Venn diagram was generated in the R package ChIPpeakAnno. The ChIP-seq data of USF1 and USF2 were provided by Dr Pakavarin Louphrasitthiphol in our group

was replaced by normal growth media, and 48 h after transfection, cells were lysed for RNA extraction. 3 µg of purified RNA was then sent to the WTCHG for library preparation and sequencing.

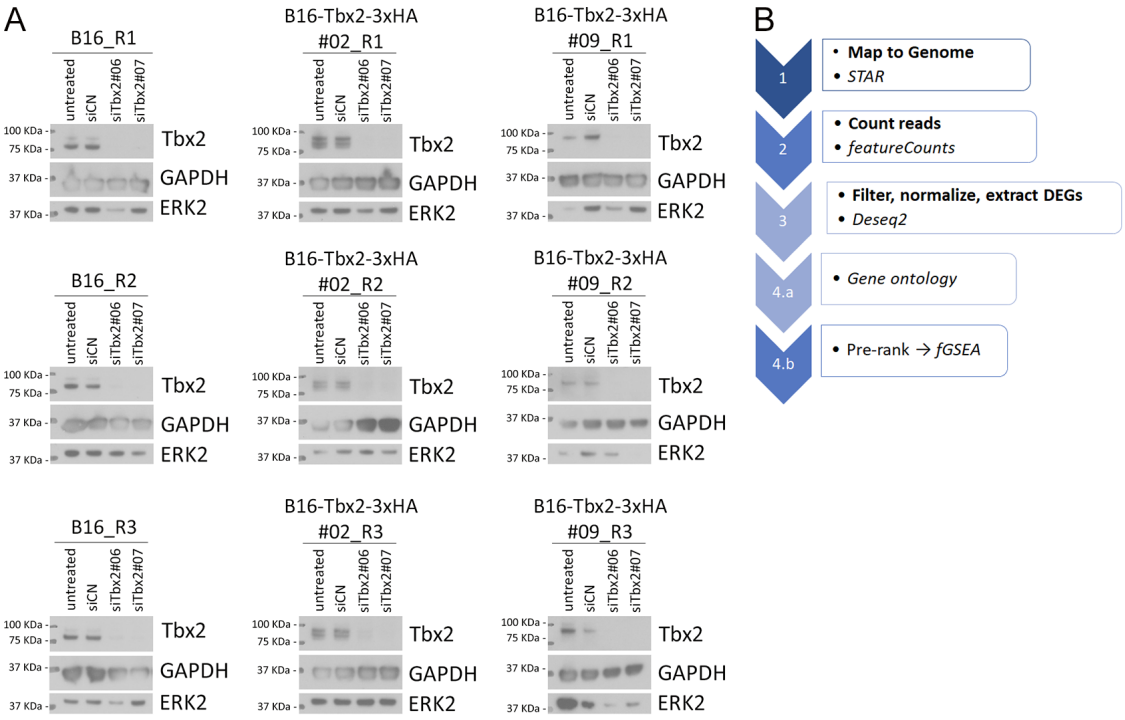



**Figure 4.5: siTbx2 experiment and RNA-seq analysis pipeline** (A) WB results of B16-Tbx2-3xHA-NeoR clones and parental B16 cells transfected with either negative control siRNA or siTbx2. GAPDH and ERK2 are used as loading controls (B) RNA-seq analysis steps and the softwares used.

### RNA-seq analysis process

The RNA-seq analysis pipeline is outlined in Figure 4.5.B, and individual steps described in section 2.6.1 on page 61. Briefly, raw reads were mapped to mm10 reference genome using STAR, then differential expressed genes were extracted using R package Deseq2.

## 4.3.2 siTbx2 RNA-seq results

### Basic analysis and DEGs

I first checked the quality of the RNA-seq samples and compared the two siRNA against Tbx2, namely siTbx2 #06 and #07. A heatmap and PCA plot were generated from the raw reads of samples. The heatmap shows that Tbx2 silencing led to a similar pattern of up- and down-regulation of genes in biological triplicates

across the two tagged clones and the B16 parental cells, however, while siTbx2 #06 appears to have changed expression of many genes, the pattern of siTbx2 #07 resembles to some extent that of the untransfected and siCN samples (Figure 4.6.A). This was confirmed in PCA — siTbx2 #07 clusters closer to the untransfected and siCN samples than siTbx2 #06 (Figure 4.6.B).

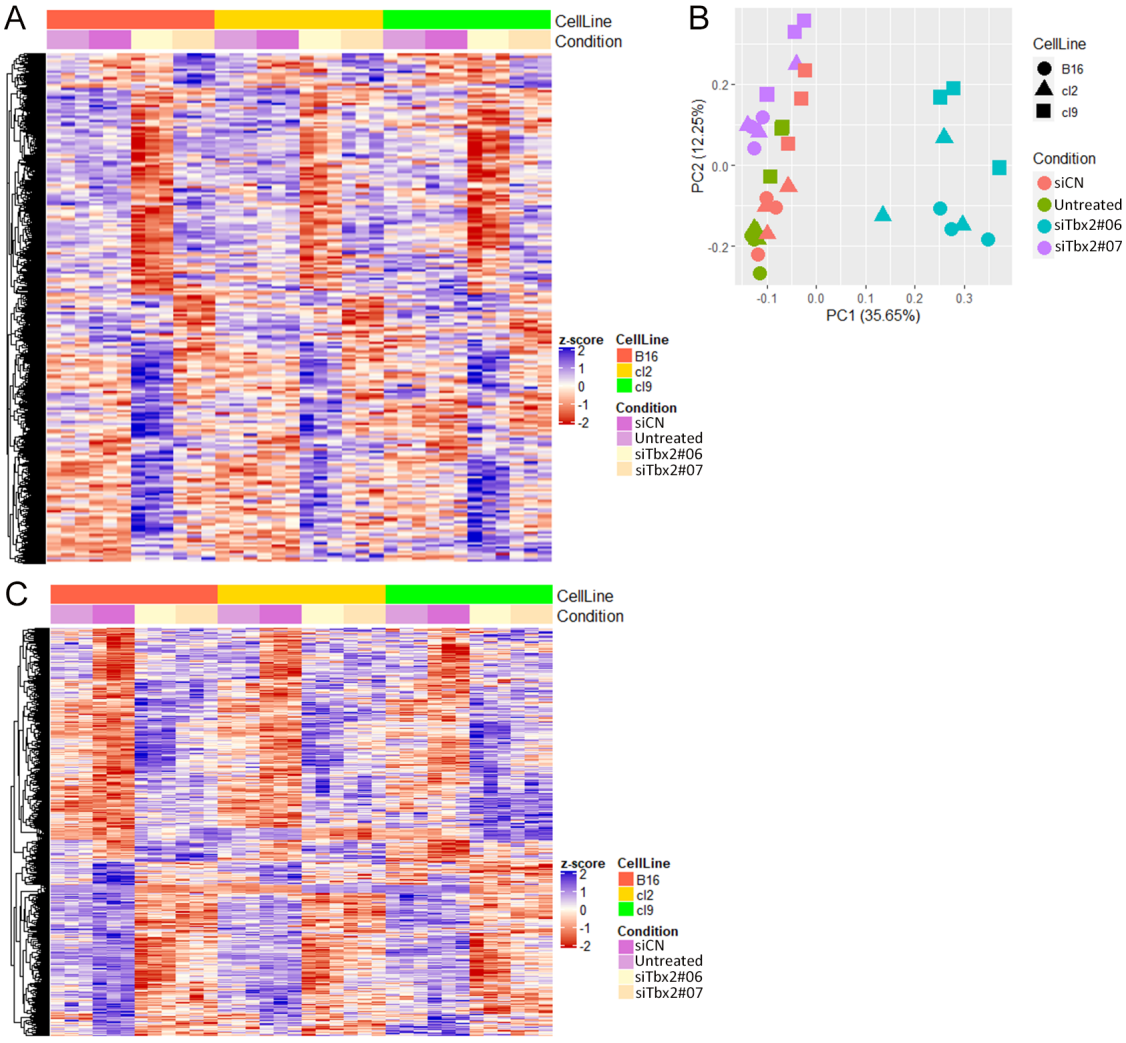



**Figure 4.6: Tbx2 RNA-seq Quality Control** (A) Heatmap of triplicates of B16-Tbx2-3xHA-NeoR clones and parental B16 cells that were transfected with negative control scramble siRNA or siTbx2. Genes shown here include statistically significant DEGs found in either siTbx2 #06 or #07 transfection. (B) Principal Component Analysis (PCA) plot of B16-Tbx2-3xHA-NeoR clones and parental B16 cells that were transfected with negative control scramble siRNA or siTbx2. Raw reads were used to generate the PCA plot in ggplot2. (C) Heatmap of triplicates of B16-Tbx2-3xHA-NeoR clones and parental B16 cells that were transfected with negative control scramble siRNA or siTbx2. Genes shown here include statistically significant DEGs that are common between siTbx2 #06 and #07 transfection.

After testing for differentially expressed genes (DEGs), it became apparent that siTbx2 #06 had more statistically significant DEGs compared to siTbx2 #07 (Table 4.3). For instance, in B16 cells siTbx2 #06 has 4,276 DEGs with an adjusted p-value less than 0.05, while #07 only has 1,708. In addition, if we examine the common significant DEGs between the two siTbx2 used, we can see that while siTbx2 #06 and #07 have similar z-score for some DEGs, siTbx2 #06 led to a greater change in gene expression in the rest (Figure 4.6.C). Taken together, these indicate siTbx2 #06 resulted in a more efficient knockdown of Tbx2.

| **Cell line** | **B16** | | **cl#02** | | **cl#09** | |
|---|---|---|---|---|---|---|
| siTbx2 | #06 | #07 | #06 | #07 | #06 | #07 |
| DEGs (p-adj $< 0.05$) | 4276 | 1708 | 3548 | 1633 | 3418 | 951 |

**Table 4.3:** Statistically significant DEGs from two siRNAs against Tbx2

In the following Gene Ontology (GO) and Gene set enrichment analysis (GSEA), a cutoff of p-value and a ranking determined by fold change and statistical significance are needed, therefore, analysis will be based on data from a single siRNA. Between the two siRNA, siTbx2 #06 proved to have superior silencing efficiency and was thus chosen for further analysis.

### Further analysis — GO, GSEA and KEGG

*Gene ontology (GO)* Genes with an adjusted p-value $< 0.05$ (from B16 cells, siTbx2 #06 vs siCN) were taken for ontology analysis, and the top 15 GO terms selected for visualisation. Enriched biological processes include multiple processes related to metabolism, biosynthesis and the cell cycle (Figure 4.7.A).

*Gene set enrichment analysis (GSEA)* DEGs were pre-ranked according to how drastic and how statistically significant the expression level change was, then taken for GSEA. Among the enriched cancer hallmark pathways, features favouring tumour growth like E2F targets, G2M checkpoint and mitotic spindle are down-regulated, while the p53 signalling pathway which arrests cell proliferation has

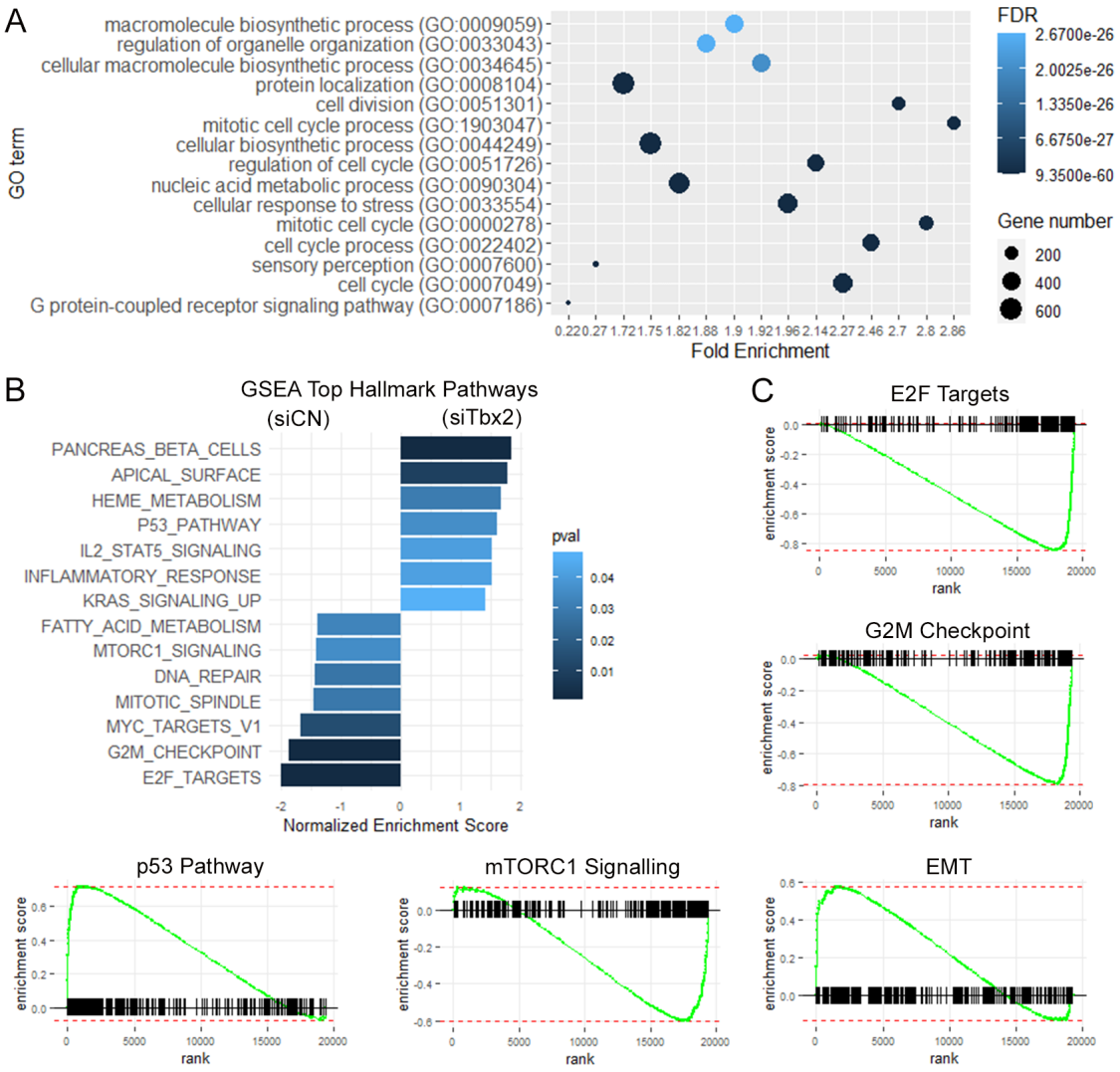



**Figure 4.7: siTbx2 RNA-seq Analysis — GO and GSEA** (A) Top 15 Gene Ontology (GO) Biological Processes (BP) as determined by FDR. Genes used for GO analysis were statistically significant DEGs (siTbx2 #06 vs siCN) from B16 cells. (B) Overview of Gene set enrichment analysis (GSEA) Hallmarks of Cancer top 14 pathways as determined by p-value. All genes from B16 cells were pre-ranked by fold change and p-value then analysed for GSEA. (C) Individual GSEA pathways shown in (B)

a positive enrichment score (Figure 4.7.B-C). The down-regulation of cell cycle related pathways is in line with previous studies showing Tbx2 contributes to senescence bypass and cell proliferation, and the changes of mTORC1 signalling could be mediated by regulating Pten (Zhu et al. 2015). Our results also revealed that Tbx2 silencing also altered fatty acid metabolism (Figure 4.7.B), shedding light on a previously unknown role of Tbx2 in maintaining melanoma growth.

Interestingly, the EMT pathway is also up-regulated (Figure 4.7.B-C). Though previous reports showed that overexpression of TBX3 will promote EMT in several types of cancer, the role of Tbx2 in melanoma invasion remains elusive, and our RNA-seq data suggest an opposite role for Tbx2, though functional assays are needed to confirm the observation.

*KEGG pathways analysis* Genes with an adjusted p-value $< 0.05$ (from B16 cells, siTbx2 #06 vs siCN) were taken for KEGG analysis, and the top pathways are listed in Figure 4.8.A. To have an intuitive understanding of whether genes in a certain pathway are up-regulated or down-regulated, I visualised individual enriched KEGG pathways and added the fold change of significant DEGs within a pathway. For example we can see that expression of the tumour suppressors p21 (encoded by *CDKN1A*) and Pten went up, while Cdk2, cyclin E and cyclin A decreased together with E2Fs, suggesting that silencing Tbx2 has a negative impact on the cell cycle (Figure 4.8.B) in agreement with loss of TBX2 promoting senescence. In addition, down-regulated ORC might block DNA replication licensing, thus preventing cells from entering S phase.

## 4.4 Identification of novel Tbx2 targets

The occupancy of a transcription factor at a certain gene, as measured by conventional ChIP experiments, does not necessarily correlate with transcriptional regulation function (Lickwar et al. 2012). On the other hand, the differentially expressed genes upon transcription factor perturbation could be either a direct target of the transcription factor or a downstream factor of a signalling cascade. In order to identify novel direct Tbx2 targets and subsequently expand our knowledge on its regulatory role in melanoma progression, it was therefore important to conduct integrative analysis of the two high-throughput sequencing datasets I obtained.

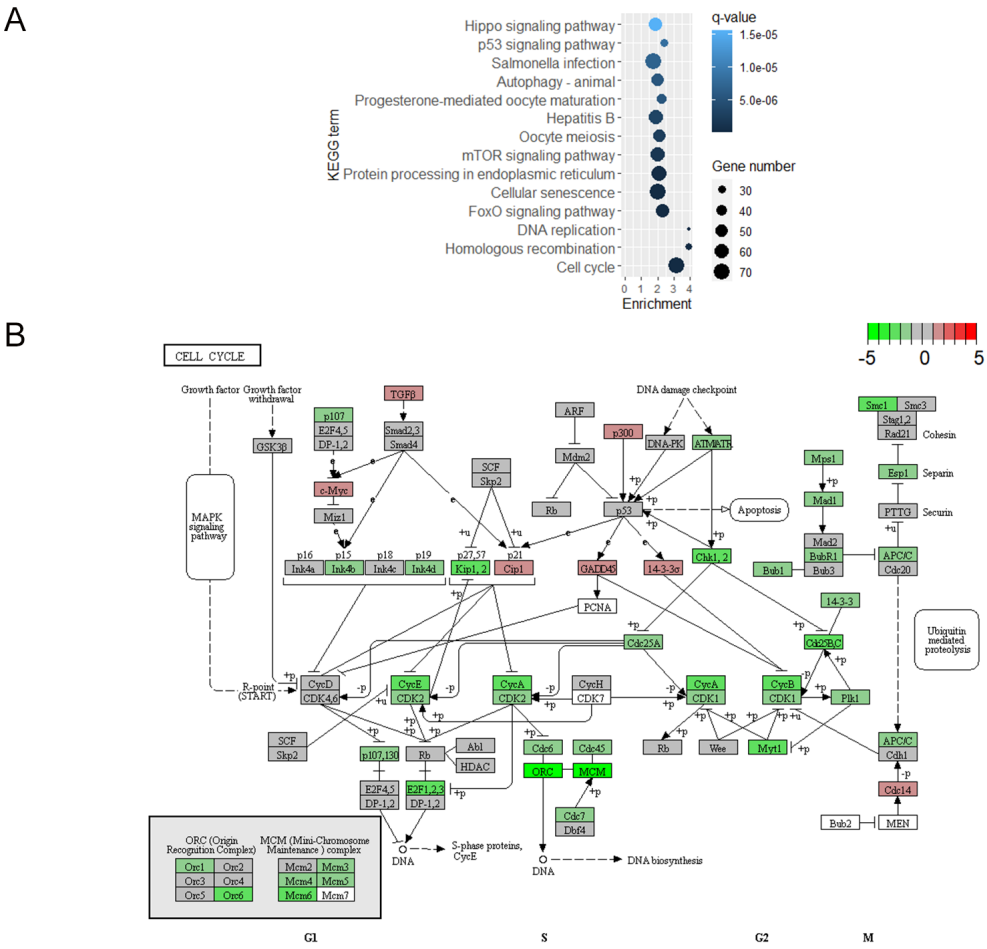



**Figure 4.8: siTbx2 RNA-seq Analysis — KEGG** (A) Overview of the top 14 KEGG pathways as determined by q-value (B) Example of siTbx2 #06 DEGs in the KEGG pathway 'cell cycle'. Colour of genes shows the degree of change in expression level, with green for decrease and red for increase.

### 4.4.1 Integration of DNA binding and gene expression profile

After comparing existing bioinformatics tools, I chose the Cistrome GO web server (http://go.cistrome.org) because it takes into account both binding site location and differential expression data, and does not rely on arbitrary cutoffs for target assignment but rather employs a gene ranking based on both datasets.

### Cistrome GO functional analysis method

In Cistrome GO ensemble mode, gene product score is derived from two ranks, $R^{RP}$ and $R^{DE}$, where the DE (differential expression) score is calculated from the p-value reflecting the significance of differential gene expression upon transcription factor perturbation, and the RP (regulatory potential) score represents how likely the transcription factor is a direct regulator of this gene based on ChIP peaks. The RP score for a certain gene is calculated as $S_g = \sum_{i=1}^{k} 2^{-\frac{d_i}{d_0}}$, where $d_0$ is the decay distance, $d_i$ is the distance between the summit of the $i$th peak and the TSS of the gene. All $k$ peaks within the distance of $15d_0$ to the TSS are included in the calculation (Figure 4.9.A). Therefore, if a gene has multiple ChIP peaks surrounding its TSS, it is likely to be a direct target of the transcription factor.

In the case of the Tbx2 ChIP-seq dataset, Cistrome GO analysed the distribution of the most significant peaks then set the decay distance $d_0$ to 1 kb, which is the default for promoter-dominant transcription factors, and sequentially derived $R^{RP}$ from it. As for the differential gene expression profile, I uploaded significant DEGs data from siTbx2 #06 experiment, the adjusted p-value from which will be used to calculate the DE score. Genes were ranked by $R^{RP}*R^{DE}$ to determine the potential of being directly regulated by Tbx2.

### Cistrome GO results

Before starting the analysis, it was necessary to confirm there was a need for integrating the datasets. For example, are the statistically significant DEGs more likely to be direct targets? Or is fold change the indicator? To test this, I plotted the RP score, representing ChIP-seq information, and the DEGs dataset to determine if any of them is a dominating parameter, thus making the integration redundant. Firstly, Figure 4.9.B shows no clear correlation between the RP score and the DEG adjusted p-value, which was used to generate the DE score for ranking. Then I

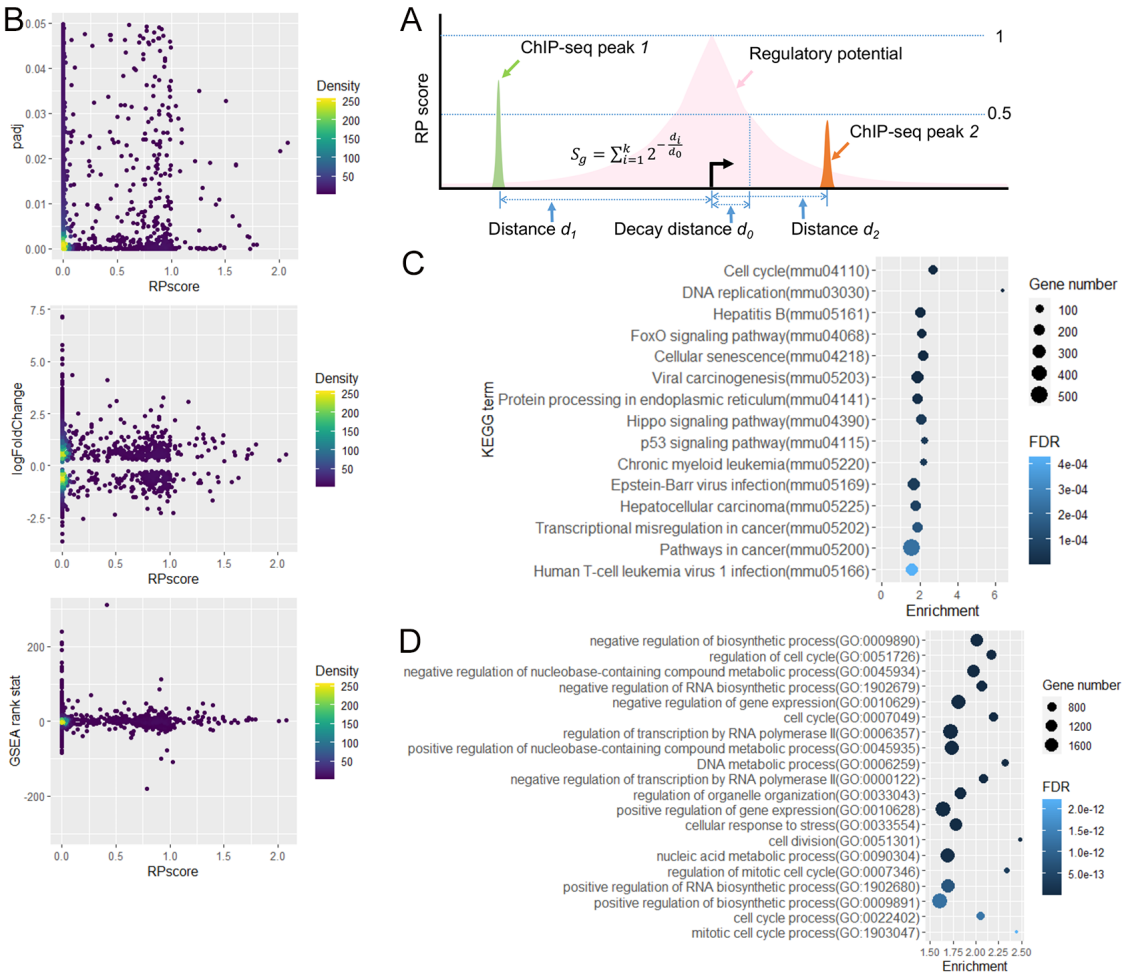



**Figure 4.9: Integrative analysis of ChIP-seq and RNA-seq** (A) Schematic illustration of regulatory potential (RP) score calculation method in Cistrome GO analysis for ChIP-seq peaks (B) Dot plots showing correlation between RP score, from ChIP-seq data, and p-value, log fold change and GSEA ranking stats, from RNA-seq data. (C) KEGG pathways and (D) GO biological processes enrichment for genes ranked by integrative analysis of ChIP-seq and RNA-seq data in Cistrome GO web server. Plots in (B), (C) and (D) were generated in R package ggplot2.

tested for the log fold change and GSEA rank statistic, which was calculated from both adjusted p-value and fold change, and found no correlation in either of them (Figure 4.9.B). Therefore, a separate rank integrating ChIP-seq and RNA-seq data is needed and will facilitate the identification of novel target genes.

Cistrome GO top ranked genes, which are most likely to be direct targets of Tbx2, were analysed for enrichment in GO terms and KEGG pathways. Figure 4.9.C shows that significantly enriched KEGG pathways include cell cycle, DNA replication, the

FoxO signalling pathway, cellular senescence, the Hippo signalling pathway and the p53 signalling pathways, resembling the analysis results of the RNA-seq. Similarity between integrated analysis and RNA-seq analysis is also observed in the GO biological processes results — cell cycle regulation and metabolic processes-related terms were enriched in both (Figure 4.9.D).

### 4.4.2 Direct target genes of Tbx2

After obtaining the rank, I went through the list and referred to the original data for additional information that was not used in generating the rank but holds importance, including peak height and statistical significance in the ChIP-seq dataset, as well as the log fold change from siTbx2 RNA-seq dataset. As the Cistrome GO derived gene rank based only on peak location and number from Tbx2 binding profile and p-value from the siTbx2 #06 gene expression profile, it is important to look at other information of the top ranked genes when choosing potential direct targets to follow up.

Among the Cistrome GO top ranked genes, there are significant DEGs with high-quality Tbx2 peaks. For example, the *Met* gene has three peaks in the promoter region, two of which are proximal to the TSS (Figure 4.10.A); *Cdk2* has a peak in an intron (Figure 4.10.B); *Telo2* has a peak in the 5' UTR spanning the TSS (Figure 4.10.C); *Pten* also has two peaks surrounding the TSS (Figure 4.3.D); *Ndrg1* has one peak spanning the TSS and two in the intron before first exon (Figure 4.3.E); *E2f1* has one peak spanning the 5' UTR and first exon and another one in the promoter (Figure 4.10.D); *Ccne2* has two peaks surrounding the TSS (Figure 4.10.E).

*Met* is a proto-oncogene encoding the tyrosine kinase receptor (c-Met) for hepatocyte growth factor (HGF). c-Met is expressed in both melanocytes and melanoma cells, and the HGF/c-Met pathways contribute to melanoma proliferation, invasion and metastasis (Czyz 2018). *Cdk2* encodes cyclin-dependent kinase 2,

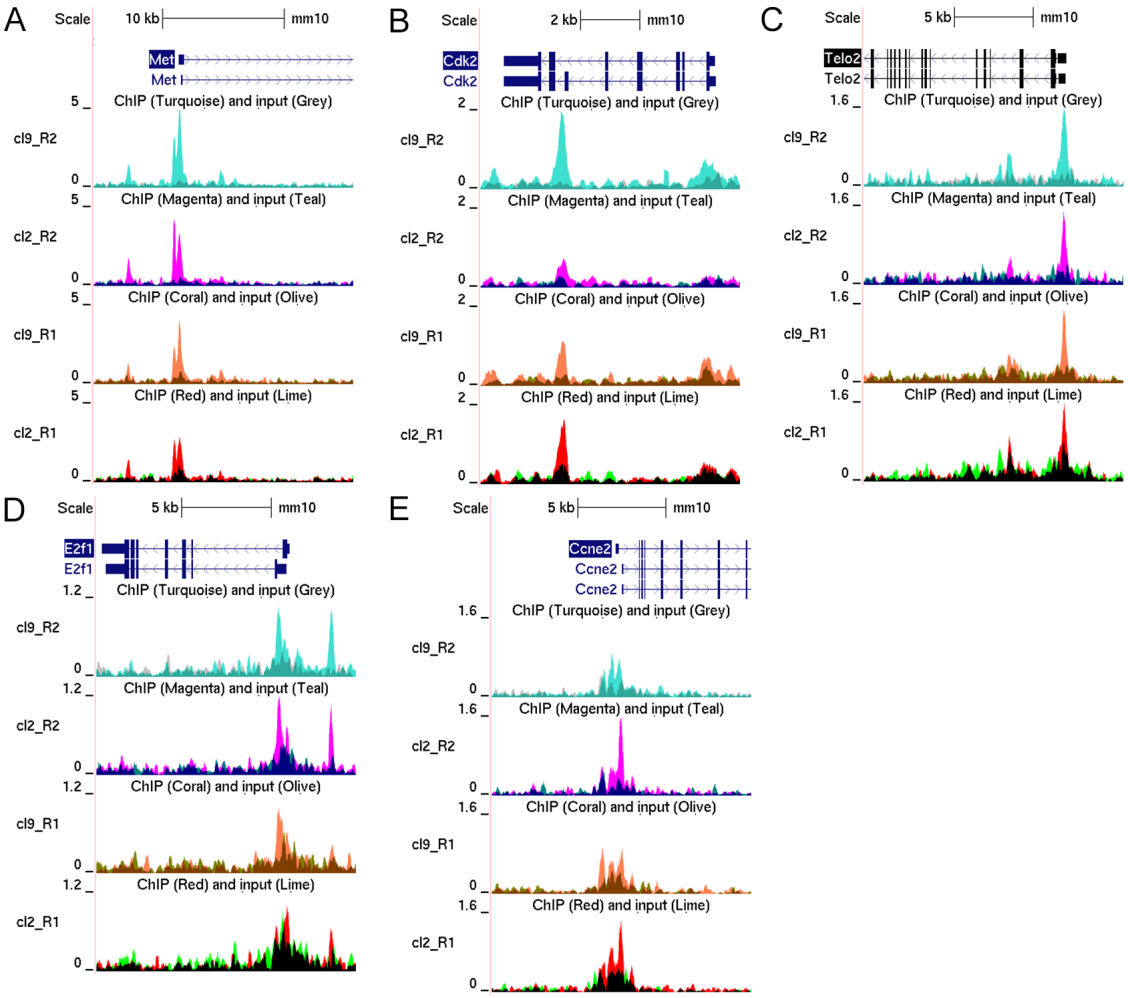
which is important for the transition of cells into the S- and M-phases of the cell cycle (Chohan et al. 2014). *Telo2* is the mammalian orthologue of *Tel2*, an *S. cerevisiae* gene involved in maintenance of telomere length, and Telo2 overexpression correlates with adverse outcome in high-grade gliomas (S.-W. Feng et al. 2015). E2f1 is a transcription factor that is important in cell cycle progression (Denechaud et al. 2017). *Ccne2* encodes cyclin E2, which binds to and activates Cdk2 (Siu et al. 2014). Among them, as shown in Figure 4.10.F, *Met*, *Cdk2* and *Telo2* were down-regulated upon Tbx2 silencing using both siRNAs; Pten was significantly up-regulated in both siTbx2 experiments, agreeing with a previous study (Zhu et al. 2015). Ndrg1 was decreased — though this only has statistical significance in siTbx2 #07 experiment — contradicting a previous report showing Tbx2 represses Ndrg1 in breast cancer (Redmond et al. 2010).

These potential direct target genes provide new insights into the regulatory mechanism of how Tbx2 promotes melanoma progression. Prevailing studies portray Tbx2 as a transcriptional repressor inhibiting tumour suppressor expression, however, my preliminary results add to our understanding of Tbx2 function by showing that Tbx2 may directly activate the expression of oncogenes like *Met*, positive regulators of cell cycle like *E2f1* and *Ccne2*, as well as *Telo2* that regulates telomere length.

### 4.4.3 Functional assays

To validate the phenotypes of Tbx2 silencing suggested by the RNA-seq data and the potential direct target genes identified in ChIP-seq data, I used siRNA to knockdown Tbx2, and then assessed the impact on melanoma cell proliferation, and tested the expression levels of relevant genes that are potential direct targets of Tbx2.

As the gene set enrichment analysis (GSEA) of RNA-seq data suggested down-regulation of E2F targets and G2M checkpoint-related genes as well as up-regulation of p53 pathway following Tbx2 knockdown, I assessed if this can lead to cell cycle




F

| | | siTbx2 #06_vs_siCN | | | siTbx2 #07_vs_siCN | | |
|---|---|---|---|---|---|---|---|
| GeneName | baseMean | log2FoldChange | pvalue | padj | log2FoldChange | pvalue | padj |
| Met | 3014.0381 | -1.22020299 | 8.6E-18 | 6.64E-16 | -0.554980449 | 9.16E-05 | 0.00157 |
| Ccne2 | 37.122708 | -1.836029747 | 7.3E-07 | 6.99E-06 | 0.701349254 | 0.020972 | 0.096523 |
| E2f1 | 363.15029 | -1.710705919 | 5.6E-17 | 3.86E-15 | 0.275058132 | 0.162498 | 0.371652 |
| Cdk2 | 1461.9515 | -1.138341098 | 4.2E-11 | 9.79E-10 | -0.650970695 | 0.000161 | 0.002491 |
| Telo2 | 314.31814 | -1.115280013 | 1.4E-16 | 9.62E-15 | -0.35051305 | 0.008094 | 0.049812 |
| Pten | 165.31722 | 1.188011262 | 2.6E-13 | 8.94E-12 | 1.544363746 | 2.78E-21 | 1.08E-18 |
| Ndrg1 | 285.8627 | -0.120895224 | 0.78014 | 0.859893 | -3.146877907 | 3.68E-12 | 4.94E-10 |

**Figure 4.10: Tbx2 potential direct targets** UCSC genome browser images showing representative peaks at (A) *Met*, (B) *Cdk2*, (C) *Telo2*, (D) *E2f1* and (E) *Ccne2* loci. (F) RNA-seq data of the selective genes generated in B16 cells in triplicates

arrest and proliferation reduction in B16 melanoma cells. Flow cytometry results in Figure 4.11.A showed that two different siRNAs against Tbx2 blocked S phase entry and thus halted the cell cycle in G1. As a result, Tbx2 knockdown drastically reduced cell growth in monolayer culture, as evidenced by crystal violet staining, which is proportional to the cell biomass that is attached to the plate and can be used to infer adherent cell number (Figure4.11.B).

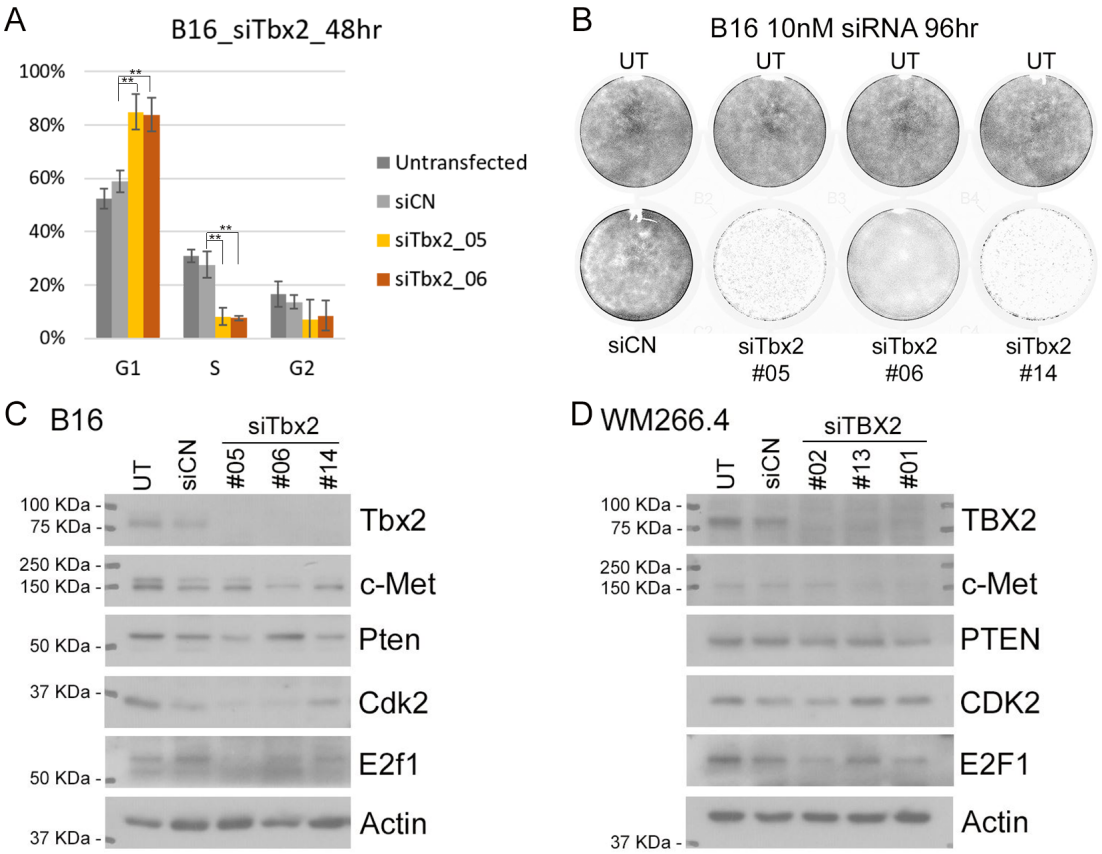



**Figure 4.11: Functional Assays of Tbx2 silencing** (A) Cell cycle analysis of B16 melanoma cells that were untransfected or transfected with 20 nM of siRNA for 48 h (n=3). DNA was stained with propidium iodide (PI) and the fluorescence intensity was measured using flow cytometry; cell cycle phases were determined using FlowJo software. Data are mean ± SD from three independent experiments (>8,000 cells/experiment); **p < 0.01 calculated by chi-square test. (B) Crystal violet staining of B16 melanoma cells that were untransfected or transfected with 10 nM of siRNA for 96 h (C-D) Western Blot of (C) B16 cells transfected with 20 nM of siRNA for 36 h and (D) WM266.4 human melanoma cells transfected with 20 nM of siRNA for 48 h

Next I sought to identify downstream effectors mediating the pro-proliferative function of Tbx2, with a special interest in potential direct target genes that were previously unknown. As shown in Figure 4.10.F, Tbx2 knockdown increased the mRNA level of Pten, and decreased that of Met, E2f1, Cdk2; in addition, Tbx2 directly binds to the promoter regions of *Met* and *E2f1*, and an intron region of *Cdk2* (Figure 4.10.A, B and D). Taken together, these results suggest Tbx2 might promote cell proliferation, at least partially, by directly activating the expression of Met, E2f1 and Cdk2. If they are indeed the key downstream factors underpinning

the observed phenotype, the expression change should be reflected in protein level as well. I therefore carried out western blot and found that all three siTbx2 all reduced E2f1 protein level in B16 cells (Figure 4.11.C), while in human melanoma cell line WM266.4, two out of three siTBX2 reduced E2F1 expression (Figure 4.11.D). For Cdk2, two out of three siRNAs against mouse Tbx2 decreased Cdk2 protein level (Figure 4.11.C), whereas there was no change in human melanoma cells (Figure 4.11.D). In addition, no consistent changes in Met or Pten were observed (Figure 4.11.C-D). Taken together, these results indicate E2f1 may be a crucial downstream effector of Tbx2 conveying the pro-proliferative signal, consistent with its role in melanoma cell cycle (Vivas-García et al. 2020).

## 4.5 Discussion

Given TBX2's important role in bypassing senescence and driving cell proliferation, it is imperative that we expand our knowledge about its downstream effectors and how TBX2 promotes melanoma progression. We therefore carried out a combination of Tbx2 ChIP-seq and RNA-seq following Tbx2 silencing, in an endeavour to identify direct target genes of Tbx2 in melanoma.

### 4.5.1 Tbx2 ChIP-seq discussion

A recent study published ChIP-seq data of TBX2 using the neuroblastoma cell line IMR-32 (Decaesteker et al. 2018). However, the data quality was compromised because of the anti-TBX2 antibody used, which from our experience is not good enough for ChIP-seq. As a result, in total they only called 557 significant peaks using a p-value of 0.05 as cutoff, and none of the peaks has a p-value under 0.01. In addition, the peak height is relatively low, and only one biological replicate was performed. Taken together, the TBX2 binding sites discovered in this study are not inclusive and might contain artefacts arising from undertaking only a single

experiment. Nevertheless, they also found an enriched E-box motif (CANNTG) apart from the T-element (AGGTGTGA), supporting the data I generated using HA-tagged endogenous TBX2.

The establishment of endogenously tagged cell line enabled me to generate high-quality Tbx2 ChIP-seq data — on the one hand, immunoprecipitating the HA tag circumvented the lack of ChIP-grade TBX2 antibody, on the other hand, targeting endogenous protein prevented artefacts arising from overexpression. As a result, my ChIP-seq data called 3,963 and 5,234 peaks in B16-Tbx2-3xHA cl#02 replicates, 3,108 and 7,558 peaks in cl#09 replicates using a p-value of 0.01 as cutoff (Figure 4.2.A). In addition, the correlation of peakscores between biological replicates is another indicator of good quality (Figure 4.2.B).

Nevertheless, my Tbx2 ChIP-seq data might still miss some parts of the Tbx2 binding profile. WB results show that after sonication and centrifugation, there is still considerable amount of Tbx2 protein in the cell pellet (Figure 4.1.A), which contains insoluble chromatin. If the chromatin-bound Tbx2 portion is excluded in the ChIP-seq experiment, which only uses the soluble fraction, we might lose the whole picture of its binding profile. TBX2 and TBX3 are reported to be transcriptionally repressive in most studies, and we have preliminary data, which I will elaborate on in Chapter 6, showing that Tbx2 interacts with key components of polycomb repressive complex 2 (PRC2). In addition, *in vitro* DNA binding assays show that Tbx2 can bind nucleosomal DNA (Demay et al. 2007). This is later independently confirmed using nucleosome consecutive affinity purification–systematic evolution of ligands by exponential enrichment (NCAP–SELEX) (F. Zhu et al. 2018). It is noteworthy that this study found both T (Brachyury) and TBX2 can bind both nucleosomal DNA gyres via dimerization. The fact that in my dataset most Tbx2 peaks overlap with ATAC-seq peaks (Figure 4.2.D) indicates the ChIP-seq experiment largely enriched for open chromatin region, therefore the data will not necessarily capture all Tbx2 binding sites and might leave out those responsible

for its repressive regulation of target genes.

### 4.5.2 siTbx2 RNA-seq discussion

While it is normal practice using multiple siRNAs for sequencing experiment in order to avoid artefacts arising from off-target silencing, different siRNAs might vary in knock-down efficiency and half-life thus resulting in different DEGs. In my experiment, silencing Tbx2 in B16 cells using siTbx2 #06 resulted in 4,276 significant differentially expressed genes (DEGs), but #07 only had 1,708 (Table 4.3). Among the set of common DEGs between them, siTbx2 #06 caused a greater change in mRNA level. These might be explained by the difference in silencing efficiency, as the average fold change of Tbx2 in siTbx2 #06 experiment is -2.9, whereas it is only -2.0 for siTbx2 #07. Alternatively, as I removed the transfection mix after 24 h to alleviate the toxicity of Lipofectamine, it could also be that siTbx2 #07 is less stable so that the transfected siRNA is quickly degraded and a small amount of Tbx2 protein, though not obvious in the WB results, remained and thus counteracted some changes caused by Tbx2 silencing.

To validate the phenotypes of Tbx2 silencing and relevant downstream target genes, I tested multiple siRNAs against Tbx2 to rule out artefacts from off-target effects of a single siTbx2.

Discrepancy among different siTbx2 was also observed in the following functional assays. For instance, siTbx2 #07 did not inhibit melanoma cell growth (data not shown). Nevertheless, the effect of siTbx2 #06 was confirmed using two other siRNAs (siTbx2 #05 and #14), indicating the cell cycle arrest and reduction in cell proliferation are not artefacts from siTbx2 #06, though the silencing efficiency of siTbx2 #05 and #14 remains to be determined (Figure 4.11.A–B). Similarly, though siTbx2 #07 did not significantly change E2f1 expression (Figure 4.10.F), siTbx2 #05, #06 and #14 all decreased E2f1 protein level (Figure 4.11.C).

I also assessed if the results obtained in B16 mouse melanoma cells can be reproduced in human melanoma cells. I chose A375M, SKmel30 and WM266.4 cell lines as they highly express TBX2. Regrettably, the knockdown efficiency in A375M and SKmel30 was not significant (data not included), even in WM266.8 where a better silencing was achieved, residual TBX2 protein is still observed in WB (Figure 4.11.D). Therefore, further optimisation of TBX2 knockdown experimental condition for different human melanoma cell line is needed.

# 5

# Binding sites of TBX3

## 5.1 Introduction

Given the important functions of TBX3 in embryonic development, pluripotency and human disease including cancer and UMS, it is essential that we explore the genes regulated by TBX3 and understand how they mediate the TBX3's effect in different biological processes. In this section, I will summarise the known direct target genes of TBX3.

### 5.1.1 Known direct target genes of TBX3

**CDKN2A** Lingbeek et al. 2002 first showed that in MEFs TBX3 can directly repress the mouse $p19^{Arf}$ (*CDKN2A*), which inhibits the p53 repressor MDM2 E3 ubiquitin ligase thus facilitating p53 activation. Later Platonova et al. 2007 reported

in mammary epithelial cells, Tbx3 inhibits $p14^{ARF}$ promoter activity, irrespective of p53 and Mdm2 status. In addition, Tbx3 expression in hepatic progenitor cells down-regulated $p19^{Arf}$ levels (Suzuki et al. 2008). These data suggest that the CDKN2A locus, which encodes important cell cycle regulators, is regulated by TBX3 during various biological processes.

**CDKN1A** Tbx3 isoforms bind to the initiator of *CDKN1A*, as shown by EMSA in 501mel cells, and represses it promoter activity and the subsequent expression of $p21^{CIP1}$, the CDK2 inhibitor (Willem M H Hoogaars et al. 2008). In chondrosarcoma cells, TBX3 also binds to and repress the promoter of $p21^{CIP1}$ (*CDKN1A*) via a conserved consensus T-element at -121 bp of the promoter, as identified by ChIP-qPCR and luciferase assays (Tarryn Willmer, Hare, et al. 2016). By repressing $p21^{CIP1}$ expression, TBX3 can reduce CDK2-mediated repression of Rb1, an inhibitor of E2F transcription factors that activate S-phase genes.

### TBX3 target genes identified in cancer

**CDKN1C** $p57^{KIP2}$ is a Cip/Kip CKIs family member critical for cell cycle regulation. TBX3 directly binds to *CDKN1C*, the coding gene of $p57^{KIP2}$, and represses its promoter activity, as evidenced by ChIP-qPCR, DNA affinity binding assays and luciferase assays. (X. Li et al. 2018)

**UCA1** The long non-coding RNA (lncRNA) UCA1 stabilises p16 mRNA and induces senescence, and silencing of UCA1 diminishes oncogene-induced senescence (OIS). TBX3 coordinates with CAPERα to directly bind the *UCA1* promoter and repress its expression. Knockdown of either TBX3 or CAPERα changes the chromatin structures of *CDKN2A* and *UCA1* promoters, reflected by an increase in the euchromatic mark H3K4me3 and a decrease in the heterochromatic mark H3K9me3 and H3K27me3 (Pavan et al. 2014). Repression of UCA1 adds to the knowledge about how TBX3 contributes to senescence bypass.

**PTEN** Overexpression of TBX3 repressed the *PTEN* promoter activity in

HeLa and HEK cell lines, though no protein level change of PTEN was shown in the paper (Burgucu et al. 2012).

**CDH1** TBX3 reportedly promotes melanoma invasiveness by directly decreasing E-cadherin levels via binding to a variant T-element close to the TSS of its coding gene *CDH1* (Rodriguez et al. 2008). TBX3 also interacts with HDAC to repress E-cadherin in hepatocellular carcinoma (HCC) cells and consequently promotes HCC metastasis (Dong et al. 2018). In addition, silencing of TBX3 increased E-cadherin protein levels in human bladder cancer cells (H. F. Du et al. 2014). Moreover, in gastric cancer tissues, the expression of TBX3 and E-cadherin inversely correlate with each other, while in gastric cancer cells TBX3 silencing up-regulated E-cadherin and decreased N-cadherin and Vimentin (Miao et al. 2016).

**SLUG** Apart from E-cadherin, TBX3 also promotes the transition from ductal carcinoma *in situ* (DCIS) to invasive breast cancer by directly binding to TSS of the EMT-related gene *SLUG*, as shown by ChIP arrays, and up-regulating SLUG expression (Krstic, Kolendowski, et al. 2019). Contrary to most studies, this is an example showing TBX3 acting as a transcriptional activator, though the authors did not go further on the underlying mechanism.

**Nodal**, **Inhbb**, **Smad4** and **Alk4** Reactivation of the embryonic Nodal/Activin signalling pathway can drive the self-renewal of cancer stem cells. In mouse ESCs, TBX3 binds to the *Nodal*, *Inhbb* (subunit of Activin A), *Smad4* and *Alk4* genes as shown in ChIP-seq profiles, while in primary cancer cells, TBX3 overexpression increases Activin A mRNA level (Perkhofer et al. 2016).

### TBX3 target genes identified in development

**Cx40** Connexin 40 (*Cx40*) is a marker for atrial and ventricular chamber-type myocardium. Tbx3 represses the *Cx40* promoter activity, while L143P and Y149S mutations in the DBD abolished the repression (W. M. Hoogaars et al. 2004).

**Gata6** In mouse ESCs, Tbx3 directly activates the expression of Gata6, the

endoderm marker, by binding to the *Gata6* promoter as shown by ChIP-PCR and luciferase assay, while its DBD mutant Y149S fails to do so. (R. Lu et al. 2010)

**Dppa3** Dppa3 is a marker of naive mESCs that is down-regulated during embryoid body differentiation. ChIP-seq data show that TBX3 binds 1.7 kb upstream of the *Dppa3* TSS and represses its expression in mESCs (Waghray et al. 2015).

**JDP2** Jun dimerization protein 2 (JDP2) is an important regulator of osteoclastogenesis. Tbx3 is induced during osteoclastogenesis, and up-regulates JDP2 expression by directly binding to the *JDP2* promoter, as shown in EMSAs and ChIP-qPCR experiments (Yao et al. 2014).

**Wnt signalling effectors** Similar to Tbx2, Tbx3 also directly represses the Wnt antagonists Frzb and Shisa3 in the lung mesenchyme, by binding to two different TBEs, as evidenced by ChIP-PCR and microarray experiments (T. H. Lüdtke et al. 2016).

**Reprogramming factors** Han et al. 2010 showed that in ESCs, Tbx3 directly binds to ESC factors Oct4 and Lefty1, as well as reprogramming factors Klf2, Klf4, Klf5 and c-Myc, as shown by ChIP-seq data. Notably, in this study only one biological replicate was performed, with 1,343 peaks called.

### 5.1.2 Identification of novel TBX3 target genes

As mentioned in section 4.1 as well as the introduction to this chapter, to date few high-quality TBX3 ChIP-seq datasets have been published, mainly resulting from lack of ChIP-grade antibody against endogenous TBX3 protein. While our previous group member Johannes performed a ChIP-seq with melanoma cells overexpressing Flag-tagged TBX3, the potential target genes identified could not be validated. This brought our attention to Flag tag used in the ChIP-seq experiment. Irwin Davidson from IGBMC Strasbourg found that anti-Flag antibody can cause artefact binding to transcription initiation sites. To overcome this issue, I sought to generate

cell lines with HA-tagged TBX3.

## 5.2 ChIP-seq experiment

### 5.2.1 Inducible TBX3-3xHA cell line

#### Generation of inducible TBX3 cell line

As I did not manage to tag endogenous TBX3 using eFlut, in order to conduct the ChIP-seq experiment with available ChIP-grade HA antibody, I established inducible 501mel cell line expressing TBX3 (+2a) with a carboxyl-terminal 3xHA tag. To allow an inducible expression of TBX3 and to maximise integration efficiency, a tetracycline-controlled transcriptional activation (Tet-on) system was used in combination with piggyBAC transposase (Figure 5.1.A).

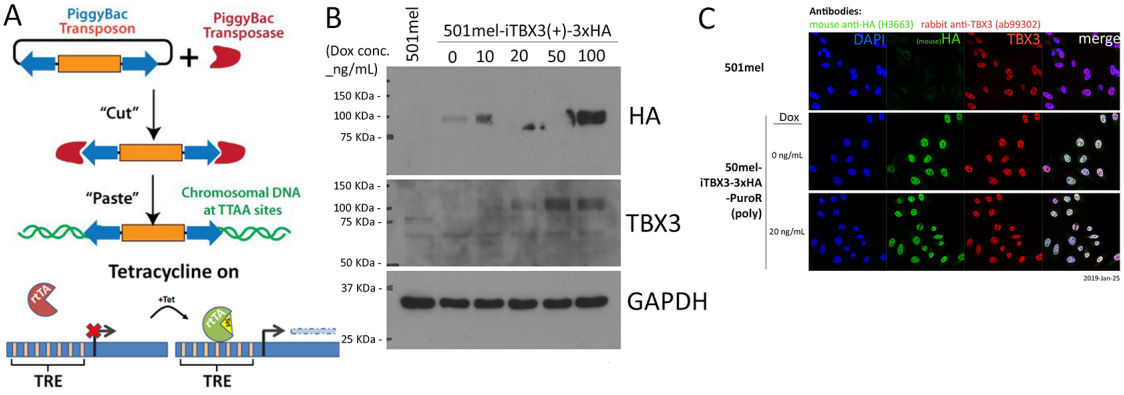



**Figure 5.1: Inducible expression of TBX3 in 501mel cells** (A) Schematic illustration of piggyBAC transposase system and Tet-on inducible expression. TRE, tetracycline response element. (B) WB result of polyclonal 501mel-iTBX3-3xHA-PuroR cells treated with different amount of doxycycline for 24 h. GAPDH was used as a loading control. (C) Immunofluorescence image of parental 501mel cells and polyclonal 501mel-iTBX3-3xHA-PuroR treated with 0 or 20 ng/ml of doxycycline (Dox).

The Tet-on system was chosen to allow a titratable expression of TBX3. TBX3 (+2a) cDNA was cloned in a CMV vector which adds an in-frame 3xHA sequence and puromycin resistance gene to the carboxy-terminus of TBX3. The puromycin N-acetyltransferase is cleaved from TBX3-3xHA by a P2A sequence and functions as

a selection marker. The TBX3(+2a)-3xHA-PuroR is under control of a tetracycline response element (TRE), which is not transcriptionally active under basal conditions and requires the binding of reverse tetracycline transactivator (rtTA) to activate downstream gene expression. Upon administration of doxycycline, a derivative of tetracycline, rtTA bound by doxycycline is now able to bind TRE and subsequently activate the expression of TBX3(+2a)-3xHA-PuroR.

The piggyBAC transposase system can efficiently integrate the gene of interest into the genome. In the CMV plasmid, the TBX3(+2a)-3xHA-PuroR sequence is flanked by inverted terminal repeat sequences (ITRs), which are recognised by the piggyBAC transposase and subsequently excised and integrated at TTAA sites randomly dispersed in the genome.

To generate the inducible cell line, three constructs were transfected into 501mel cells: the previously described TRE-TBX3(+2a)-3xHA-PuroR plasmid, the rtTA construct that is under the control of constitutively active CAG promoter and the piggyBAC transposase plasmid. The polyclonal 501mel-TRE-TBX3(+2a)-3xHA-PuroR (501mel-iTBX3-3xHA from here on) cell line was generated through puromycin selection. As the piggBAC transposase integrates the transposon at random TTAA sites throughout genome, depending on where the integration site is, the addition of TRE-TBX3(+2a)-3xHA-PuroR may lead to clone-specific variance in expression. Hence, we chose to use polyclonal cells for the ChIP-seq experiment to avoid any bias that might arise from clone picking.

### Characterisation of an inducible TBX3 cell line

To characterise the polyclonal 501mel-iTBX3-3xHA cell line, WB was used to examine the induction of the TBX3 protein and IF to check the subcellular location of the exogenous TBX3. I also tested different doxycycline concentration to choose one that induces HA-tagged TBX3 expression to levels similar to the endogenous protein.

As shown in Figure 5.1.B, TBX3 levels increase when the doxycycline concen-

tration increases. In the WB result, the TBX3 and HA bands have the same trend, confirming the induction of HA tagged TBX3. As highly overexpressed TBX3 might titrate out important interacting factors thus diluting the ones that can bind DNA with cofactors or binding to non-specific sites, I aimed for a TBX3-3xHA protein level that was comparable to that of endogenous TBX3, and chose 20 ng/ml of doxycycline for TBX3 ChIP-seq experiment.

TBX3 as a transcription factor needs to be in the nucleus to be able to bind DNA, thus it is important to test if HA tag affects its subcellular localisation. In addition, an excessive amount of TBX3 protein in a subset of cells is not desired and may be an issue if the polyclone is too heterogeneous — some cells overexpress TBX3 while others do not have it at all, and we need to ensure the protein level seen in WB reflects the whole population. The IF result in Figure 5.1.C shows that HA tagged TBX3 remains nuclear, and most cells in the polyclone are inducible.

### 5.2.2 TBX3 ChIP experiment

#### TBX3 ChIP-seq sample preparation

Experimental details are presented in section 2.5.1 on page 54. Briefly, DNA-protein complexes were cross-linked in dish using 1% formaldehyde or after trypsinization using 0.4% PFA, and chromatin was sheared by sonication into DNA fragments of proximate 200–500 bp. The insoluble fraction was then removed by high-speed centrifugation. Complexes containing TBX3 in the soluble fraction was immunoprecipitated using the HA antibody and magnetic Dynabeads. Protein and RNA were digested using enzymes, and DNA fragments were purified and sequenced by Illumina Truseq in Wellcome Trust Centre for Human Genetics.

#### TBX3 ChIP-seq analysis

Analysis process is presented in section 2.6.2 on page 63. Briefly, raw reads were mapped to hg38 reference genome using bowtie2, and peaks called using MACS2 taking a p-value of 0.01 as threshold.

## 5.3 ChIP-seq results

### 5.3.1 ChIP experiment validation

First, I tested whether using anti-HA antibody and magnetic beads successfully immunoprecipitated TBX3-3xHA from 501mel cells and if the sonication sheared DNA to the desired fragment size.

WB results show that in replicates for both crosslinking methods, there are bands at expected size in the IP samples, indicating a reasonable IP efficiency, as shown clearly by both anti-HA and anti-TBX3 antibody (Figure 5.2.A).

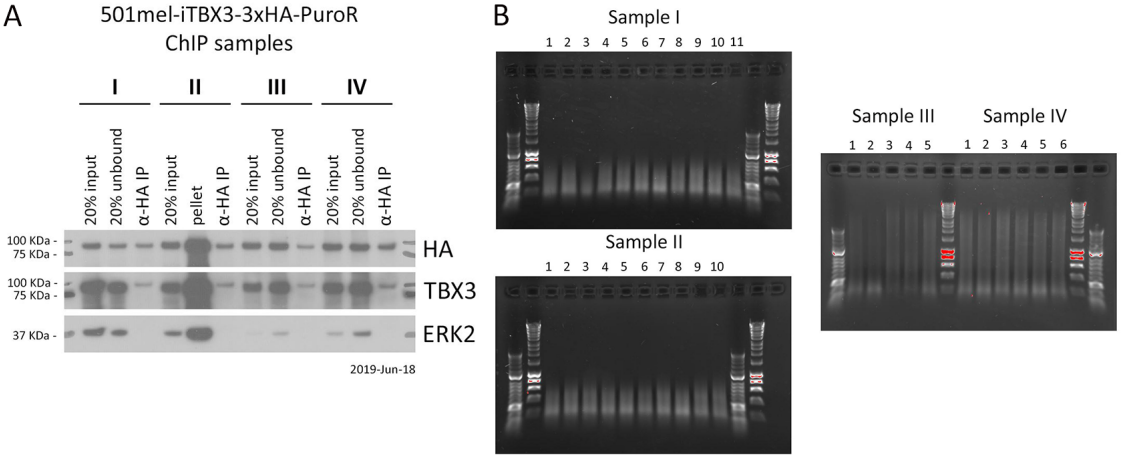



**Figure 5.2: ChIP experiment results** (A) WB results from HA immunoprecipitation of 501mel-iTBX3-3xHA-NeoR. ERK2 was used as a loading control. (B) DNA fragments after sonication were run on an agarose gel for size checking. Samples I and II were biological replicates, consisting of 30 150mm-dishes of cells, that were crosslinked with 0.4% PFA at room temperature for 10 min. Samples III and IV were biological replicates, consisting of 15 or 18 150mm-dishes of cells, that were crosslinked with 1% PFA at room temperature for 10 min.

To check the fragment length, I ran input DNA from every three dishes on

agarose gel. The results show that the fragment lengths all peak at around 200 bp (Figure 5.2.B), which is ideal for the paired-end 150 bp flow cell we chose, as it will give good coverage for DNA fragments under 300 bp. While the PFA crosslinked samples have a uniform fragment length, in the samples that were crosslinked in the dish with formaldehyde the peak is tailing towards a longer size (Figure 5.2.B), indicative of excessive crosslinking. Upon finishing the downstream analysis comparing the two crosslinking methods, I found that over-crosslinking is necessary for the TBX3 ChIP-seq — formaldehyde crosslinked samples called more reproducible peaks (Figure 5.3.A), in spite of a similar IP efficiency shown by WB (Figure 5.2.A). This might suggest that TBX3 binds DNA via interacting with other factors, or it comes off DNA so quickly that the trypsinization process was long enough to dissociate the interaction.

### 5.3.2 ChIP-seq quality check

To assess quality and reproducibility of the peaks, I used ChIPpeakAnno to overlap peaks called from each replicate. We can see that in PFA-crosslinked replicates, only 1/3 of peaks overlap; while in formaldehyde-crosslinked replicates, most peaks overlap, showing a better reproducibility (Figure 5.3.A). Peak score dot plot also shows a strong correlation between replicates (Figure 5.3.B). As the following analyses require the exact peak sequence and peak score and thus cannot combine multiple datasets, I chose to work with overlapping peaks of formaldehyde-crosslinked samples.

Next, to test if there are artefact peaks arising from HA antibody non-specific binding, I compared ChIP-seq results from 501mel-iTBX3-3xHA-PuroR cells and parental 501mel cells. Figure 5.3.C shows that 501mel-iTBX3-3xHA-PuroR cells called many more peaks and fewer than 1% of peaks overlapped with those called in parental cells, showing that the vast majority of peaks are indeed enriched DNA

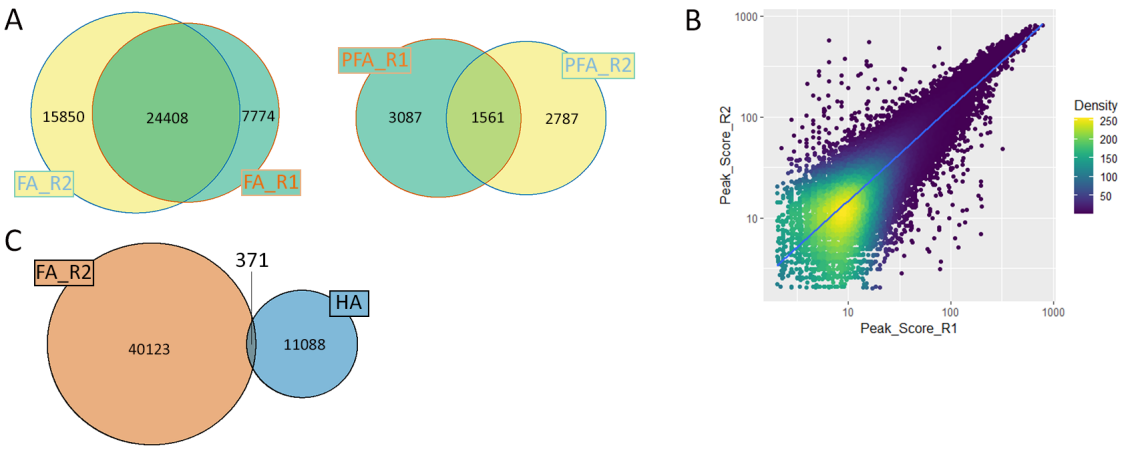



**Figure 5.3: TBX3 ChIP-seq quality control** (A) Venn diagram of peaks overlapping between biological replicates using different crosslinking methods. (B) Peak score correlation of formaldehyde crosslinked replicates. Dot plot and linear regression line were generated in R package ggplots. (C) Venn diagram of peaks overlapping between formaldehyde crosslinked 501mel-iTBX3-3xHA-NeoR replicate 2 (FA R2) and the negative control 501mel (HA). All samples went through same HA antibody IP and same process. (A) and (C) were generated in R package ChIPpeakAnno.

fragments crosslinked to immunoprecipitated TBX3-3xHA fusion protein.

### 5.3.3 ChIP-seq analysis

#### Representative peaks

The 501mel-iTBX3-3xHA ChIP-seq experiment generated high-quality peaks that were previously unknown, including one peak in the *MC1R* 5' UTR, one peak in a *NOTCH1* intron, one peak in a *RUNX3* exon, and one peak in a *BCL9L* intron (Figure 5.4.A-D). Melanocortin 1 receptor (MC1R) binds melanocyte-stimulating hormone and is overexpressed on the cell surface of most human melanomas (Rosenkranz et al. 2013). *NOTCH1* encodes a transmembrane receptor. NOTCH signalling is induced at cell-cell contact between a signal-sending and signal-receiving cell, and plays important role in regulating tumour microenvironment (TME) (Meurette and Mehlen 2018). The runt-related transcription factor 3 (RUNX3) lies downstream of the TGF-β signalling pathway, and has been shown to suppress cancer metastasis (CHEN et al. 2016). B-cell lymphoma 9-like (BCL9L) is a

coactivator of β-catenin, and loss of BCL9/9L blocks tumourigenesis in colorectal cancer models (Gay et al. 2019).

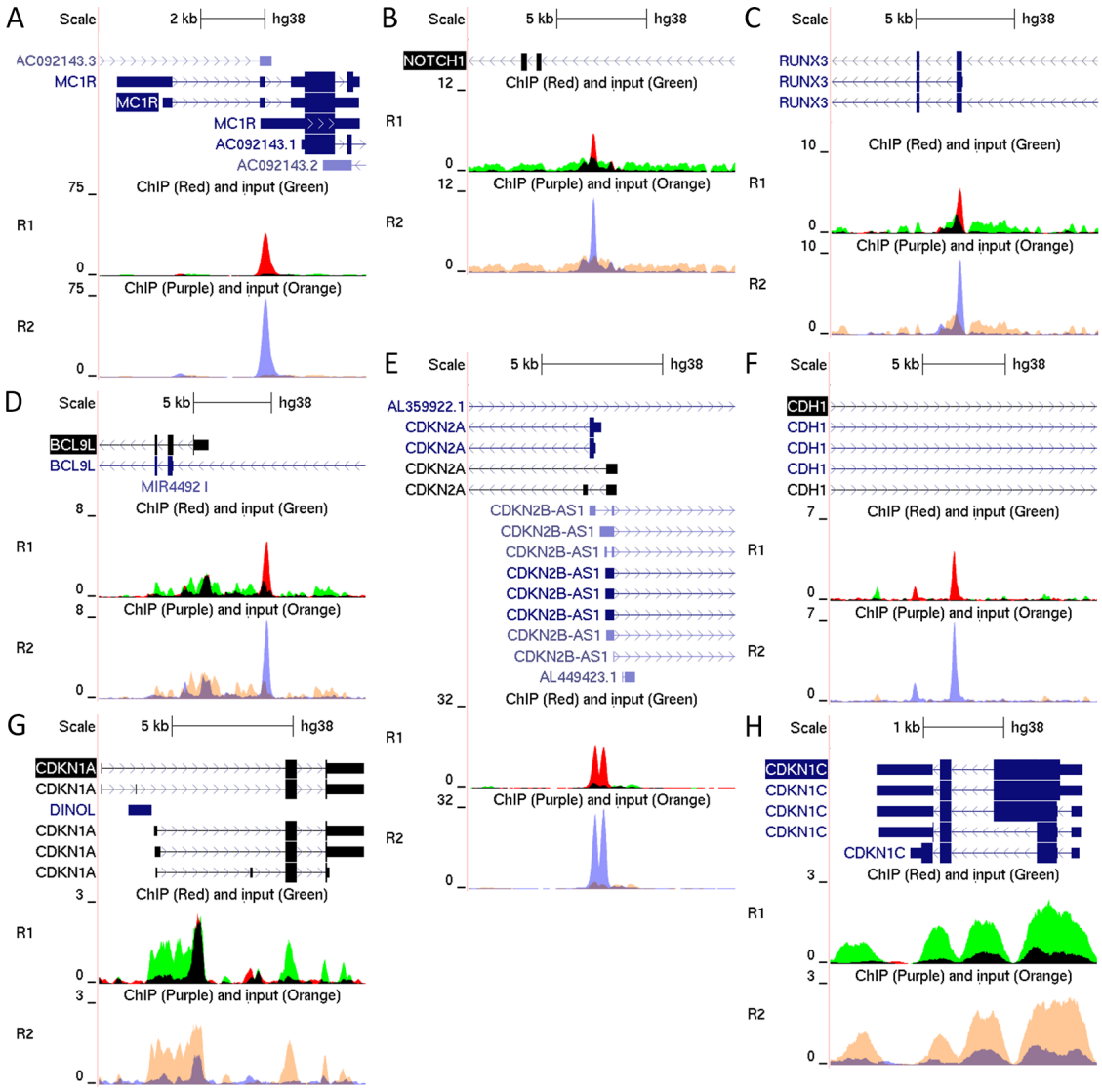



**Figure 5.4: TBX3 ChIP-seq peak examples** UCSC genome browser images showing representative peaks at (A) *MC1R* (B) *NOTCH1* (C) *RUNX3* (D) *BCL9L* (E) *CDKN2A-2B* (F) *CDH1* loci from formaldehyde-crosslinked replicates. ChIP and Input are overlayed in the same track for peak quality visualisation.

## Known TBX3 binding sites

It was revealed that the -19 to +54 region of *CDKN2A* initiator sequence contains a binding site for TBX2 (Jacobs et al. 2000; Lingbeek et al. 2002; S. Kumar et al. 2011) and TBX3 (Lingbeek et al. 2002; Yarosh et al. 2008). Confirming these reports, my

TBX3 ChIP-seq data also show a double peak spanning the TSS of *CDKN2A*, or one for *CDKN2A* and one for *CDKN2B-AS1* (Figure 5.4.E). Apart from $p14^{ARF}$/$p19^{Arf}$, another study (Tarryn Willmer, Hare, et al. 2016) reported that a *CDKN1A* binding site, containing a T-element at position -121 of the promoter, is responsible for TBX3 repression of $p21^{WAF1}$ in chondrosarcoma cells. However, my ChIP-seq data, which was generated in melanoma cells, did not call any peak at the *CDKN1A* gene (Figure 5.4.G). Our group's previous work established a TBX3 binding site in mouse *CDH1* promoter region that mediates the repression of E-cadherin (Rodriguez et al. 2008), however, in my ChIP-seq data, which was generated in human 501mel cells, the peak is in the second intron (Figure 5.4.F), suggesting a different mechanism in these melanoma cells. In addition, X. Li et al. 2018 reported that in human thyroid carcinoma cells, TBX3 represses $p57^{KIP2}$ via directly binding to two T half sites, -1,279 bp and -800 bp, in the *CDKN1C* promoter, which is not observed in the 501mel-iTBX3-3xHA data (Figure 5.4.H).

Taken together, a comparison of published binding sites and my data suggests the TBX3 repression mechanism differs, at least partially, in different tumour contexts. This may also be because previous studies all needed to make assumptions as to the location of potential binding sites, whereas the analysis presented here is unbiased. A similar discordance between many published binding sites based on ChIP-PCR versus those identified by unbiased ChIP-seq has been shown for the MITF transcription factor (Louphrasitthiphol et al. 2020).

### Binding site distribution

To determine where in the genome TBX3 binds, I analysed peak distribution using ChIPseeker. The results show that only 12.75% of peaks locate in promoter regions (less than 3 kb from TSS) (Figure 5.5.A), among which the read counts peak at the TSS (Figure 5.5.B). Intron binding accounts for 34.94% of peaks, while the majority of peaks (46.96%) are in distal intergenic regions.

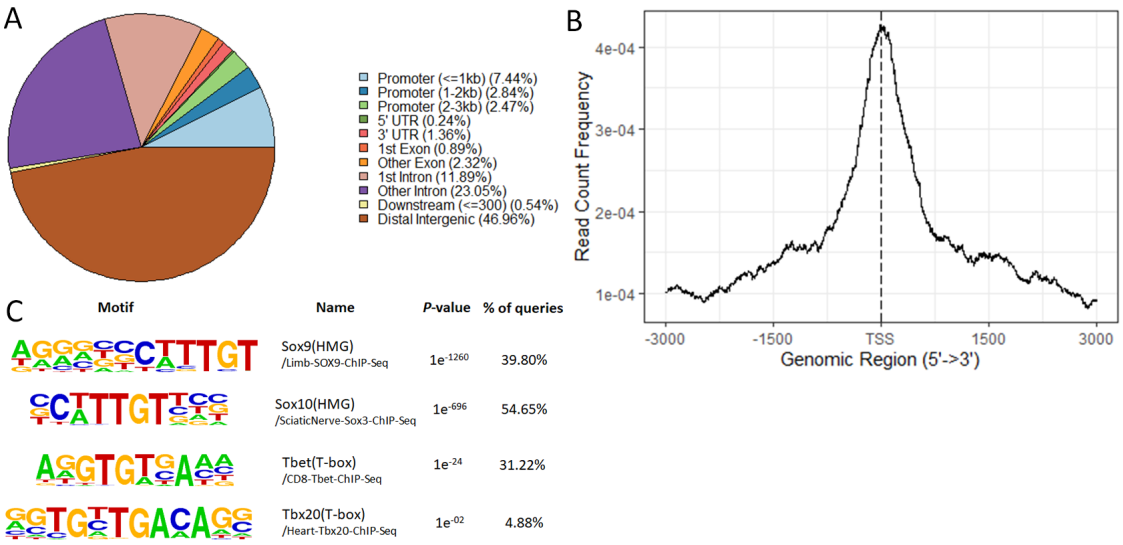



**Figure 5.5: TBX3 ChIP-seq analysis** (A) Genomic distribution of TBX3 ChIP-seq peaks (B) Peak read count in transcription start site (TSS) +/- 3 kb region (C) Enriched known motifs and respective p-value and enrichment percentage, identified by HOMER. (A) and (B) are generated in R package ChIPseeker, (C) in HOMER, using overlapping peaks of formaldehyde-crosslinked replicates of 501mel-iTBX3-3xHA.

### Motif enrichment

The published TBX3 binding motif (MA1566.1/Jaspar), the experiment of which was conducted in human HepG2 cells, has a variant T-element NNGGTGTGA, nevertheless it is important to assess if TBX3 binds the same motif in melanoma cells. As expected, HOMER known motif enrichment results include a T-bet binding motif (CD8-Tbet-ChIP-Seq(GSE33802)/Homer) that is enriched in 31.22% of peaks with a p-value of $1e^{-24}$ (Figure 5.5.C). To our surprise, among top enriched motifs there are SOX9 (Limb-SOX9-ChIP-Seq(GSE73225)/Homer) and SOX10 (Scia ticNerve-Sox3-ChIP-Seq(GSE35132)/Homer), with p-value of $1e^{-1260}$ and $1e^{-696}$ respectively. This unexpected result might unveil a previously unknown binding motif or interacting factors for TBX3.

## 5.4 Discussion

Identification of TBX3 direct targets involved in tumourigenesis is essential to understanding how it contributes to different stages of cancer progression, including promoting EMT and conferring therapy resistance. To date, there is only one reported TBX3 ChIP promoter array dataset that was generated in pre-invasive breast cancer cells (Krstic, Kolendowski, et al. 2019), and no TBX3 ChIP-seq dataset has been published in the context of cancer. Though there are Tbx3 ChIP-seq datasets generated in mESCs (Waghray et al. 2015; Han et al. 2010), they might not reflect the TBX3 binding profile in cancer. For example, in mammary glands Tbx3 does not suppress the expression of E-cadherin (Kunasegaran et al. 2014), which has been shown in several types of cancer is a TBX3 direct target, suggesting the target genes of TBX3 are cell type specific. Therefore, the ChIP-seq dataset generated in melanoma cells will add to our knowledge of TBX3 target genes in cancer progression by discovering previously unknown TBX3 binding sites.

### 5.4.1 Comparison of TBX2 and TBX3 binding sites

The DNA-binding domains and the C-terminal repression domains of TBX2 and TBX3 share high homology, and many individual studies suggested similar roles for TBX2 and TBX3 in various biological processes. Therefore, one can speculate that they regulate the same subsets of genes and have similar functions. Indeed, previous studies reported that common direct target genes of TBX2 and TBX3 include *CDKN2A* ($p14^{ARF}$/$p19^{Arf}$) (Lingbeek et al. 2002; Jacobs et al. 2000), *CDKN1A* ($p21^{CIP1}$) (Vance, Suzanne Carreira, et al. 2005; Prince et al. 2004; T. Lüdtke et al. 2013; Willem M H Hoogaars et al. 2008; Tarryn Willmer, Hare, et al. 2016), *PTEN* (Zhu et al. 2015; Burgucu et al. 2012), *CDH1* (E-cadherin) (B. Wang et al. 2012; Rodriguez et al. 2008) and *CX40* (Connexin-40) (W. M. Hoogaars et al. 2004; Christoffels et al. 2004).

Nevertheless, there are reports seeking to test if TBX2 and TBX3 have redundant functions, and the most significant difference lies in their ability to directly repress E-cadherin and impact cell invasion. For example, Peres, E. Davis, et al. 2010 addressed the difference between the role of TBX2 and TBX3 in VGP melanoma cells and MCF7 breast cancer cells, and found that while TBX2 contributes more to cell proliferation, Tbx3 suppresses E-cadherin expression and promotes cell migration. In addition, while silencing of TBX3 increased melanoma cells invasion and calcium-dependent cell-cell adhesion, silencing of TBX2 did not have similar effect. Consistent with this, Rodriguez et al. 2008 reported though overexpression of both TBX2 and TBX3 repressed the *CDH1* promoter and decreased E-cadherin protein level, only silencing of TBX3 increased E-cadherin expression.

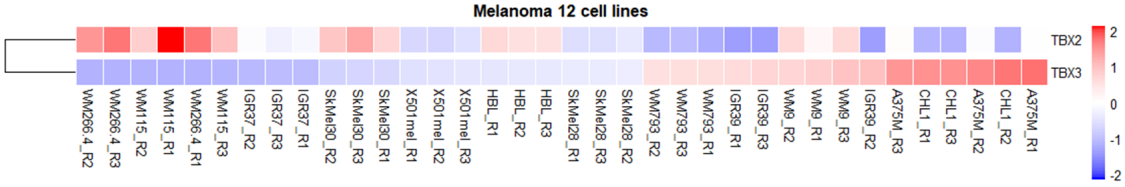



**Figure 5.6: Expression levels of TBX2 and TBX3 in melanoma cell lines** Heatmap of TBX2/3 normalised gene expression from triplicate 12 validated melanoma cell lines ranked by TBX3 expression, provided by Dr Pakavarin Louphrasitthiphol.

In line with the above reports, we also observed that the expression of TBX2 and TBX3 inversely correlate with each other in a panel of human melanoma cell lines (Figure 5.6), with high expression of TBX3 conferring a more invasive phenotype and high TBX2 level correlating with a more proliferative phenotype. Proliferation and invasion were assessed by cell number increase and Matrigel invasiveness assay respectively (Falletta et al. 2017).

Moreover, the newly generated ChIP-seq datasets seem to confirm the notion that closely related TBX2 and TBX3 only share a subset of their binding sites in melanoma cells (Figure 5.7.E). For example, both TBX2 and TBX3 bind to the promoter and TSS of *MC1R* and *MET*, the same intron region of *CDH1* , *BCL9L*

and *CDK2* (Figure 4.3, 4.10, 5.4 and 5.7.A, B, D, H and I). However, TBX2 does not bind the TBX3 site at a *RUNX3* exon (Figure 5.7.C), and TBX3 does not bind to other TBX2 binding loci such as those in *PTEN*, *NDRG1*, *TELO2*, *E2F1* and *CCNE2* (Figure 5.7.F–G, J–L). In addition, while TBX3 binds to the second intron in *NOTCH1*, the TBX2 binding site spans the second exon (Figure 5.4.B and 5.7.B). Furthermore, the *CDKN2A–CDKN2B* locus in B16 cells was deleted, so whether TBX2 can bind there in other melanoma cells remains unclear. As for the genomic distribution of ChIP peaks, most of TBX3 binding sites (46.96%) are in distal intergenic regions (Figure 5.5.A), while the majority of Tbx2 binding sites (37.93%) are in promoter regions (Figure 4.3.A).

*TBX2* and *TBX3* genes originated from tandem duplication of an ancestral locus at the base of vertebrate lineage (Horton et al. 2008). Although sharing 95% of homology in the T-box DNA-binding domains, TBX2 and TBX3 have a large portion of distinct binding sites, indicating that they may rely on other domains that interact with cofactors for DNA-binding specificity.

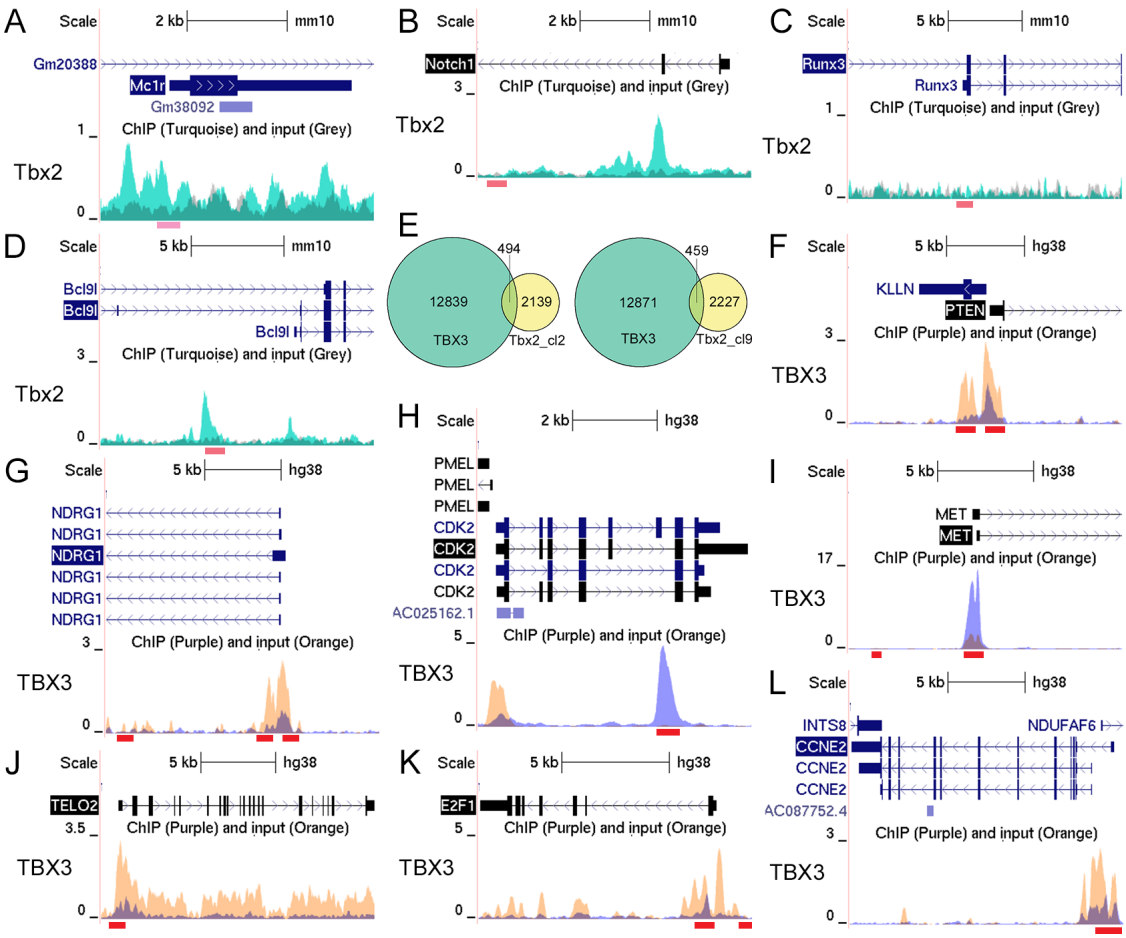



**Figure 5.7: Comparison of TBX3 and TBX2 binding sites** (A–D) UCSC mm10 genome browser images showing representative Tbx2 peaks at (A) *Mc1r*, (B) *Notch1*, (C) *Runx3* and (D) *Bcl9l* loci. Pink boxes highlight the relative location of TBX3 binding sites. (E) Venn diagram of overlapping of Tbx2 peaks that are common between biological replicates of clone#09 and TBX3 peaks that are common between biological replicates of formaldehyde-crosslinked samples; the binding sites of TBX3 were lifted over to mm10 using UCSC tools; the Venn diagram was generated in R package ChIPpeakAnno. (F–L) UCSC hg38 genome browser images showing representative TBX3 peaks at (F) *PTEN* (G) *NDRG1* (H) *CDK2* (I) *MET* (J) *TELO2* (K) *E2F1* (L) *CCNE2* loci. Red boxes highlight the relative location of Tbx2 binding sites.

# 6

# Interacting factors of Tbx2

## 6.1 Introduction

While most studies suggest a transcriptional repression role for TBX2 and TBX3, there are also cases where they seem to transactivate the target genes. How TBX2 and TBX3 switch between different transcriptional regulatory roles is thus an intriguing question. One answer might be the different interacting factors they recruit to the different DNA-binding sites. In this section, I will introduce the known interactors of TBX2 and TBX3.

### 6.1.1 Known TBX2 interactors

**HDAC** Vance, Suzanne Carreira, et al. 2005 showed the C-terminus of Tbx2 associates with the histone deacetylase 1 (HDAC1) *in vitro* and *in vivo* as evidenced

by GST pull-down assay and Co-IP in COS cells. A dominant negative form of Tbx2 protein (dnTbx2), which contains the DBD but lacks the C-terminal residues required for HDAC1 interaction displaced HDAC1 from the *CDKN1A* promoter as shown by ChIP-PCR, and leads to increased p21 expression. Later, Zhu et al. 2015 also reported that TBX2 recruits HDAC1 to the *PTEN* promoter. Though they did not show evidence of physical interaction between TBX2 and HDAC1, TBX2 and HDAC1 occupy the same region in the *PTEN* promoter and TBX2 expression level determines the binding ability of HDAC1 as shown by ChIP-qPCR. In addition, TBX2 recruits HDAC1 to the promoters of muscle-specific genes (Bo Zhu et al. 2014), and HDAC1 is required for repression of the *ADAM10* promoter by TBX2 in human neuroblastoma cells (Reinhardt et al. 2019).

**HP1** The Heterochromatin Protein 1 (HP1) family consists of 3 isoforms (α, β and γ). HP1β and HP1γ have been found at euchromatic sites, where they play a role in transcriptional repression. Co-IP shows that TBX2 associates with HP1 proteins (α, β and γ), and they all bind to the *NDRG1* promoter (-80/+4) in MCF7 cells as shown by ChIP-qPCR. TBX2 interacts with HP1α at 2 conserved HP1-binding motifs (PxVxL) located at 95-99 and 439-443 amino acids (Crawford et al. 2019). In addition, GST pull down assays show that Tbx2, like Tbx4, Tbx5 and Tbx6, can bind the histone H3 N-terminal tail (Demay et al. 2007).

**PRC2 and DNMT3b** Co-IP shows that TBX2 associates with SUZ12, EZH2 (components of PRC2) and DNMT3b (a DNA methyltransferase) to establish repressive H3K9me3 and H3K27me3 marks. Induction of dnTBX2 abrogated the recruitment of SUZ12 to the *NDRG1* promoter, and silencing of TBX2 abolishes H3K9me3 mark at the *NDRG1* promoter (Crawford et al. 2019).

**Rb1** GST pull down assay and Co-IP show that Tbx2 associates with the retinoblastoma protein (Rb1) via amino acids 287-301, which are immediately C-terminal to the T-box domain (Vance, Shaw, et al. 2010). Within the interacting sequence, L294 and L296 are most important for the interaction, as alanine

substitutes of them abolished the interaction and Tbx2's transcriptional repression function. Strikingly, in Rb1-negative cells, Tbx2 activates the p21 promoter.

**EGR1** TBX2 also associates with early growth response 1 (EGR1) to directly repress the *NDRG1* promoter as shown in Co-IP and ChIP-PCR experiments (Redmond et al. 2010). The same group also showed that EGR1 is required for the binding of TBX2 to the proximal *CST6* promoter (-153/+30) and subsequent transcriptional repression (D'Costa et al. 2014). This was also confirmed in Mohamad et al. 2018, where EGR1 physically associates with TBX2 in rhabdomyosarcoma cells, as shown in Co-IP.

**NRAGE** NRAGE physically interacts with the repression domain of TBX2, as shown in Co-IP experiments, and the interaction is required for NRAGE to bind and repress the *CDKN2A* promoter (S. Kumar et al. 2011).

**MyoD** and **Myogenin** MyoD and myogenin are myogenic regulatory factors (MRFs) that are required during normal myogenesis, and are used as diagonostic markers for rhabdomyosarcomas. TBX2 interacts with Myogenin and MyoD in rhabdomyosarcoma cells as shown by Co-IP (Bo Zhu et al. 2014). This interaction represses the activity of Myogenin and MyoD on target gene promoter, and counteracts MRF-driven activation of genes, including p21. Myogenin is required for the binding of TBX2 to the promoters of *CDKN1A* and muscle-specific genes. Notably, MyoD and myogenin both contain a bHLH domain and bind E-box motif (Shklover et al. 2007).

**Msx1/2** Msx1 and Msx2 are transcription factors that are induced by epithelial Bmp4 and in turn induces and maintains Bmp4 expression in dental mesenchyme. Tbx2 interacts with Msx1 and Msx2 in an *in vitro* maltose-binding protein (MBP) pull-down assay, and the T-box domain alone is able to retain this interaction (K.-J. Boogerd et al. 2008). Confirming this, GST pulldown and Co-IP using mouse mesenchymal cells show that Tbx2 and Msx1 proteins physically interact (Saadi et al. 2013). In addition, in $Msx1^{-/-}$ embryos, tooth development arrests at early

bud stage, and this can be ameliorated by Tbx2 gene dosage.

**Sox4** C. J. J. Boogerd et al. 2011 showed that Tbx2 interacts with Sox4 via the T-box domain in an *in vitro* MBP assay, and that Tbx2 counteracts the up-regulation capacity of Sox4 at a putative enhancer region in a *Gja1* intron.

**PML** TBX2 also interacts with promyelocytic leukemia protein (PML), a tumour suppressor important for establishment of cellular senescence, via the C-terminal repression domain as shown in Co-IP (Martin et al. 2012).

**Smad** Finally, Co-IP shows Tbx2 physically associates with Smad1/2/4 (Reich et al. 2020). The T-box domain together with the 'spacer' region after the N-terminal repression domain, are sufficient for Smad interaction, and repression of the BMP/GDF and Activin/Nodal pathways. R122 in the T-box domain is necessary for the DNA-binding activity for Tbx2, and substantially its repression function.

### 6.1.2 Known TBX3 interactors

**HDACs** and **PRC2** Yarosh et al. 2008 showed TBX3 interacts with HDAC1, 2, 3 and 5 in MCF7 breast cancer cell line in Co-IP. They also reported that while HDAC1, 2 and 3 interacted with both N-terminal and C-terminal half of TBX3, HDAC5 lost the interaction when the repression domain (amino acids 566–624) or the C-terminus was depleted. This experiment was performed in COS-7 cells where TBX3 and HDACs were ectopically expressed and labelled with Myc and Flag respectively, with immunoprecipitation performed using anti-Myc antibody. In addition, the interaction with HDACs is important for TBX3 to repress $p14^{ARF}$ expression, as shown in a luciferase assay. Dong et al. 2018 also reported that TBX3 interacts with HDAC5 via its repression domain, within which the $^{585}$LFSYPYT$^{591}$ and $^{604}$HRH$^{606}$ motifs are essential for this interaction and the repression of E-cadherin and consequently hepatocellular carcinoma (HCC) cell metastasis.

TBX3 recruits PRC2 and HDAC1/2 to two binding sites in the *CDKN1C*

promoter, as shown by ChIP-qPCR experiment performed in papillary thyroid carcinoma (PTC) cells — EZH2, a PRC2 component, and HDAC1/2 occupied TBX3 binding sites, but when TBX3 was stably knocked-down the occupancy was barely detectable (X. Li et al. 2018). The physical interaction was confirmed in Co-IP experiment performed in HEK293T cells expressing Flag-tagged TBX3, or in K1 cells using antibodies against endogenous proteins. The recruitment of PRC2 and HDAC1/2 is required for TBX3 repression of $p57^{KIP2}$, as dual inhibition of H3K27me3 and HDAC1/2 relieved the transcriptional repression of the *CDKN1C* promoter in a luciferase assay and restore $p57^{KIP2}$ protein level. A previous report also mentioned that Tbx3 activates Gata6 expression by repressing the H3K27me3 histone mark and inhibiting recruitment of PRC2 proteins EZH2, EED and SUZ12 to the *Gata6* promoter (R. Lu et al. 2010).

**RNA-binding proteins** An interesting report showed TBX3 interacts with RNA-binding proteins hnRNP C, DDX17, DDX3 and hnRNP U and regulates alternative splicing (P. Kumar et al. 2014).

**CAPERα** Coactivator of AP1 and Estrogen Receptor (CAPERα) acts as a pre-mRNA splicing factor. The above TBX3-interacting proteins were found in mass spectrometry (MS) data following Co-IP of TBX3. From the same MS dataset, the authors also identified CAPERα as a potential TBX3 cofactor, then confirmed the result in Co-IP performed in mouse embryos and *in vitro* MBP pull-down assay, and elucidated the interaction is mediated by the repression domain of TBX3 (Pavan et al. 2014).

**Nkx2.5** Tbx3 isoforms physically interact with homeobox factor Nkx2.5, as shown in GST pull down assay, and abrogate the synergistic activation of the *Nppa* promoter by Tbx5 and Nkx2.5 (Willem M H Hoogaars et al. 2008).

**Msx2** K.-J. Boogerd et al. 2008 showed Tbx3 interacts with muscle segment homeobox protein Msx2 in an *in vitro* MBP pull down assay as well as in Co-IP experiment, and that Tbx3 is required for the down-regulation of Connexin43

(*Cx43*) expression by Msx1 and Msx2.

**Sox4** C. J. J. Boogerd et al. 2011 showed that Tbx3 isoforms interact with Sox4 via the T-box domain in an *in vitro* MBP assay. ChIP data showed the binding sites of Tbx3 and Sox4 overlap in the first intron of *Gja1*, though this interaction has antagonistic potential — while Sox4 increases the luciferase expression, the addition of Tbx3 down-regulates it. Interestingly, in HEK293 cells, transfected Tbx3 isoforms are primarily in the cytoplasm, and co-transfection of Sox4 or Nkx2.5 translocates Tbx3 into the nucleus, though Tbx3 shows an exclusively nuclear localisation in other cell lines such as COS-7 or cardiac H10.

**Kif7** Finally, Tbx3 decreases Gli3 protein stability by interacting with Kif7, which is a ciliary protein required for the processing of Gli3 to its repressor form, as evidenced by Tbx3 Co-IP performed in mouse embryo lysates (Emechebe et al. 2016). In addition, IHC showed loss of Tbx3 led to aberrant Kif7 localisation.

### 6.1.3 Identifying novel TBX2/3 interactors

While some published interactors of TBX2 and TBX3, like HDAC, were confirmed in multiple biological contexts, others seem to only interact with TBX2 or TBX3 during a certain stage of normal development or pathological processes. Therefore, identifying novel TBX2 and TBX3 interacting factors with respect to their effect in melanoma is vital to our understanding of how they regulate the target genes expression and influencing the disease progression.

## 6.2 Protein-protein interaction of TBX2

### 6.2.1 Mass spectrometry data

As my ChIP-seq experiment suggested that Tbx2 relies on cofactors to bind DNA, it is important to expand our knowledge of its interactors and how the

interaction facilitates Tbx2's regulatory function. Before I joined the lab, mass spectrometry (MS) was carried out for an unbiased screen for Tbx2 interactors based on immunoprecipitation of epitope tagged TBX2 in HEK293 cells (Table 6.1).

| **Gene Name** | **SpecSum** |
|---|---|
| TBX2 | 346 |
| **BCOR** | 309 |
| TUBA1B | 271 |
| KIAA0182 | 147 |
| NCOR2 | 132 |
| CHD7 | 130 |
| ZFHX4 | 112 |
| NCOR1 | 101 |
| ZFHX3 | 88 |
| **KDM2B** | 74 |
| ZNF281 | 56 |
| TLE3 | 46 |
| GPATCH8 | 37 |
| SATB1 | 35 |
| SATB2 | 34 |
| ZNF609 | 31 |
| TCF7L2 | 29 |
| TRPS1 | 27 |
| FOXP1 | 26 |
| TFAP2A | 26 |
| TNRC18 | 26 |
| MORC4 | 23 |
| ARID5B | 22 |
| CHD6 | 22 |
| **PCGF1** | 22 |
| SOX5 | 19 |
| ZNF608 | 19 |
| TLE1 | 18 |
| HOXA13 | 18 |
| ZNF362 | 18 |
| **BCORL1** | 17 |

**Table 6.1: AP-MS result for Tbx2** SpecSum indicates the numbers of peptides corresponding to the interacting proteins listed. Bold indicates components of PRC1.

Included in the potential interactors are, DNA-binding transcription factors that might facilitate recruitment of Tbx2 to specific targets, such as SOX5, Forkhead

box protein P1 (FOXP1), transcription factor AP-2-alpha (TF2P2A); transcription cofactors implicated in chromatin modification, such as Special AT-rich Sequence Binding Protein 1/2 (SATB1/2), nuclear receptor corepressor 1/2 (NCOR1/2), and chromodomain-helicase-DNA-binding protein 6/7 (CHD6/7). I used STRING analysis to classify the potential interacting proteins and found three components from the polycomb repressive complexes PRC1, BCOR, KDM2B and PCGF1, which form a non-canonical PRC1.

KDM2B demethylates H3K36me3 that sequentially recruits PRC1 for monoubiquitination of H2AK119, leading to chromatin compaction and subsequent transcriptional silencing of target genes (Gil and O'Loghlen 2014). Besides, abundant evidence showed canonical PRC1 regulates senescence through the *CDKN2A/2B* locus, and similar immortalising function via repressing the same locus was reported for KDM2B (Tzatsos, Paskaleva, et al. 2011; Tzatsos, Pfau, et al. 2009; J. He et al. 2008).

Many studies postulated a transcriptional repressor role for TBX2/3, while only few explored the mechanism behind it. A recent study from X. Li et al. 2018 reported that recruitment of PRC2 and HDAC1/2 is needed for TBX3 to repress $p57^{KIP2}$. Hence, the potential interaction of TBX2, which belong to the same subfamily of T-box transcription factors with TBX3, with PRC1 to establish repressive histone marks would suggest a related but complementary mechanism of regulation of its target genes.

### 6.2.2 Interaction between Tbx2 and PRC1

The mass spectrometry data obtained following immunoprecipitation (IP) of TBX2 suggest that components of non-canonical KDM2B-PRC1 interact with TBX2, which requires further validation. As I mentioned in section 4.5.1 that a lot of TBX2 remains associated with the insoluble fraction, and thus Co-IP only enriches for TBX2 fraction that is soluble. Therefore, specific interactions related to chromatin

may be missed, as frequently the conditions required to extract Tbx2 for a Co-IP might disrupt complex formation. Nevertheless, the MS essentially was a Co-IP experiment using sheared chromatin and may indicate factors associated with TBX2 or with nearby chromatin. Therefore, I chose to validate the interaction *in vivo* using a LacO/LacI tethering assay, where a 'bait protein' is used to test if a 'prey protein' interacts with it. This assay would be complementary to the Co-IP assay used in MS.

The bait protein is fused to Lac repressor protein (LacI/LacR) that binds to arrays of Lac operator (LacO) sequence integrated in the genome, so the bait protein accumulates in LacO repeats in a small nuclear focus or 'dot'. If the prey protein interacts with the bait, it will also accumulate at the same locus, otherwise it will be evenly diffused. When bait and prey are labelled with different fluorescent proteins, we will see a bright dot of bait protein, and if prey is recruited there will be a second overlapping dot. If there is no interaction between the two, the prey fluorescence will be evenly distributed across the nucleus (Figure 6.1.A).

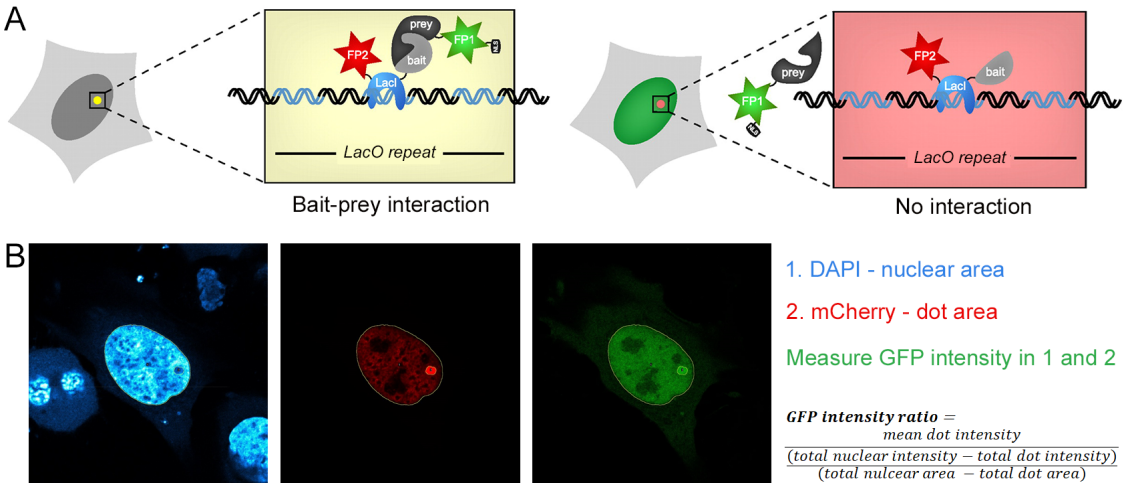



**Figure 6.1: LacO/LacI tethering assay illustration** (A) Schematic illustration of LacO/LacI nuclei dot assay. FP, fluorescent protein; LacI, Lac repressor protein. Adapted from Zolghadr et al. 2012. (B) Dot area fluorescence enrichment quantification method

In my experiment, I used full-length, N-terminal half (1-372 aa) or C-terminal half (373-711 aa) of Tbx2 protein as bait labelled with mCherry, with the PRC1 component Pcgf1 as the 'prey' labelled with eGFP on either end of the protein.

We can see that Tbx2 forms a red dot within the nucleus reflecting its localisation at the LacO repeats. When Pcgf1-eGFP is co-expressed, the red Tbx2 and green Pcgf1 signals co-localise (Figure 6.2.A and B). By contrast, no Pcgf1 dot was observed when a negative control — mCherry fused to LacR without TBX2 — was transfected into cells (Figure 6.2.A and B); in addition, eGFP alone does not interact with mCherry-Tbx2-LacR (Figure 6.3.C top two panels). To identify the domain responsible for the interaction, I first divided Tbx2 into two parts. Figure 6.2.C shows that while the C-terminus of Tbx2 still recruits Pcgf1, the accumulation of green fluorescence is much weaker when the N-terminus of Tbx2 was used, suggesting the interacting domain predominantly locates in the C-terminus of Tbx2, consistent with original observations that indicate the TBX2 repression domain is within the C-terminal region (Carreira et al. 1998).

To confirm the observation and avoid bias, I also measured the green fluorescence in the bait dot area and compared it to other parts of the nucleus to determine whether Pcgf1 protein was enriched in that location. The dot/nuclear fluorescence intensity ratio was calculated as $\frac{GFP_{\mathrm{d}}}{(GFP_{\mathrm{tn}}-GFP_{\mathrm{td}})\cdot(ns-ds)}$, where *ns* is nuclear size, *ds* is dot area size, $GFP_n$ and $GFP_d$ stand for mean GFP intensity in nuclear and dot area. If there is accumulation of Pcgf1 in the dot area, the ratio should be larger than 1 (Figure 6.1.B). The result is shown in Figure 6.2.D, confirming the observation that C-terminus of Tbx2 recruits Pcgf1 to a level similar to full length, while the N-terminal half does so much less. Notably, the number for the negative control is bigger than one (Figure 6.2.D and Figure 6.3.D), despite a clear image of even diffusion (Figure 6.2.B and Figure 6.3.C top two panels). This is due to the 'black cavity' within nucleoli, making the mean nuclear GFP intensity lower than the actual nucleoplasmic intensity.

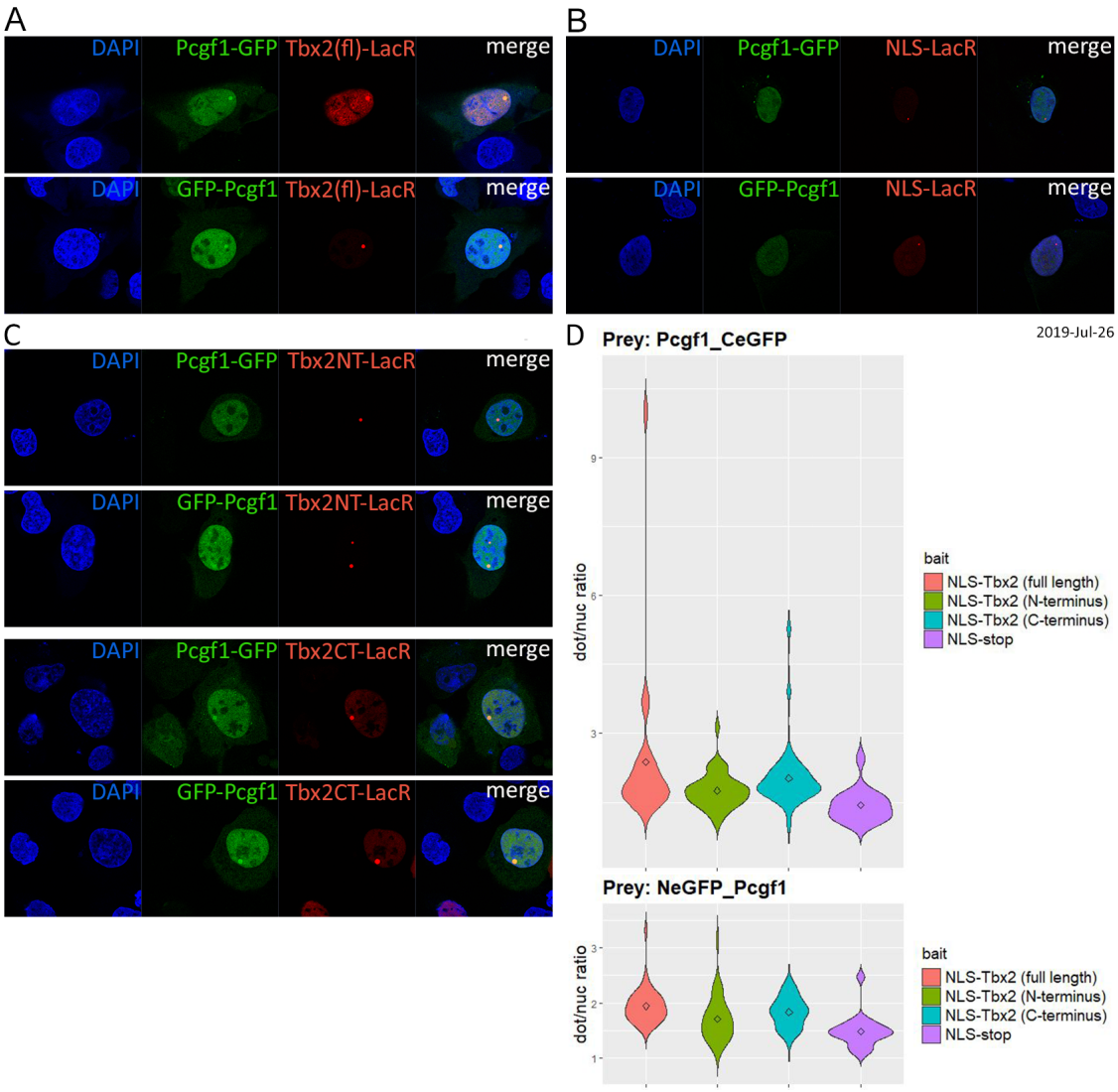



**Figure 6.2: LacO/LacI tethering assay result of Tbx2 and Pcgf1** (A-C) Representative images from U2OS-LacO cells transfected with (A) Pcgf1-eGFP and mCherry-Tbx2-LacR (B) Pcgf1-eGFP and negative control mCherry-NLS-LacR (C) Pcgf1-eGFP and N-terminus (1–372 aa) or C-terminus (373–711 aa) of Tbx2 (D) Quantification of eGFP fluorescence intensity dot/nuclear ratio, data from individual cells were plotted in violin graph using R package ggplot2. In the violin plot, the width indicates cell number, while the height indicates the eGFP fluorescence intensity dot/nuclear ratio. Note: Pcgf1 is labelled with eGFP on either the C-terminus or the N-terminus.

## 6.3 Other cofactors of TBX2

My ChIP-seq data showed that E-box motif is significantly enriched in Tbx2-binding sites (Figure 4.3.C). However, Tbx2 protein does not contain a basic helix-loop-helix (bHLH) DNA-binding domain, which recognises E-box motif. It seems therefore

unlikely that Tbx2 can directly bind to E-box, but more likely does so by interacting with cofactors containing a bHLH domain.

As shown in Figure 4.4, comparison between Tbx2 and USF1/2 ChIP-seq datasets revealed that they share common binding sites. This led us to the hypothesis that Tbx2 might bind to E-box motifs by interacting with USF1/2.

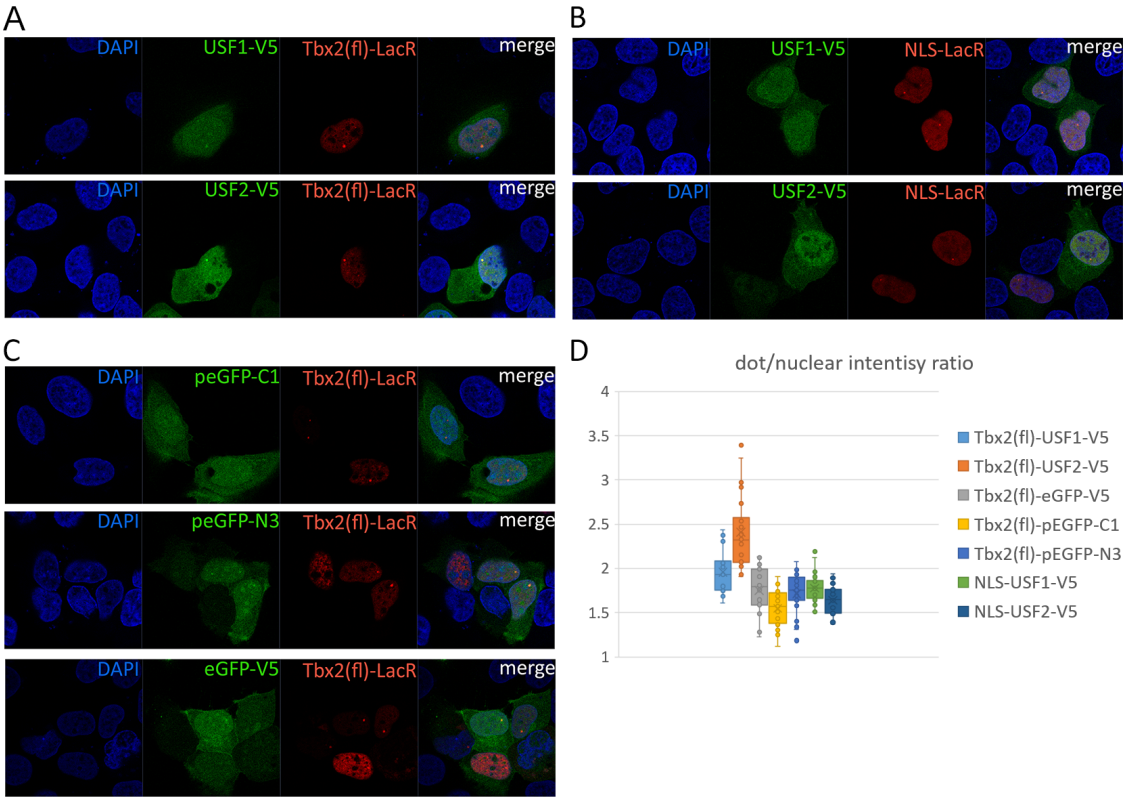



**Figure 6.3: LacO/LacI tethering assay of Tbx2 and Usf1/2** (A-C) Representative images from U2OS-LacO cells transfected with (A) mCherry-Tbx2(fl)-LacR and USF1/2-eGFP-V5 (B) negative control mCherry-NLS-LacR and USF1/2-eGFP-V5 (C) mCherry-Tbx2(fl)-LacR and eGFP-MCS, MCS-eGFP or eGFP-V5 (D) Quantification of eGFP fluorescence intensity dot/nuclear ratio, data from individual cells were plotted in Excel.

Therefore, I conducted LacO/LacI tethering assay to test this hypothesis. We can see that USF2-eGFP-V5 forms a bright green dot overlapping that of mCherry-Tbx2(fl)-LacR; USF1-eGFP-V5 also forms a dot, though the interaction is relatively weak (Figure 6.3.A). When negative control mCherry-NLS-LacR was used, USF1/2-eGFP-V5 is evenly diffused in the nuclei (Figure 6.3.B). Similarly, empty eGFP vector peGFP-C1 did not interact with mCherry-NLS-LacR, though when the

multiple cloning site (MCS) is moved to the N-terminus of eGFP (peGFP-N3) or then eGFP is fused with V5 tag, the interaction became slightly stronger (Figure 6.3.C). The background dot/nuclear area eGFP intensity ratio from different negative control varies, nevertheless, quantification clearly shows USF2 interacts with full-length Tbx2 (Figure 6.3.D).

## 6.4 Discussion

Prevailing data suggested a transcription repressive function for TBX2 and TBX3, however, there are also some reports showing they can also directly activate gene expression (Yao et al. 2014; Nandana et al. 2017; Krstic, Kolendowski, et al. 2019; R. Lu et al. 2010). How TBX2 and TBX3 switch between repressor and activator has not been elucidated.

One possibility is that they inhibit transcription by recruiting histone-modifying cofactors to initiate chromatin compaction in the promoter of target genes. Indeed, it has been shown that TBX3 recruits PRC2 to two binding sites in the *CDKN1C* promoter and promote methylation of H3K27, which is required for TBX3 repression of $p57^{KIP2}$ (X. Li et al. 2018).

Our TBX2 IP-MS data suggest that TBX2 might interact with a non-canonical PRC1, comprising BCOR, KDM2B and PCGF1, which ubiquitinates H2AK119 and leads to chromatin compaction and subsequent gene silencing. I validated the interaction between Tbx2 and Pcgf1 using the LacO/LacI tethering assay, and found the interaction is likely to be mediated by the C-terminus of Tbx2. To further confirm this, interaction with other components of PRC1 is also needed. In addition, it remains undetermined whether Tbx2 recruits PRC1 to target gene sites and establish a repressive histone mark.

As for the opposite function, TBX2 and TBX3 might activate gene expression by recruiting cofactors that associate with RNA polymerase to initiated transcription.

The E-box motif was enriched in both Tbx2 and TBX3 binding sites, therefore, we speculate that they might interact with bHLH transcription factors that bind E-box motif. Indeed, preliminary results from the LacO/LacI tethering assay indicate that the bHLH-containing USF2 interacts with TBX2. However, we did not see USF2 in the TBX2 IP-MS data. Co-immunoprecipitation (Co-IP) is widely-used for protein-protein interaction identification and validation, and its advantages include proteins that are interacting are in post-translational modified and are in their native conformation. However, Co-IP is not sensitive enough for low-affinity or transient protein-protein interaction, which might explain why USF2 was not observed in Tbx2 IP-MS data. Nevertheless, additional experiments are needed to validate the interaction between USF2 and TBX2, and to further assess if the interaction confers a transactivating function for TBX2.

# 7

# Regulation of TBX2 and TBX3

## 7.1 Introduction

The majority of published studies show TBX2 and TBX3 are overexpressed in tumours and that they contribute to different stages of tumour progression through bypassing oncogene-induced senescence, promoting tumour growth, driving metastasis and conferring therapy resistance. Therefore, down-regulating TBX2 and TBX3 can potentially be a novel anti-cancer therapeutic strategy, especially since interfering with TBX2 function using a dnTBX2 could induce senescence in human and mouse melanoma cell lines (Vance, Suzanne Carreira, et al. 2005), and silencing of TBX2 caused G1/S cell cycle arrest and reduced melanoma cell number (Suzanne Carreira et al. 2000), which were also observed in my experiment (Figure 4.11.A–B). However, transcription factors are often considered to be 'undruggable' due to their large and 'featureless' DNA-interacting surface. Alternatively, targeting factors that

regulate TBX2 and TBX3 expression would potentially serve the same purpose. It is therefore crucial to expand our knowledge on regulators of TBX2 and TBX3. In this section, I will review the known factors and signalling pathways that regulate TBX2 and TBX3 expression.

### 7.1.1 Transcriptional regulation of TBX2/3

As mentioned in section 5.4, TBX2 and TBX3 expression levels are inversely correlated in a panel of melanoma cell lines (Figure 5.6), and multiple reports showed that they repress each other's expression. For example, J. Li, D. Ballim, et al. 2014 reported that TBX3 binds a T-element in the *TBX2* proximal promoter and represses its transcription, as shown in ChIP-qPCR and luciferase assay. Agreeing with this, Oh et al. 2019 also showed that TBX3 represses TBX2 mRNA and protein levels in rhabdomyosarcoma cells, and ChIP-qPCR shows TBX3 binds the *TBX2* promoter. In addition, the mRNA levels of TBX2 and TBX3 in sarcoma patients inversely correlate with each other. Furthermore, silencing of TBX3 in melanoma cells lead to substantial increase in TBX2 expression (Rodriguez et al. 2008).

Several studies reported regulators of TBX2 and/or TBX3 that directly bind to their promoter and alter their expression, which are listed below.

**PAX3** F. Liu et al. 2013 showed PAX3 activates TBX2 expression by directly binding to and activates the *TBX2* promoter as evidenced by ChIP-PCR and luciferase assay.

**MITF** and **Usf1** EMSA shows that Usf1 and Mitf bind the E-box sequence in the *Tbx2* and activate Tbx2 expression. (Suzanne Carreira et al. 2000)

**Retinoic acid receptors** While maternal injection of retinoic acid (RA) to mice represses Tbx2 transcription in murine development through a retinoic acid response element (RARE) in the *Tbx2* promoter (Sakabe et al. 2011), in B16 melanoma cells RA up-regulated Tbx2 mRNA level (Boskovic and Niles 2004).

These results suggest the regulation of Tbx2 by RA varies in different biological contexts. In human melanoma cell lines, RA increases TBX3 mRNA and protein levels (R. Ballim et al. 2012). Luciferase assays showed the activation of TBX3 is mediated by a RARE half site at -75 bp, and DAI and ChIP-PCR suggest RA receptors directly bind the *TBX3* promoter.

**Smads** Yang et al. 2006 showed TBX3 promoter contains a Smad-binding element and is directly bound by Bmp Smads, as shown by *in vivo* ChIP-PCR. Luciferase assays in mouse embryonal carcinoma cells revealed that the *Tbx3* promoter is activated by Smad1 and -4. In addition, J. Li, M. S. Weinberg, et al. 2013 showed Smad3/4 and JunB bind the *TBX3* promoter region and activate TBX3 expression.

**c-Myc** Tarryn Willmer, Peres, et al. 2015 showed that TBX3 mRNA level peaks in S-phase, when c-Myc binds the *TBX3* promoter region containing two E-box sites and activates its transcription.

**PML** Promyelocytic leukemia protein (PML) is a tumour suppressor that is frequently mutated or lost in cancers. PML isoform IV binds to the *TBX2* promoter, as evidenced by luciferase assay. While PML represses TBX2 expression in $Ras^{V12}$-induced senescence, overexpression of TBX2 counteracts the induction of senescence and reduction in cell proliferation driven by PML-IV, for which the repression domain in the TBX2 C-terminus is required (Martin et al. 2012).

**HAND2** HAND2, a bHLH transcriptional regulator, activates Tbx3 in mouse limb buds. ChIP-seq and ChIP-qPCR show that HAND2 binds to Tbx3 -58 kb and -19 kb region (Osterwalder et al. 2014).

**Dpf2** and **Eed** Dpf2 is a subunit of Baf45 complex that maintains pluripotency and differentiation potential of ESCs. Dpf2 binds to two distinct enhancers at the *Tbx3* locus and activates Tbx3 expression by maintaining the H3K27ac histone mark in these regions. By contrast, PRC2 subunit Eed opposes Dpf2's regulation by repressing Tbx3 expression, and its binding at the *Tbx3* intronic enhancer is

increased when Dpf2 is depleted (W. Zhang et al. 2019).

**β-catenin** In human HCCs and hepatoblastomas, Tbx3 expression is activated by mutant β-catenin both *in vitro* and *in vivo*. β-catenin and the T-cell factor Tcf4 are recruited to the *Tbx3* promoter and activate Tbx3 expression, as shown by ChIP-PCR and luciferase assay (Renard et al. 2007).

**AP-1 factors** Phorbol ester PMA leads to activation of activator protein-1 (AP-1) factors via protein kinase C (PKC)-mediated transcriptional activation. Mowla et al. 2011 reported that AP-1 components c-Jun/JunB bind to the TBX3 promoter and mediate the up-regulation of *TBX3* induced by PMA in in normal human fetal lung fibroblasts and prostate epithelium.

**PRC2** EZH2 is the catalytic unit of PRC2 that methylates histone H3 lysine 27 (H3K27me), while JARID2 is a non-catalytic, substoichiometric component. Silencing of EZH2 or JARID2 enhanced the expression of TBX3 and inhibited that of TBX2 (Oh et al. 2019). ChIP-qPCR in mouse muscle myoblasts shows that H3K27me3 mark is enriched in the *TBX3* promoter, and depletion of JARID2 abrogates this enrichment. Taken together these results indicate TBX3 is directly repressed by PRC2.

### 7.1.2 Post-transcriptional regulation of TBX2/3

TBX2 and TBX3 expression is also regulated by microRNAs. For example, microRNA-7 (miR-7) is reduced in glioblastoma multifome (GBM) and affects the invasive ability of GBM cells. Pan et al. 2020 reported miR-7 inhibits the expression of TBX2 by repressing the 3' UTR region of TBX2.

Tbx3 is normally down-regulated at the time of hindlimb bud initiation, however, when Dicer, an enzyme processing microRNAs to their mature forms, is inactivated, Tbx3 expression increased, causing a delay and posterior shift of limb bud. MicroRNA miR-363 was identified as a key mediator of the inhibition via the Tbx3 3'

UTR as shown by luciferase assay (Zhen Zhang et al. 2011). Another microRNA regulator of Tbx3 is miR-137, which is highly expressed in the mitotic phase of ESCs. miR-137 activates ES cells differentiation and inhibits their proliferation *in vitro*, partly via reducing Tbx3 and Klf4 mRNA and protein levels (Jiang et al. 2013). Among primary melanoma patients, low levels of miR-137 and high levels of TBX3 associate with poor prognosis in terms of overall survival (Peres, Kwesi-Maliepaard, et al. 2017). Luciferase reporter assay shows that miR-137 represses the TBX3 3'UTR via a miR-137 binding site. miR-137 inhibitor treatment leads to increased activity of the TBX3 3'UTR as well as higher protein level, while miR-137 mimic does the opposite. In addition, Luciferase assays showed that miR-206 directly binds the TBX3 3' UTR and represses its expression, and subsequently impairs breast cancer cells growth in Matrigel and tumoursphere formation (Amir et al. 2016). Moreover, miR-93 represses the Tbx3 3' UTR activity and thereby inhibits the self-renewal of adipocyte precursors (Cioffi et al. 2015).

### 7.1.3 Post-translational modification of TBX2/3

Post-translational modifications (PTMs) of TBX2 and TBX3 can change their protein stability or subcellular localisation and consequently affect their regulatory function.

**p38** Abrahams, Mowla, et al. 2008 reported that ultraviolet C (UVC) irradiation increases TBX2 protein levels in breast cancer cells, which is mediated by p38 phosphorylation at S336, S623 and S675, as shown in *in vitro* kinase assays. Collective alanine substitution of the three sites abolished UVC-mediated up-regulation of TBX2, and resulted in a shorter TBX2 half life. In addition, the authors also claimed that UVC irradiation promotes TBX2 nuclear localisation and consequently the ability to repress *p21* gene. This claim seems less convincing as the Human Protein Atlas and other reports show TBX2 is constitutively nuclear, and

in this report there is no quantification for the localisation change. Furthermore, Yano et al. 2011 reported that p38 phosphorylates Tbx3 at S692 and that a S692A mutant is less capable of repressing E-cadherin expression.

**AKT3** Peres, Mowla, et al. 2014 identified TBX3 as a substrate for AKT3 in melanoma development, showing that AKT3 phosphorylates TBX3 protein at S720, which is in a putative AKT motif "RsRsaS$^{720}$", stabilising it and promoting its nuclear localisation. IF results showed that a TBX3 S720A mutant is 100% cytoplasmic, and has less transcriptional repression of E-cadherin, a well-established TBX3 target gene.

**CDK2** Another report showed during S-phase TBX3 is phosphorylated by cyclin A-CDK2 in the N-terminus, and the two consensus CDK2 recognition sites are S190 and S354 (Tarryn Willmer, Peres, et al. 2015). In addition, in chondrosarcoma cells and normal human prostate epithelium, silencing of CDK2 or pharmacological inhibition of cyclin A -CDK2 diminished TBX3 protein level, suggesting that CDK2-mediated phosphorylation of TBX3 leads to increased protein stability.

While the above report indicates phosphorylation of TBX3 S190 contributes to protein stability, another report shows that pseudo-phosphorylation of S190 abolishes TBX3 binding to the *CDKN1A* promoter and consequently repression of p21 expression is reduced, as shown using DNA affinity immunoblotting (DAI) and luciferase assay respectively (Tarryn Willmer, Hare, et al. 2016).

### 7.1.4 Other known regulators of TBX2/3

Tbx2 protein level in B16 mouse melanoma cells is regulated by cell cycle, being low in G1, increasing in S-phase and staying high in G2 and reducing at mitosis (Bilican and Colin R Goding 2006). Interesting, the mRNA level of Tbx2 is lower in G2 compared to G1/S, suggesting the cell cycle regulation of Tbx2 expression mainly happens at post-transcriptional level.

Apart from cell cycle, The following factors and signalling pathways also regulate TBX2 and TBX3 expression, though detailed mechanisms are not known.

**BRAF$^{V600E}$ mutation** TBX3 expression is induced by MAPK signalling downstream from the melanoma associated BRAF$^{V600E}$ mutation in melanocytes, leading to TBX3-mediated repression of CDH1 (Boyd et al. 2012). Exogenous expression of BRAF$^{V600E}$ increased the activity of the *TBX3* promoter, though the transcription factor(s) mediating the regulation of TBX3 expression are not yet defined. Contradictory to this report, Niwa et al. 2009 showed TBX3 expression is repressed by MAPK signalling pathway in mouse ESCs.

**BMP** Yamada et al. 2000 showed BMP2 induces the expression of Tbx2 in early avian and murine embryos, and the expression pattern of chicken *cTbx2* and *cTbx3* overlaps with that of *Bmp2* in the developing heart. Tbx2 is expressed in mouse dental mesenchyme and can be induced by Bmp4 (Saadi et al. 2013). Tbx2 expression is induced by Bmp signalling in the limb bud mesenchyme (Farin et al. 2013). Similarly, Cho et al. 2006 showed overexpression of Bmp4 inhibited Tbx3 expression in mouse embryos.

**FGF signalling** *In vivo* study showed Fgfr1 and Wnt inhibitors abolished Tbx3 expression in mammary-forming region in embryos, suggesting that Tbx3 is downstream of Fgf and Wnt signalling in mammary gland initiation (Eblaghie et al. 2004). TBX2 expression in SW13 cells is up-regulated by FGF-4 (Ismail and Bateman 2009). Another study showed Tbx3 mRNA and protein levels in MCF7 cells increased when stimulated with estrogen and FGF9, which can be abrogated by FGFR signalling inhibitor (Fillmore et al. 2010).

**Shh signalling** Sonic hedgehog (Shh) is an important protein in embryonic lung development. In Shh$^{-/-}$ lungs, *Tbx2* and *Tbx3* transcript levels were significantly reduced (Y. Li et al. 2004). T. H. Lüdtke et al. 2016 also showed in the lung mesenchyme Tbx2 and Tbx3 function downstream of Shh signalling.

**PI3K-AKT signalling** In mouse ESCs, under the control of LIF signal Tbx3

mRNA and protein levels are activated by the PI3K–AKT pathway (Niwa et al. 2009). In agreement with this, Ke et al. 2018 also showed PI3K–AKT pathway is involved in LIF-mediated activation of Tbx3 expression. In addition, the PI3K inhibitor LY29400 also decreased TBX2 mRNA level in SW13 adrenal carcinoma cells (Ismail and Bateman 2009).

**Snail** EMT-related factors Snail1/2 activate Tbx3 expression in squamous cell carcinoma cells (Humtsoe et al. 2011).

**PLCε** Phospholipase C-ε (PLCε) can activate protein kinase C (PKCα/β). shPLCε decreased TBX3 mRNA and protein levels in human bladder cancer cells. (H. F. Du et al. 2014)

**CSF1** Colony-stimulating factor 1 (CSF1) is an important regulator of osteoclastogenesis, and CSF1 induces TBX3 expression. (Yao et al. 2014)

**TGF-β** B. Wang et al. 2012 showed TGF-β stimulation in primary human mammary epithelial cells (HMECs) induced TBX2 expression. Similarly, TBX3 is up-regulated by TGF-β1 in breast epithelial cells and keratinocytes (J. Li, M. S. Weinberg, et al. 2013), and mesangial cells (Wensing and Campos 2014). In breast cancer cells, TGF-β1 stimulation up-regulates TOPK, in turn TOPK induces TBX3 expression (Y.-J. Lee et al. 2019). However, a contradictory report showed that, in breast epithelial cells and mouse melanoma B16 cells, TGF-β1 reduced TBX2 mRNA and protein level by repressing the TBX2 promoter activity (J. Li, D. Ballim, et al. 2014).

## 7.2 Screen for TBX2 and TBX3 regulators

TBX2 and TBX3 are overexpressed in several types of cancer, including melanoma, breast and liver cancers (Juan Lu et al. 2010; Wansleben et al. 2014). Previous work has established that TBX2 and TBX3 are important in melanoma senescence bypass and invasion (Prince et al. 2004; Vance, Suzanne Carreira, et al. 2005;

Lingbeek et al. 2002; Peres and Prince 2013). As such, the identification of drugs that down-regulate them or translocate TBX2 or TBX3 to the cytoplasm might induce senescence or reduce melanoma invasion, which would be of therapeutic benefit. By contrast drugs that up-regulate TBX2 and/or TBX3 could also help us discover new upstream regulators, if their original targets are known. With the aim of finding drugs that can alter the protein level or sub-cellular localisation of TBX2 and/or TBX3, we used melanoma cell lines to carry out a high-throughput drug screen, in collaboration with Daniel Ebner's group from the Target Discovery Institute in University of Oxford.

### 7.2.1 Drug screen experiment

The initial aim was to use melanoma cell lines expressing TBX2 and/or TBX3 in a high content imaging screen of a Pharmakon 1600 compound FDA-approved drug library using immunofluorescence as a readout. As this would require specific anti-TBX2 and anti-TBX3 antibodies we undertook a preliminary screen of commercial and in-house antibodies for specificity in this assay.

At the time when the screen was being optimised (April 2017), we quickly tested eight TBX3 antibodies that were commercially available at that time (ab99302, Thermos Fisher 42-4800, sc-17871, ab89220, H000069-M03, MAB10089, 210606-002, sc-166625) and found that they were unlikely to be specific, as several siTBX3 we tested decreased TBX3 mRNA levels but the antibody staining remained unchanged. On the other hand, our homemade anti-TBX2 monoclonal antibody is specific but too dilute (1:50 dilution was required for a good dynamic window), and we did not have sufficient for the drug screen. Therefore, we chose to use 501 melanoma cell lines that ectopically express TBX2 or TBX3 with a FLAG tag, which were generated by a previous DPhil student in our group, Johannes Schmidt. The cell lines have a combined piggyBAC-Tet-on system and can express TBX2 or TBX3 upon

doxycycline administration. In order to have a good dynamic window, we titrated the doxycycline concentration and chose to use 30 ng/mL doxycycline for the screen (Figure 7.1.A). Although a screen based on FLAG-tagged TBX2 or TBX3 protein levels would not enable us to identify drugs that regulate TBX2/3 transcription, control of TBX2/3 post-translational modification was likely to be important.

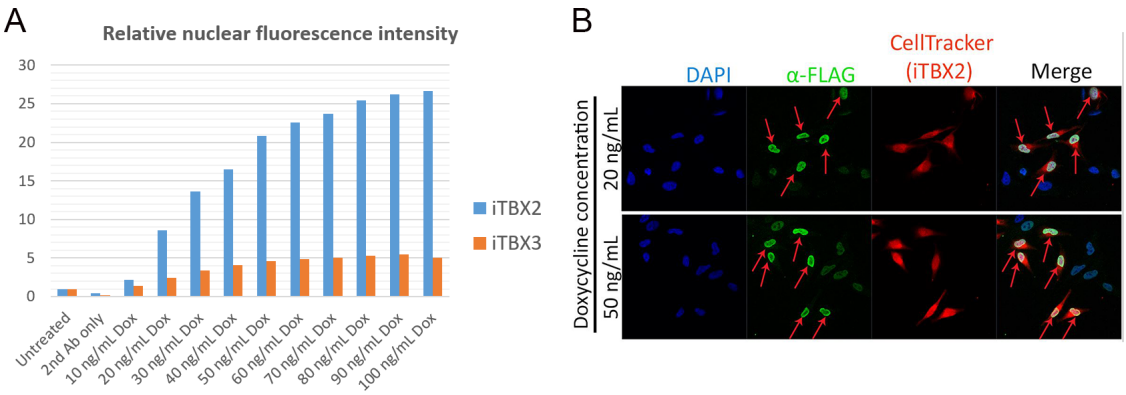



**Figure 7.1: Assay development for the drug screen** (A) TBX2 and TBX3 protein expression were induced with different concentration of doxycycline for 24 h. Data were averaged from 16 wells of cells in µclear 384-well plates, cell number > 16,000. Fluorescence intensity was measured by PE Operatta, data were analysed by Columbus. (B) iTBX2 was pre-stained with CellTracker Orange and co-cultured with iTBX3. Arrows: iTBX2 cells. Images were obtained from Zeiss 710. iTBX2: 501mel-iTBX2-FLAG, iTBX3: 501mel-iTBX3-FLAG.

To perform the screen, we first co-cultured 501mel-iTBX2-FLAG and 501mel-iTBX3-FLAG in the same well and induced TBX2 and TBX3 expression by doxycycline. The 501mel-iTBX2-FLAG cell line was pre-stained with Cell Tracker Orange, a live cell dye, which was used to differentiate the two cell lines during analysis (Figure 7.1.B). Second, cells were treated with different drugs at 10 µM for three time points – 4 h, 12 h and 24 h – in replicates. After treatment, cells were fixed with 4% PFA, permeabilised with 0.2% Triton-X100, the nuclei stained with DAPI, and TBX2/3 proteins detected with anti-FLAG antibody and Alexa488-conjugated secondary antibody. Finally, the fluorescence intensity and localisation were measured using a GE InCell Analyzer 6000.

In the analysis process, the readouts included cell number, fluorescence intensity

(FI) in the nucleus and in the whole cell, as well as the ratio of nuclear FI to cytoplasmic FI. We used cell number as the major indicator of cell health. We initially plated 400 cells in each well, and at the end of screen we used 250 cells as threshold for compound toxicity. Other indicators included nuclear area size and form factor. FI represents the protein level of TBX2 and TBX3, while the nucleus/cytoplasm ratio reveals changes in protein subcellular localisation. As we did not have a small molecule positive control that regulates the post-translational modification of TBX2/3, we used the Z-score to determine how many standard deviations an observation is above or below the mean and whether the difference is large enough, which is calculated as below, where Z = Z-score, $\chi$ = observed value, $\mu$ = mean of the sample, $\sigma$ = standard deviation of the sample.

$$Z = \frac{\chi - \mu}{\sigma}$$

### 7.2.2 Drug screen results – overview

From the screening, we found 42 hits that up-regulated TBX2 and 18 for TBX3 (Figure 7.2 A and B). This is determined by a Z-score of more than 3 and a cell number no less than 250, excluding random variations arising from cell death. Similarly, hits that down-regulate the target proteins are determined by a Z-score less than -2 and a cell number no less than 250. Of these there were 23 for TBX2 and 12 for TBX3 (Figure 7.2 A and B). Some 'hits' only affect either TBX2 or TBX3 but not both, which are of interest as their effects are specific and less likely to be an artefact affecting the Tet-on promoter that drives TBX2/3 expression (Figure 7.2.C). Some changes in protein subcellular localisation were also observed (Figure 7.2 D and E), however, further data analysis showed that there was no significant shift from the nucleus to the cytoplasm and the majority of protein remained nuclear and the change mainly arose from drug toxicity.

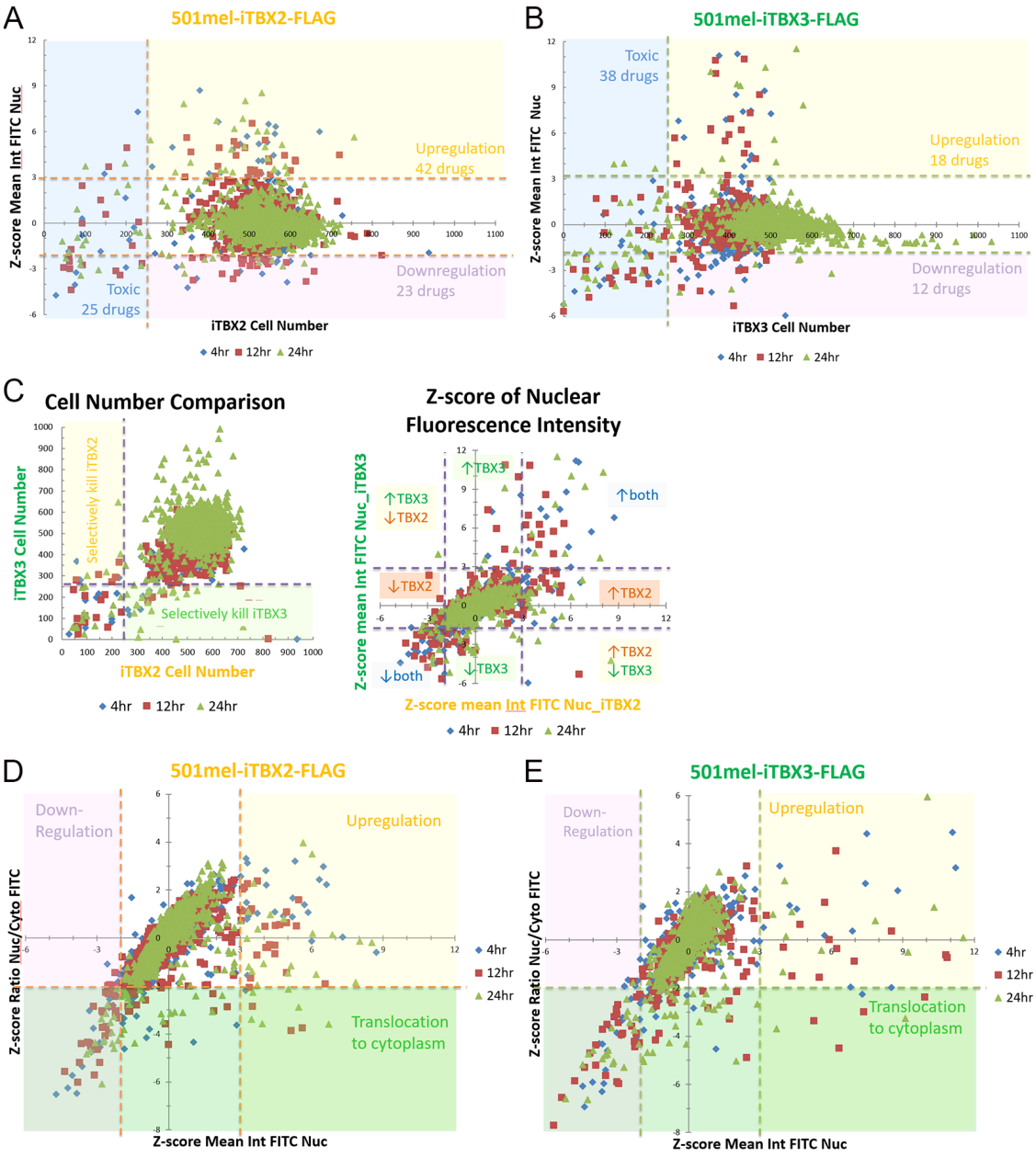



**Figure 7.2: Overview of the drug screen results** (A) and (B) Z-score of nuclear protein level was plotted against cell number for both cell lines. (C) Differential effects on cell number and nuclear protein level of the two cell lines were compared. (D) and (E) Z-score of nuclear/cytoplasmic protein ratio was plotted against Z-score of nuclear protein level.

## Drug screen hit compounds classification

Hit compounds were classified according to their known targets or function and these are summarised in Table 7.1. Known targets of the hits include calcineurin, cyclooxygenase, dopamine receptor, adrenergic receptor, androgen receptor, pro-

gesterone receptor, oestrogen receptor and retinoic acid receptor. We were not particularly interested in drugs that hit generic targets such as DNA or protein synthesis, or those that do not have known targets in mammalian cells, because it is not likely that they will have specific effects on TBX2 or TBX3.

| **Target or Function** | **Hits number** |
|---|---|
| inhibits DNA synthesis | 14 |
| inhibits protein synthesis | 9 |
| inhibits tubulin polymerisation or blocks microtubules | 10 |
| acts on bacterial cell wall | 4 |
| inhibits folic acid synthesis | 5 |
| inhibits cyclooxygenase (COX) | 3 |
| dopamine agonist/antagonist | 7 |
| α1-adrenergic receptor | 4 |
| androgen receptor | 4 |
| progesterone receptor | 3 |
| oestrogen receptor | 4 |
| $Na^+$-$K^+$ exchanging ATPase | 5 |
| retinoic acid receptor | 2 |
| inhibits calcineurin | 2 |
| ion channel blocker | 3 |
| blocks histamine H1-receptor | 6 |
| protein kinase C (iota/delta) | 2 |
| glucocorticoid receptor | 2 |
| cholinesterase | 3 |

**Table 7.1:** Known target and/or function of hit drugs

### Drug screen top hits

Among all the drugs that can potentially influence the protein levels of TBX2 and/or TBX3, we were particularly interested in the calcineurin inhibitors (CaNis), tacrolimus (TAC) and cyclosporin A (CsA), as their Z-scores are among the greatest and they do not have acute toxicity to melanoma cells (Figure 7.3.A).

Regarding the mechanism of action, since they decrease TBX2 and TBX3 protein levels, one hypothesis is that the calcineurin is protecting TBX2 and TBX3 protein from degradation. Calcineurin is a calcium-calmodulin-dependent phosphatase, and

its function is important in the immune response. For example, when transcription factor Nuclear factor of activated T-cells (NFAT) is dephosphorylated by calcineurin, it can translocate into nucleus and activate genes involved in immune response, such as *IL2* (Fantini et al. 2006). Previous work from our group found that there are PEST-like sequences in TBX2 and TBX3 (unpublished) likely to be implicated in their degradation. It has been shown in other proteins that phosphorylation of PEST sequence often contributes to a shorter protein half-life (Rechsteiner and Rogers 1996; García-Alai et al. 2006). If calcineurin were to dephosphorylate the PEST-like sequences, it might stabilise TBX2 and TBX3 proteins. When tacrolimus and cyclosporin A enter cells, they bind to the immunophilins cyclophilin and FKBP12 respectively, inhibiting the phosphatase activity of calcineurin (Fantini et al. 2006), which might result in phosphorylation of the PEST-like sequences and fast degradation of TBX2 and TBX3 proteins.

### 7.2.3 Drug screen validation

#### CaNis initial validation

To validate the HTS results using other assays, we undertook a time course and titration of the two CaNis using In-Cell ELISA (Cat. No. 62200). In-Cell ELISA is quantitative way to determine relative protein levels. The experiment was carried out according to the manufacturer's instructions: 501mel-iTBX2-FLAG and 501mel-iTBX3-FLAG cell lines were seeded separately in 96-well plates, and TBX2 and TBX3 expression was induced by adding 30 ng/mL doxycycline. After drug treatment, cells were fixed and permeabilised, then TBX2 and TBX3 proteins were probed with anti-FLAG antibody and a horseradish peroxidase (HRP)-conjugated secondary antibody. Then the TMB substrate was added to the wells and absorbance measured as a quantification of proteins. The results were normalised to cell number with a whole cell stain.

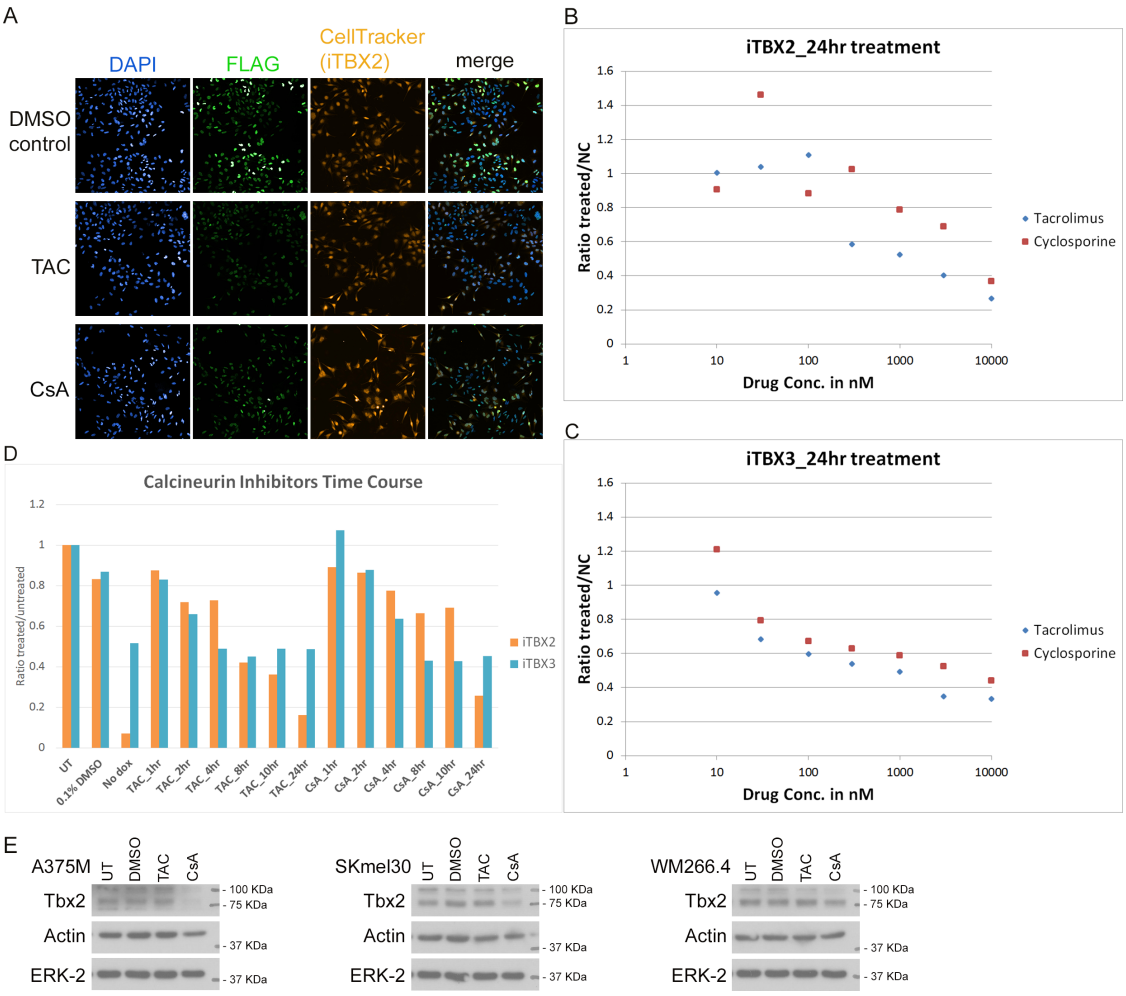



**Figure 7.3: Preliminary validation of Calcineurin inhibitors** (A) representative images from the drug screen. Cells were treated with DMSO or CaNis for 24 h. (B-D) TBX2 and TBX3 expression was induced by 30 ng/mL of doxycycline. After treatment, the relative amount of TBX2 or TBX3 protein was detected using InCell ELISA. Results were normalised by cell number and averaged from three 96-well plate wells. (B-C) iTBX2 and iTBX3 cells were treated with different concentration of CaNis for 24 h. (D) iTBX2 and iTBX3 cells were treated with 10 µM of CaNis for different time periods. iTBX2: 501mel-iTBX2-FLAG, iTBX3: 501mel-iTBX3-FLAG. (E) WB result from human melanoma cell lines treated with 10 µM of CaNis for 24 h; endogenous TBX2 protein was detected using HPA-validated antibody (Cat# HPA008586). Actin and ERK-2 were used as loading controls. TAC, tacrolimus; CsA, cyclosporin A

From the preliminary results (Figure 7.3 B and C), we can see that both tacrolimus and cyclosporin A exhibit a dose-dependent inhibition of TBX2 and TBX3 protein expression. While tacrolimus can decrease TBX2 at a lower concentration compared with cyclosporin A, the two CaNis have similar effects on TBX3. Longer treatment exhibits a greater inhibition on TBX2, whilst 8 h treatment is enough

to decrease TBX3 to a level compared to that with no doxycycline induction (Figure 7.3.D).

**Tacrolimus artefact**

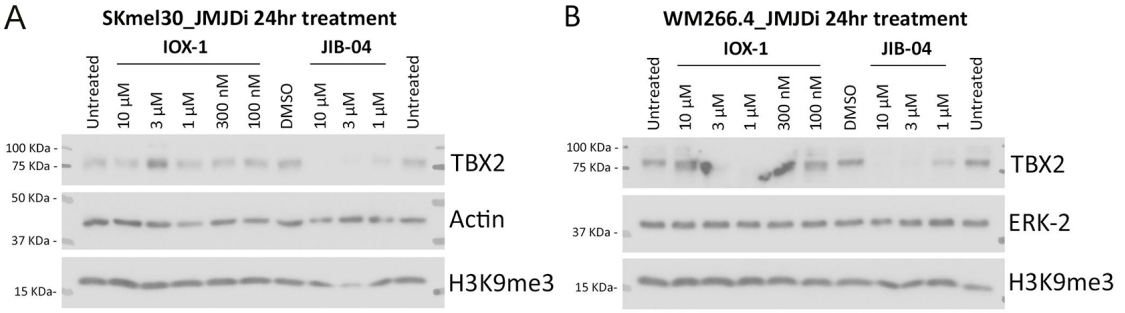
As the drug screen was done in cells that ectopically express TBX2 or TBX3, the hit compounds we found could be either affecting TBX2 and/or TBX3 protein stability or having 'off-target' effects on the Tet-on promoter. Unfortunately, it was shown by our collaborator Professor Sharon Prince's group that tacrolimus also decreased mRNA level of Cas9 that is under a Tet-on promoter (data not included), indicating that the CaNi may have generic effects on the Tet-on system and interrupt exogenous protein expression.

It is therefore essential that we validate the 'hits' on endogenous proteins. However, at the time when the drug screen was finished in 2017, the lack of reliable antibodies still posed a challenge to study changes in endogenous protein expression. Later a new TBX2 antibody (Cat# HPA008586), which was developed and validated by The Human Protein Atlas (https://www.proteinatlas.org), became commercially available in 2019. My WB experiment also showed that this antibody can detect changes in TBX2 protein level following TBX2 knockdown in human melanoma cell line (Figure 4.11.D).

Interestingly, while tacrolimus has no effect on TBX2, cyclosporin A decreased endogenous TBX2 protein level in a panel of human melanoma cell lines (Figure 7.3.E). Cells were treated with 10 µM of Calcineurin inhibitors for 24 h, which were the concentration and longest time point used in the initial screen. The effects of cyclosporin A varied among different cell lines — in A375M and SKmel30 melanoma cell lines the effects on TBX2 expression were evident, the effect in the WM266.4 line was less robust. Further experiments are needed to determine the concentration and treatment time for each cell line.

## 7.3 Epigenetic control of TBX2 expression

In senescent cells, the repressive H3K9me3 marks in the promoters of proliferation-related genes leads to transcriptional silencing and subsequent growth arrest. Y. Yu et al. 2018 reported that H3K9-active demethylases LSD1 and JMJD2C (KDM4C) disable oncogene-induced senescence and promote malignant transformation of $BRAF^{V600E}$-driven melanocytic naevi. Interestingly, they also showed JMJD2C silencing decreased TBX2 mRNA levels in WM266.4 melanoma cells. Two other reports showed JMJD2C knockdown decreased TBX3 mRNA level in mESCs, in which the mRNA of TBX2 was not detected (Das et al. 2014; Tomaz et al. 2017). In addition, ChIP-seq data show that JMJD2C binds to the *TBX2* promoter in myeloid leukemia cells and mESCs (Agger et al. 2016; Pedersen et al. 2014).




C

| $IC_{50}$ value | JARID1A | JMJD1A | JMJD2D | JMJD2E | JMJD2B | JMJD2A | JMJD3 | JMJD2C | KDM6A | KDM2A | PHF8 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| JIB-04 | 230 nM | | 290 nM | 340 nM | 435 nM | 445 nM | 855 nM | 1.1 μM | | | |
| IOX-1 | 25 μM | 170 nM | 200 nM | 300 nM | | 200 nM | 140 nM | 600 nM | 1.1 μM | 15.4 μM | 15 μM |

**Figure 7.4: Histone demethylases inhibitor reduces TBX2 expression** (A-B) WB result from human melanoma cell lines (A) SKmel30 and (B) WM266.4 cells treated with different concentration of IOX-1 or JIB-04 for 24 h (C) $IC_{50}$ values of IOX-1 or JIB-04 against different histone demethylases

Therefore I explored the possibility of using JMJD2C inhibitors to establish transcriptional repressive marks in the *TBX2* promoter and thus reduce TBX2 expression. As there is currently no specific inhibitor for JMJD2C, I experimented on a pan-selective inhibitor of Jumonji histone demethylases, JIB-04, and a broad-spectrum inhibitor of the JMJD2 family, IOX-1. Interestingly, while JIB-04 drastically decreased TBX2 expression in two human melanoma cell lines, IOX-1 did

not have any consistent effect (Figure 7.4.A-B). The difference between IOX-1 and JIB-04 suggests maybe the inhibition of TBX2 expression by JIB-04 is not mediated by JMJD2C, but other histone demethylases affected by JIB-04. Figure 7.4.C shows while IOX-1 is more sensitive for JMJD2 family, JIB-04 is highly selective of JARID1A (KDM5A), which specifically demethylates tri- and di-methylated H3K4. In addition, the concentration of JIB-04 I tested did not impact H3K9me3 level (Figure 7.4.A-B), further supporting the hypothesis that JIB-04 induced TBX2 repression is not mediated by JMJD2 family of histone demethylase.

## 7.4 Discussion

In contrast to the interest in upstream regulators of TBX2 and TBX3, few reports set out to pharmacologically inhibit the expression of TBX2 and TBX3 in cancer, where they are often overexpressed, and assess the potential therapeutic benefits.

One example is AKT inhibitor VIII. Peres, Mowla, et al. 2014 showed AKT3 phosphorylation at S720 increases the TBX3 protein stability. The authors also tested one small molecule — the AKT Inhibitor VIII, which is selective for AKT1/2 but not AKT3 — at a high concentration of 20 µM, and saw a decrease in TBX3 expression. Given the fact that no selective AKT3 inhibitor has been developed, it might be preferable to test some pan-AKT inhibitors with a relative low $IC_{50}$ for AKT3. Apart from this report, several other lines of evidence indicate that the PI3K-AKT signalling pathway is involved in the activation of TBX3 expression (Niwa et al. 2009; Ke et al. 2018; Ismail and Bateman 2009). In addition, a previous student in our group, Johannes Schmidt showed PI3K inhibitors can reduced the expression of both TBX2 and TBX3 in melanoma cells.

On the other hand, transcription factors like TBX2 and TBX3 are often considered to be "undruggable", and thus the lack of interest in developing small molecule compounds specifically targeting them. While there is no established

TBX2/3 inhibitor, Sahm et al. 2020 identified compounds that can bind the DNA-binding domain of TBX2 protein *in vitro* using a reverse-affinity approach. They found chromomycins A5 (CA5) and A6 (CA6) to interact with TBX2 *in vitro*, and reduce the growth of melanoma, breast cancer and rhabdomyosarcoma cell lines, with $IC_{50}$ values in the nM range. The cytotoxicity of CA5 is dependent on TBX2 expression, as silencing of TBX2 increased the sensitivity of 501mel melanoma cells, and overexpression of TBX2 deceased it. Regrettably, the authors did not expand on the mechanisms underlying if and how chromomycins affect TBX2 expression, protein stability or DNA-binding function.

Preliminary data indicate the calcineurin inhibitor cyclosporin A and the histone demethylase inhibitor JIB-04 reduce TBX2 expression in melanoma cells (Figure 7.3 and Figure 7.4). As cyclosporin A was discovered in a screen using ectopically expressed TBX2, it is likely the mechanism of inhibition is via destabilisation of TBX2 protein. To test this hypothesis, further experiments testing TBX2 protein half-life are needed. If cyclosporin A does affect TBX2 protein stability, the dephosphorylation site of calcineurin in TBX2 will be mapped using mutagenesis method. On the other hand, JIB-04 is likely to reduce TBX2 expression by epigenetic control of the *TBX2* promoter. This hypothesis too needs to be tested by detecting changes in mRNA level, followed by identification of histone-modifying factors that mediate the repression.

# 8

# Discussion

## 8.1 Mechanism of bifunctional transcription factors

Preliminary data from the Tbx2 ChIP-seq and RNA-seq following siTbx2 indicate that apart from the well-established transcription repressive function, Tbx2 may directly activate expression of a broad range of genes. Why some transcription factors possess a dual-function has not been extensively studied.

One example is the existence of multiple DNA-binding domains. Unique among the T-box transcription factors, Mga contains a single-exon T-box domain and a basic helix-loop-helix (bHLH) zipper domain (Hurlin, Steingrìmsson, et al. 1999). Mga represses transcription of reporter genes containing Brachyury-binding T-element, and is converted to an activator when heterodimerising with the small

bHLHZip protein Max to bind E-box motifs.

However, most transcription factors only have a single DNA-binding domain, therefore, alternate mechanisms might depend on the interacting factors or the local DNA structure and genomic features (Inukai et al. 2017). The importance of cofactors is manifest by the capability of Max itself to switch between transcriptional activator and repressor by recruiting cofactors with opposing effects. Max functions as a repressor when interacting with Mnt or Mxd1, but as an activator when forming a heterodimer with MYC (Hurlin and J. Huang 2006).

### 8.1.1 TBX2/3 can be a transactivator

Prevailing reports suggest a transcriptional repressor role for TBX2 and TBX3. For example, in cancer progression, TBX2 and TBX3 can repress tumour suppressors such as PTEN, $p21^{CIP1}$ and $p14^{Arf}$, and inhibit E-cadherin expression to promote invasion and metastasis.

Interestingly, several reports have pointed out a transcriptional activating role for TBX2 and TBX3. For instance, Krstic, Kolendowski, et al. 2019 showed that TBX3 promotes the transition from ductal carcinoma *in situ* (DCIS) to invasive breast cancer, by directly binding to the TSS of EMT-related gene *SLUG* and up-regulating its expression, though the authors did not identify the protein domain mediating the transcriptional activation. In addition, in mouse ESCs, Tbx3 directly activates Gata6 expression by binding to the *Gata6* promoter (R. Lu et al. 2010). Furthermore, Tbx3 up-regulates *JDP2* expression by directly binding to the *JDP2* promoter (Yao et al. 2014), and TBX2 directly binds to the *WNT3A* promoter and activates its transcription (Nandana et al. 2017).

Mechanistically, R. Lu et al. 2010 showed Tbx3 reduces the repressive H3K27me3 histone mark in the *Gata6* promoter, whereas Nandana et al. 2017 reported TBX2 increased the occupancy of RNA pol II in the *WNT3A* promoter. However, it was

not shown whether Tbx3 recruits H3K27-active demethylase, or whether TBX2 can directly recruit RNA pol II. While the above studies did not specify any domain in TBX2/3 protein that is mediating the transactivating function, Paxton et al. 2002 identified a weak activation domain within the T-box domain of Tbx2 that can activate promoters containing multiple T-elements. In addition, Butz et al. 2004 reported they found an activation domain within the N-terminus of TBX2. However the evidence provided is not so convincing as their model fuses the N-terminus of TBX2 to the VP16 transactivation domain to show up-regulation of a reporter.

### 8.1.2 TBX2/3 function depends on cofactors

Interestingly, the DNA-binding specificity and transcriptional regulatory function of TBX2 and TBX3 sometimes depend on the interacting cofactors.

For example, in mouse embryonic forelimbs, 62% of Tbx3-bound regions are also occupied by Bcl9, as shown in ChIP-seq data (Zimmerli et al. 2020). Notably, enriched motifs in the Tbx3-Bcl9 common binding loci include the TCF/LEF and homeobox TF consensus sequences, but not the T-element; on the other hand, T-element motif enrichment was only detected in the Tbx3 exclusive peaks. In addition, Tbx3 and Bcl9 co-occupancy was detected at many Wnt-responsive-elements (WREs) within known Wnt target genes, with luciferase assay showing that under basal condition Tbx3 represses WREs, and this turned into activation when the Wnt pathway is switched on. Deletion of β-catenin or TCF abolishes the effect of Tbx3 on WREs, supporting the notion that the binding of Tbx3 to these sites relies on cofactors.

A similar dependence on cofactors was also found for TBX2. D'Costa et al. 2014 showed that TBX2 represses the *CST6* proximal promoter. Although the two putative TBX2 binding sites (-857 and -172) are dispensable for the repression of *CST6* promoter by TBX2, silencing of EGR1, an established TBX2

interactor in breast cancer, abrogated this repression, leading to increased CST6 expression. The same was reported for TBX2-mediated repression of the proximal *NDRG1* promoter, where no putative T-element was found (Redmond et al. 2010). Instead, TBX2 bound via a GCGGGGGCG consensus binding site for EGR1. Furthermore, EGR1 is required for TBX2's binding to the *NDRG1* proximal promoter as shown in ChIP-PCR experiments.

Considering the above examples, we speculate the enrichment of the E-box motif in Tbx2 binding sites might involve transcription factors containing basic helix-loop-helix (bHLH) domain that binds E-box motifs. This is supported by the preliminary results from LacO/LacI tethering assay, which showed Tbx2 interacts with Usf2, however the potential cooperative DNA binding has yet to be established.

## 8.2 Targeting TBX2/3 in cancer treatment

### 8.2.1 Cancer hallmarks and therapies

Based on the hallmarks of cancer proposed by Hanahan and Weinberg, many therapies have been developed and delivered clinical efficacy.

The first generation of anti-cancer targeted therapies, including small molecule inhibitors such as Imatinib, a BCR-ABL inhibitor used in chronic myeloid leukemia (Mughal and Goldman 2006), Vemurafenib and Dabrafenib, $BRAF^{V600E}$ inhibitors used in melanoma (Chapman et al. 2011; McArthur et al. 2014), Trametinib , a MEK inhibitor used in melanoma (Keith T Flaherty et al. 2012a), Gefitinib , an epidermal growth factor receptor (EGFR) inhibitor used in non-small cell lung cancer (NSCLC) (Nurwidya et al. 2016), as well as monoclonal antibodies such as Trastuzumab , an antibody targeting human epidermal growth factor receptor 2 (HER2) used in breast cancer (Garnock-Jones et al. 2010), are targeting the pro-proliferative oncogenic pathways.

There are also drugs suppressing irregular cell growth, such as Palbociclib, Ribociclib and Abemaciclib, which are inhibitors of CDK4 and CDK6, serine-threonine kinases that regulate the cell cycle progression (Serra et al. 2019); some drugs reactivate apoptosis, such as Bcl-2 inhibitors, Navitoclax (ABT-263) and Venetoclax (ABT-199), used in chronic lymphocytic leukemia (CLL); the replicative immortality is also a target for cancer treatment, for instance, Tertomotide (GV1001), a peptide derived from a reverse-transcriptase subunit of telomerase (hTERT) and developed as a vaccine against various cancers, including NSCLC, melanoma, hepatocellular carcinoma (H. Kim et al. 2016).

Angiogenesis inhibitor marks another line of anti-cancer treatments, with one of the successful efforts being the VEGF-A-targeting monoclonal antibody Bevacizumab, which has been approved for treatment of metastatic colorectal cancer and breast cancer, NSCLC, glioblastoma, renal cell carcinoma, ovarian cancer and cervical cancer (Garcia et al. 2020).

The DNA damage repair deficiency in cancer cells that gives rise to genome instability can also be exploited for cancer treatment. Based on synthetic lethal concept, PARP inhibitors, Olaparib, Rucaparib, Niraparib, and Talazoparib, have been developed to target BRCA1/2 mutated tumour cells (Yi et al. 2019).

Furthermore, immunotherapy opened a new era for cancer therapy, with immune checkpoint inhibitors (ICIs) being the most well-characterised ones. Ipilimumab, which blocks CTLA-4, was the first approved ICI for treatment of metastatic melanoma; this was followed by development of Pembrolizumab and Nivolumab, which target PD-1, as well as Atezolizumab and Durvalumab, which target PD-L1, and they have been authorised for treating NSCLC, melanoma, Hodgkin's lymphoma and renal cell carcinoma (Darvin et al. 2018; Havel et al. 2019).

### 8.2.2 Inhibiting TBX2/3 for cancer treatment

The cancer therapies described above indicate that the hallmarks may indeed serve as a roadmap to the ultimate goal of treating cancer. However, these therapies are only focused on a certain aspect of the big picture. It is anticipated that future cancer treatments will be multi-scale, systematic and interdisciplinary. We can speculate that targeting different aspects of the hallmarks will provide better clinical efficacy than targeting solely one. Therefore it is important to study the fundamentals of cancer development, including the mechanisms of senescence bypass, metabolic reprogramming and cell death resistance, in hope of finding new therapeutic strategies and bringing synergy to existing ones.

TBX2 and TBX3 transcription factors are crucial for embryonic development and they have been implicated in tumourigenesis. TBX2 is overexpressed in several types of cancer including breast, pancreatic, colorectal, cervical, ovarian, endometrial cancer, melanoma and rhabdomyosarcoma; TBX3 overexpression has been observed in breast, pancreatic, lung, ovarian, liver, head and neck cancers and melanoma (Wansleben et al. 2014).

TBX2 and TBX3 contribute to melanoma progression by facilitating senescence bypass and promoting invasiveness, and disruption of TBX2 function in melanoma cells can induce senescence (Vance, Suzanne Carreira, et al. 2005). Mechanistically, TBX2 and TBX3 can directly repress human $p14^{Arf}$ and mouse $p19^{Arf}$ (*CDKN2A*) (Jacobs et al. 2000; Lingbeek et al. 2002; Brummelkamp et al. 2002), which inhibit the p53 repressor MDM2 E3 ubiquitin ligase thus facilitating p53 activation. In addition, TBX2 and TBX3 bind to and repress the promoter of the CDK2 inhibitor $p21^{CIP1}$ (*CDKN1A*) (Tarryn Willmer, Hare, et al. 2016; Vance, Suzanne Carreira, et al. 2005). By repressing $p21^{CIP1}$ expression, TBX2/3 can reduce CDK2-mediated repression of Rb1, an inhibitor of E2F transcription factors that activate S-phase genes (Sharpless and Sherr 2015).

Considering the important role of TBX2 and TBX3 in melanoma, pharmacological inhibition of them might serve as a therapeutic strategy. For this purpose, a high-throughput drug screen was conducted to identify FDA-approved drugs that inhibit TBX2/3 expression and thus may be re-purposed for melanoma treatment. Preliminary data indicate the calcineurin inhibitor cyclosporin A and the histone demethylase inhibitor JIB-04 reduce TBX2 expression in melanoma cells (Figure 7.3 and Figure 7.4). Further work is needed to elucidate the mechanisms of inhibition — whether it is repressing transcription or destabilising the protein — as outlined in Chapter 7. Here I discuss the prospective therapeutic benefits and potential risks of utilising cyclosporin A and JIB-04 in treating melanoma.

Calcineurin and its downstream signalling pathways are activated in colorectal, breast, lung, prostate, bladder, ovarian, pancreatic, and liver cancer, as well as glioblastoma, melanoma and leukemia, and calcineurin inhibitors (CaNis) exhibit anti-tumour effects in lymphoma, leukemia, bladder cancer, prostate cancer and breast cancer, using both *in vitro* cell experiment as well *in vivo* mouse model (Brun and Godbout 2016). As CaNis are widely used as immunosuppressants after organ transplantation, it is not surprising to find evidence for a positive correlation between CaNis and increased cancer risk (Hojo et al. 1999; Datta et al. 2009). However, a recent post-marketing surveillance study, involving more than 93,000 patients, showed that long-term topical use of CaNis did not increase keratinocyte carcinoma risk (Asgari et al. 2020). In addition, the toxicity of CaNis was mainly related with systemic use of CaNis in organ transplant recipients (Azzi et al. 2013). As topical calcineurin inhibitors are licensed for the treatment of atopic dermatitis, application of such formula to pre-malignant lesions or early stages of melanoma might complement surgical removal.

Up-regulation of the Jumonji domain-containing histone demethylases (KDMs) have been implicated in tumour progression, and small molecule inhibitors of KDMs have been developed and tested for cancer treatment (Sainathan et al. 2015). JIB-04

was identified in a screen for small molecules that specifically inhibit the activity of the Jumonji family of histone demethylases (L. Wang et al. 2013). The authors showed that in prostate or lung cancer cells, but not in patient-matched normal cells, JIB-04 induces apoptosis and blocks viability, while in mice, JIB-04 reduces tumour burden and prolongs survival. In addition, breast cancer patients with high levels of Jumonji demethylases have significantly lower survival. Later, M. S. Kim et al. 2018 showed JIB-04 significantly attenuated self-renewal and stemness of colorectal CSCs, where CD133 was used as CSCs marker, as evidenced by reduction of CSC tumoursphere formation, relapse, invasion, and migration *in vitro*. In addition, JIB-04-treated colorectal cancer cells showed reduced tumourigenic activity *in vivo*. Overall, the effects of JIB-04 were equal to or greater than those of salinomycin, a known anti-colorectal CSC drug, and JIB-04 used a much lower concentration compared with that of salinomycin. Furthermore, JIB-04 increased DNA damage in Ewing Sarcoma cells and thus potently inhibits their growth and viability, and also impairs tumor xenograft growth (Parrish et al. 2018). Effects on histone methylation vary among cell lines, with most cell lines exhibiting increased total H3K27me3 levels, and some increased H3K4me3 and H3K9me3.

Mechanistically, Parrish et al. 2018 showed JIB-04 treatment widely alters expression of oncogenic and tumour suppressive pathways, including down-regulation of known oncogenic members of the Homeobox B and D clusters. JIB-04 also disrupts the expression signature EWS/Fli1, which is the most common fusion oncoprotein in Ewing sarcoma. Regarding histone methylation status, despite variance among cell lines,increase in total H3K27me3, H3K4me3 and H3K9me3 levels were observed. In addition, M. S. Kim et al. 2018 showed JIB-04 inhibits Wnt/β-catenin-regulated target genes associated with colorectal CSC function. It is not surprising that JIB-04, as a pan-selective inhibitor of histone demethylases, affects the expression different genes in different tumours. It is noteworthy that these studies reported different mechanisms of action underpinning the anti-cancer properties of JIB-04.

Therefore, if JIB-04 reduces melanoma cell growth and whether TBX2 is a key mediator remain to be determined.

Moreover, recent advances in siRNA delivery through lipid nanoparticle (LNP) technology are inspiring, should small compounds specifically targeting TBX2/3 prove difficult to discover, the potential of gene therapy is worthy of note.

## 8.3 Limitations and future perspectives

### 8.3.1 Lack of appropriate tools

There are some parts of my DPhil project that were limited by lack of appropriate tools. To begin with, the drug screen project was a collaboration with Professor Sharon Prince's group in 2017. At that time, we could not find suitable antibodies for the screen — our home-made TBX2 antibody was not sufficient for the high-throughput screen, whereas commercial antibodies we tested seem to have a narrow detection window for TBX2 and high background staining for TBX3 and thus could not detect an decrease in TBX3 protein levels. Therefore, we chose to use melanoma cell lines with ectopically expressed, FLAG-tagged TBX2 or TBX3, which are not optimal as drugs that regulate the transcription of TBX2/3 are omitted. Later only in 2018 and 2019 did I manage to tag endogenous TBX2 protein with a 3xHA tag in the IGR37 human melanoma cell line and in mouse B16 melanoma cells using CRISPR-Cas9 strategy. Though repeating the drug screen using these cell lines would add the missing piece of transcriptional regulators, the high cost of screening 1,600 drugs imposes a barrier.

Later, a new TBX2 antibody which was developed and validated by The Human Protein Atlas (HPA) became commercially available in 2019. This antibody works well in human melanoma cells, and it does detect TBX2 decrease when cells were transfected with siTBX2 (Figure 4.11.D). This enabled us to further validate the

initial hits identified in the drug screen (Figure 7.3.E).

However, I tested a panel of commercial TBX3 antibodies, but none could detect a decrease in TBX3 protein level after siTBX3 transfection in WB. To rule out the possibility of poor transfection efficiency or siTBX3 specificity, I performed qRT-PCR and could confirm that the siTBX3 decreased mRNA level of TBX3 (data not shown). The lack of reliable TBX3 antibody hindered the validation of drug screen hits, as well as RNA-seq experiment assessing the transcriptome of melanoma cells after TBX3 silencing, which is crucial for identifying TBX3 direct target genes.

### 8.3.2 Impact of the Covid-19 pandemic

Owing to the Covid-19 pandemic, I could not access the research building for 4 months from March to July in 2020. This disturbed my experimental plans and thesis writing schedule. For example, I obtained the Tbx2 ChIP-seq data in late January 2020, and completed analyses in February, but did not have time to validate any of binding sites using ChIP-qPCR or luciferase assay, as producing three biological replicates for ChIP-qPCR is time-consuming and luciferase assay conditions need to be optimised. These two experiments are important for validating the potential direct targets of Tbx2 found in ChIP-seq and RNA-seq experiments.

Likewise, after I tested the new HPA-validated TBX2 antibody at the end of 2019, there was only time for some quick experiments to assess the impact of calcineurin inhibitors and histone demethylase inhibitors on TBX2 protein levels. Experiments that are indispensable to confirm their regulation of TBX2 expression include assessment of TBX2 mRNA level and protein half-life. Further experiments will focus on identifying the genomic region mediating the transcription control by JIB-04 and the protein domain that is dephosphorylated and stabilised by calcineurin.

### 8.3.3 Future perspectives

Apart from the above missing experiments that are essential for validation of E2F1 as a direct target gene of TBX2, understanding mechanisms of cyclosporin A and JIB-04 regulation of TBX2 expression, and identification of novel direct target genes of TBX3, additional assays are also needed to explore other parts of my project.

First, to test if TBX2 interacts with USF2 to bind E-box sequence, *in vitro* cooperation assays such as EMSA are needed. This could be coupled with USF2 silencing experiments to assess whether USF2 is required for TBX2 regulation of key genes and whether the interaction converts TBX2 to a transcriptional activator. Similarly, to validate the interaction between TBX2 with PRC1, other components of the non-canonical KDM2B-PRC1 will also be tested in the LacO/LacI tethering assay. Further experiments will test whether in ChIP assays depletion of TBX2 abrogates recruitment of PRC1 components to target genes, and in turn, whether depletion of PRC1 abolishes the transcriptional repression of direct TBX2 target genes. The interaction with PRC1 would add to our understanding of how TBX2 inhibits target gene expression by establishing repressive histone marks, complementing the known mechanisms such as recruiting HDAC and HP1 to the promoter of target genes (Vance, Suzanne Carreira, et al. 2005; Zhu et al. 2015; Crawford et al. 2019), while the interaction with bHLH-containing transcription factor USF2 would unravel the mechanistic basis of E-box binding capacity and transcriptional activation function of TBX2.

Further work could expand on integrating the novel findings of TBX2 into the big picture of melanoma progression. E2F1, a transcription factor initially identified as a cell cycle regulator, has emerged as a multifunctional player in control of DNA damage response and metabolism (Denechaud et al. 2017). E2F1 activates the transcription of genes involved in progression into lat G1/S phase, such as cylins and CDKs (Roworth et al. 2015). Other target genes are involved in

DNA replication, which include MCM proteins and topoisomerase II, or nucleotide synthesis such as dihydrofolate reductase and thymidine kinase. In addition, E2F1 can facilitate the metabolic reprogramming in cancer cells by enhancing glycolysis through activation of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase which synthesises a glycolysis stimulator, and at the same time suppressing oxidative metabolism via promoting expression of pyruvate dehydrogenase kinases which decrease glucose oxidation (Denechaud et al. 2017). In line with TBX2's role in oncogene-induced senescence, E2F1 has also been reported to promote bypass of senescence (Rowland et al. 2002; Dimri et al. 2000), supporting the notion that E2F1 is an important mediator of TBX2 function. Therefore, future experiments should also test if pharmacological inhibition of TBX2, using cyclosporin A or JIB-04, can also repress E2F1 expression and subsequently arrest tumour cell growth, recapitulating the observations in TBX2 knockdown experiments.

# Appendices

The full list of drug screen hits is attached as additional files to the the thesis.

# Works Cited


Abrahams, Amaal, Shaheen Mowla, et al. (2008). "UV-mediated Regulation of the Anti-senescence Factor Tbx2". In: *Journal of Biological Chemistry* 283.4, pp. 2223–2230.

Abrahams, Amaal, M. Iqbal Parker, and Sharon Prince (2010). "The T-box transcription factor Tbx2: Its role in development and possible implication in cancer". In: *IUBMB Life* 62.2, pp. 92–102.

Acosta, Juan and Jesús Gil (2012). "Senescence: a new weapon for cancer therapy". In: *Trends in Cell Biology* 22.4, pp. 211–219.

Agger, Karl et al. (2016). "Jmjd2/Kdm4 demethylases are required for expression of Il3ra and survival of acute myeloid leukemia cells". In: *Genes & Development* 30.11, pp. 1278–1288.

Amir, Sumaira et al. (2016). "Regulation of the T-box transcription factor Tbx3 by the tumour suppressor microRNA-206 in breast cancer". In: *British Journal of Cancer* 114.10, p. 1125.

Anbari, Kevin K. et al. (1997). "Melanoma of unknown primary site". In: *Cancer* 79.9, pp. 1816–1821.

Andtbacka, Robert H.I. et al. (2015). "Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma". In: *Journal of Clinical Oncology* 33.25, pp. 2780–2788.

Asgari, Maryam M. et al. (2020). "Association Between Topical Calcineurin Inhibitor Use and Keratinocyte Carcinoma Risk Among Adults With Atopic Dermatitis". In: *JAMA Dermatology* 156.10.

Atkins, Michael B et al. (1999). "High-Dose Recombinant Interleukin 2 Therapy for Patients With Metastatic Melanoma: Analysis of 270 Patients Treated Between 1985 and 1993". In: *Journal of Clinical Oncology* 17.7, pp. 2105–2105.

Aydoğdu, Nurullah et al. (2018). "TBX2 and TBX3 act downstream of canonical WNT signaling in patterning and differentiation of the mouse ureteric mesenchyme". In: *Development* 145.23, dev171827.

Azzi, Jamil R., Mohamed H. Sayegh, and Samir G. Mallat (2013). "Calcineurin Inhibitors: 40 Years Later, Can't Live Without . . . " In: *The Journal of Immunology* 191.12, pp. 5785–5791.

Balch, Charles M. et al. (2009). "Final Version of 2009 AJCC Melanoma Staging and Classification". In: *Journal of Clinical Oncology* 27.36, pp. 6199–6206.

Ballim, Reyna et al. (2012). "The ulnar-mammary syndrome gene, Tbx3, is a direct target of the retinoic acid signaling pathway, which regulates its expression during mouse limb development". In: *Molecular Biology of the Cell* 23.12, pp. 2362–2372.

Bamshad, Michael et al. (1997). "Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome". In: *Nature Genetics* 16.3, pp. 311–315.

Bannister, Andrew J and Tony Kouzarides (2011). "Regulation of chromatin by histone modifications". In: *Cell Research* 21.3, pp. 381–395.

Benci, Joseph L et al. (2016). "Tumor Interferon Signaling Regulates a Multigenic Resistance Program to Immune Checkpoint Blockade." In: *Cell* 167.6, 1540–1554.e12.

Bilican, Bilada and Colin R Goding (2006). "Cell cycle regulation of the T-box transcription factor tbx2". In: *Experimental Cell Research* 312.12, pp. 2358–2366.

Boogerd, C J J et al. (2011). "Sox4 mediates Tbx3 transcriptional regulation of the gap junction protein Cx43." In: *Cellular and Molecular Life Sciences : CMLS* 68.23, pp. 3949–61.

Boogerd, Kees-Jan et al. (2008). "Msx1 and Msx2 are functional interacting partners of T-box factors in the regulation of Connexin43." In: *Cardiovascular Research* 78.3, pp. 485–93.

Boskovic, Goran and Richard M Niles (2004). "T-box binding protein type two (TBX2) is an immediate early gene target in retinoic-acid-treated B16 murine melanoma cells". In: *Experimental Cell Research* 295.2, pp. 281–289.

Boyd, Suzanah C et al. (2012). "Oncogenic B-RAF(V600E) signaling induces the T-Box3 transcriptional repressor to repress E-cadherin and enhance melanoma cell invasion." In: *The Journal of Investigative Dermatology* 133.5, pp. 1269–77.

Brenner, Michaela and Vincent J. Hearing (2008). "The Protective Role of Melanin Against UV Damage in Human Skin†". In: *Photochemistry and Photobiology* 84.3, pp. 539–549.

Brummelkamp, Thijn R et al. (2002). "TBX-3, the Gene Mutated in Ulnar-Mammary Syndrome, Is a Negative Regulator of p19 ARF and Inhibits Senescence". In: *Journal of Biological Chemistry* 277.8, pp. 6567–6572.

Brun, Miranda and Roseline Godbout (2016). "Activation of calcineurin in cancer: many paths, one hub". In: *Translational Cancer Research* 5.S3, S497–S506.

Burgucu, Durmus et al. (2012). "Tbx3 represses PTEN and is over-expressed in head and neck squamous cell carcinoma." In: *BMC cancer* 12.1, p. 481.

Butz, Nataliya V, Christine E Campbell, and Richard M Gronostajski (2004). "Differential target gene activation by TBX2 and TBX2VP16: evidence for activation domain-dependent modulation of gene target specificity". In: *Gene* 342.1, pp. 67–76.

Campbell, Kirsteen J. and Stephen W. G. Tait (2018). "Targeting BCL-2 regulated apoptosis in cancer". In: *Open Biology* 8.5, p. 180002.

Canon, Jude et al. (2019). "The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity". In: *Nature* 575.7781, pp. 217–223.

Carlson, Hanqian, Sara Ota, Christine E. Campbell, et al. (2001). "A dominant repression domain in Tbx3 mediates transcriptional repression and cell immortalization: relevance to mutations in Tbx3 that cause ulnar-mammary syndrome". In: *Human Molecular Genetics* 10.21, pp. 2403–2413.

Carlson, Hanqian, Sara Ota, Youngsup Song, et al. (2002). "Tbx3 impinges on the p53 pathway to suppress apoptosis, facilitate cell transformation and block myogenic differentiation". In: *Oncogene* 21.24, pp. 3827–3835.

Carreira, S et al. (1998). "Brachyury-Related Transcription Factor Tbx2 and Repression of the Melanocyte-Specific TRP-1 Promoter". In: *Molecular and Cellular Biology* 18.9, pp. 5099–5108.

Carreira, Suzanne, Baoguo Liu, and Colin R Goding (2000). "The gene encoding the T-box factor Tbx2 is a target for the Microphthalmia-associated transcription factor in melanocytes". In: *Journal of Biological Chemistry* 275.29, pp. 21920–21927.

Carson, Karen F et al. (1996). "Nodal Nevi and Cutaneous Melanomas". In: *The American Journal of Surgical Pathology* 20.7, pp. 834–840.

Chapman, Paul B et al. (2011). "Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation". In: *The New England Journal of Medicine* 364.26, pp. 2507–2516.

CHEN, FEIFEI et al. (2016). "The emerging role of RUNX3 in cancer metastasis (Review)". In: *Oncology Reports* 35.3, pp. 1227–1236.

Chen, Jiji et al. (2014). "Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells". In: *Cell* 156.6, pp. 1274–1285.

Cho, Kyoung-Won et al. (2006). "Molecular interactions between Tbx3 and Bmp4 and a model for dorsoventral positioning of mammary gland development". In: *Proceedings of the National Academy of Sciences* 103.45, pp. 16788–16793.

Chohan, Tahir et al. (2014). "Cyclin-Dependent Kinase-2 as a Target for Cancer Therapy: Progress in the Development of CDK2 Inhibitors as Anti-Cancer Agents". In: *Current Medicinal Chemistry* 22.2, pp. 237–263.

Christoffels, Vincent M et al. (2004). "T-box transcription factor Tbx2 represses differentiation and formation of the cardiac chambers". In: *Developmental Dynamics* 229.4, pp. 763–770.

Cioffi, Michele et al. (2015). "MiR-93 Controls Adiposity via Inhibition of Sirt7 and Tbx3". In: *Cell Reports* 12.10, pp. 1594–1605.

Clapier, Cedric R. and Bradley R. Cairns (2009). "The Biology of Chromatin Remodeling Complexes". In: *Annual Review of Biochemistry* 78.1, pp. 273–304.

Colak, Selcuk and Peter ten Dijke (2017). "Targeting TGF-β Signaling in Cancer". In: *Trends in Cancer* 3.1, pp. 56–71.

Coll, Miquel, Jonathan G Seidman, and Christoph W Müller (2002). "Structure of the DNA-Bound T-Box Domain of Human TBX3, a Transcription Factor Responsible for Ulnar-Mammary Syndrome". In: *Structure* 10.3, pp. 343–356.

Costin, Gertrude-E. and Vincent J. Hearing (2007). "Human skin pigmentation: melanocytes modulate skin color in response to stress". In: *The FASEB Journal* 21.4, pp. 976–994.

Courtois-Cox, S, SL Jones, and K Cichowski (2008). "Many roads lead to oncogene-induced senescence". In: *Oncogene* 27.20, p. 1210950.

Crawford, N T et al. (2019). "TBX2 interacts with heterochromatin protein 1 to recruit a novel repression complex to EGR1-targeted promoters to drive the proliferation of breast cancer cells." In: *Oncogene* 38.31, pp. 5971–5986.

Curto, M and A I McClatchey (2008). "Nf2/Merlin: a coordinator of receptor signalling and intercellular contact". In: *British Journal of Cancer* 98.2, pp. 256–262.

Czyz, Malgorzata (2018). "HGF/c-MET Signaling in Melanocytes and Melanoma". In: *International Journal of Molecular Sciences* 19.12, p. 3844.

D'Costa, Zenobia C. et al. (2014). "TBX2 represses CST6 resulting in uncontrolled legumain activity to sustain breast cancer proliferation: a novel cancer-selective target pathway with therapeutic opportunities." In: *Oncotarget* 5.6, pp. 1609–1620.

Darvin, Pramod et al. (2018). "Immune checkpoint inhibitors: recent progress and potential biomarkers". In: *Experimental & Molecular Medicine* 50.12, pp. 1–11.

Das, Partha et al. (2014). "Distinct and Combinatorial Functions of Jmjd2b/Kdm4b and Jmjd2c/Kdm4c in Mouse Embryonic Stem Cell Identity". In: *Molecular Cell* 53.1, pp. 32–48.

Datta, Dipak et al. (2009). "Calcineurin Inhibitors Activate the Proto-Oncogene Ras and Promote Protumorigenic Signals in Renal Cancer Cells". In: *Cancer Research* 69.23, pp. 8902–8909.

Davar, Diwakar et al. (2017). "High-dose interleukin-2 (HD IL-2) for advanced melanoma: a single center experience from the University of Pittsburgh Cancer Institute". In: *Journal for ImmunoTherapy of Cancer* 5.1, p. 74.

Davenport, Todd G., Loydie A. Jerome-Majewska, and Virginia E. Papaioannou (2003). "Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in human ulnar mammary syndrome". In: *Development* 130.10, pp. 2263–2273.

Davis, Lauren E., Sara C. Shalin, and Alan J. Tackett (2019). "Current state of melanoma diagnosis and treatment". In: *Cancer Biology & Therapy* 20.11, pp. 1–14.

DeBerardinis, Ralph J. and Navdeep S. Chandel (2020). "We need to talk about the Warburg effect". In: *Nature Metabolism* 2.2, pp. 127–129.

Decaesteker, Bieke et al. (2018). "TBX2 is a neuroblastoma core regulatory circuitry component enhancing MYCN/FOXM1 reactivation of DREAM targets." In: *Nature Communications* 9.1, p. 4866.

Demay, Florence et al. (2007). "T-box factors: targeting to chromatin and interaction with the histone H3 N-terminal tail". In: *Pigment Cell Research* 20.4, pp. 279–287.

Denechaud, Pierre-Damien, Lluis Fajas, and Albert Giralt (2017). "E2F1, a Novel Regulator of Metabolism". In: *Frontiers in Endocrinology* 8, p. 311.

Dimri, Goberdhan P. et al. (2000). "Regulation of a Senescence Checkpoint Response by the E2F1 Transcription Factor and p14ARF Tumor Suppressor". In: *Molecular and Cellular Biology* 20.1, pp. 273–285.

Dong, Liang et al. (2018). "Novel HDAC5-interacting motifs of Tbx3 are essential for the suppression of E-cadherin expression and for the promotion of metastasis in hepatocellular carcinoma". In: *Signal Transduction and Targeted Therapy* 3.1, p. 22.

Dranoff, Glenn (2002). "GM-CSF-based cancer vaccines". In: *Immunological Reviews* 188.1, pp. 147–154.

Du, Hong Fei et al. (2014). "A new PKCα/β/TBX3/E-cadherin pathway is involved in PLCε-regulated invasion and migration in human bladder cancer cells". In: *Cellular Signalling* 26.3, pp. 580–593.

Du, Wen-Liang et al. (2017). "Effect of silencing the T-Box transcription factor TBX2 in prostate cancer PC3 and LNCaP cells". In: *Molecular Medicine Reports* 16.5, pp. 6050–6058.

Eblaghie, Maxwell C et al. (2004). "Interactions between FGF and Wnt signals and Tbx3 gene expression in mammary gland initiation in mouse embryos". In: *Journal of Anatomy* 205.1, pp. 1–13.

Ecco, Gabriela, Michael Imbeault, and Didier Trono (2017). "KRAB zinc finger proteins". In: *Development* 144.15, pp. 2719–2729.

Emechebe, Uchenna et al. (2016). "T-box3 is a ciliary protein and regulates stability of the Gli3 transcription factor to control digit number." In: *eLife* 5, e07897.

Etcheverry, Amandine et al. (2010). "DNA methylation in glioblastoma: impact on gene expression and clinical outcome." In: *BMC genomics* 11.1, p. 701.

Falletta, Paola et al. (2017). "Translation reprogramming is an evolutionarily conserved driver of phenotypic plasticity and therapeutic resistance in melanoma". In: *Genes & Development*.

Fan, Weiwei et al. (2004). "TBX3 and Its Isoform TBX3+2a Are Functionally Distinctive in Inhibition of Senescence and Are Overexpressed in a Subset of Breast Cancer Cell Lines". In: *Cancer Research* 64.15, pp. 5132–5139.

Fantini, Massimo C et al. (2006). "Drug Insight: novel small molecules and drugs for immunosuppression". In: *Nature Clinical Practice Gastroenterology & Hepatology* 3.11, pp. 633–644.

Fares, Jawad et al. (2020). "Molecular principles of metastasis: a hallmark of cancer revisited". In: *Signal Transduction and Targeted Therapy* 5.1, p. 28.

Farin, Henner F et al. (2013). "Tbx2 Terminates Shh/Fgf Signaling in the Developing Mouse Limb Bud by Direct Repression of Gremlin1". In: *PLoS Genetics* 9.4, e1003467.

Feng, Shao-Wei et al. (2015). "Overexpression of TELO2 decreases survival in human high-grade gliomas". In: *Oncotarget* 7.29, pp. 46056–46066.

Feng, Xianguang et al. (2018). "T-box Transcription Factor Tbx3 Contributes to Human Hepatocellular Carcinoma Cell Migration and Invasion by Repressing E-Cadherin Expression". In: *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics* 26.6, pp. 959–966.

Fillmore, Christine M. et al. (2010). "Estrogen expands breast cancer stem-like cells through paracrine FGF/Tbx3 signaling". In: *Proceedings of the National Academy of Sciences* 107.50, pp. 21737–21742.

Finn, Robert D. et al. (2016). "The Pfam protein families database: towards a more sustainable future". In: *Nucleic Acids Research* 44.D1, pp. D279–D285.

Fischer, Kathrin and Gert O Pflugfelder (2015). "Putative Breast Cancer Driver Mutations in TBX3 Cause Impaired Transcriptional Repression". In: *Frontiers in Oncology* 5, p. 244.

Flaherty, Keith T et al. (2012a). "Improved Survival with MEK Inhibition in BRAF-Mutated Melanoma". In: *The New England Journal of Medicine* 367.2, pp. 107–114.

Flaherty, Keith T. et al. (2012b). "Combined BRAF and MEK Inhibition in Melanoma with BRAF V600 Mutations". In: *The New England Journal of Medicine* 367.18, pp. 1694–1703.

Frank, Deborah U et al. (2013). "Mouse TBX3 mutants suggest novel molecular mechanisms for Ulnar-mammary syndrome." In: *PloS one* 8.7, e67841.

Garcia, Josep et al. (2020). "Bevacizumab (Avastin) in cancer treatment: a review of 15 years of clinical experience and future outlook". In: *Cancer Treatment Reviews* 86, p. 102017.

García-Alai, Maria M et al. (2006). "Molecular Basis for Phosphorylation-Dependent, PEST-Mediated Protein Turnover". In: *Structure* 14.2, pp. 309–319.

García-Jiménez, Custodia and Colin R. Goding (2018). "Starvation and Pseudo-Starvation as Drivers of Cancer Metastasis through Translation Reprogramming". In: *Cell Metabolism* 29.2, pp. 254–267.

García-Borrón, José C., Berta L. Sánchez-Laorden, and Celia Jiménez-Cervantes (2005). "Melanocortin-1 receptor structure and functional regulation". In: *Pigment Cell Research* 18.6, pp. 393–410.

Garnock-Jones, Karly P., Gillian M. Keating, and Lesley J. Scott (2010). "Trastuzumab: A Review of its Use as Adjuvant Treatment in Human Epidermal Growth Factor Receptor 2 (HER2) Positive Early Breast Cancer". In: *Drugs* 70.2, pp. 215–239.

Gay, David M. et al. (2019). "Loss of BCL9/9l suppresses Wnt driven tumourigenesis in models that recapitulate human cancer". In: *Nature Communications* 10.1, p. 723.

Gil, Jesús and Ana O'Loghlen (2014). "PRC1 complex diversity: where is it taking us?" In: *Trends in Cell Biology* 24.11, pp. 632–641.

Ha, Nati, Maria Polychronidou, and Ingrid Lohmann (2012). "COPS: detecting co-occurrence and spatial arrangement of transcription factor binding motifs in genome-wide datasets." In: *PloS one* 7.12, e52055.

Hamid, Omid et al. (2013). "Safety and Tumor Responses with Lambrolizumab (Anti–PD-1) in Melanoma". In: *The New England Journal of Medicine* 369.2, pp. 134–144.

Han, Jianyong et al. (2010). "Tbx3 improves the germ-line competency of induced pluripotent stem cells". In: *Nature* 463.7284, p. 1096.

Hanahan, Douglas and Robert A Weinberg (2000). "The Hallmarks of Cancer". In: *Cell* 100.1, pp. 57–70.

— (2011). "Hallmarks of Cancer: The Next Generation". In: *Cell* 144.5, pp. 646–674.

Havel, Jonathan J., Diego Chowell, and Timothy A. Chan (2019). "The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy". In: *Nature Reviews Cancer* 19.3, pp. 133–150.

Hayflick, L. (1965). "The limited in vitro lifetime of human diploid cell strains". In: *Experimental Cell Research* 37.3, pp. 614–636.

He, Jin et al. (2008). "The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15(Ink4b)." In: *Nature Structural & Molecular Biology* 15.11, pp. 1169–75.

Hen, Omri and Dalit Barkan (2019). "Dormant disseminated tumor cells and cancer stem/progenitor-like cells: similarities and opportunities". In: *Seminars in Cancer Biology* 60, pp. 157–165.

Hodi, F. Stephen et al. (2010). "Improved Survival with Ipilimumab in Patients with Metastatic Melanoma". In: *The New England Journal of Medicine* 363.8, pp. 711–723.

Hojo, Minoru et al. (1999). "Cyclosporine induces cancer progression by a cell-autonomous mechanism". In: *Nature* 397.6719, pp. 530–534.

Hong, David S. et al. (2020). "KRASG12C Inhibition with Sotorasib in Advanced Solid Tumors". In: *New England Journal of Medicine* 383.13, pp. 1207–1217.

Hoogaars, W M H et al. (2007). "Tbx3 controls the sinoatrial node gene program and imposes pacemaker function on the atria". In: *Genes & Development* 21.9, pp. 1098–1112.

Hoogaars, Willem M H et al. (2008). "TBX3 and its splice variant TBX3 + exon 2a are functionally similar". In: *Pigment Cell & Melanoma Research* 21.3, pp. 379–387.

Hoogaars, Willem M.H et al. (2004). "The transcriptional repressor Tbx3 delineates the developing central conduction system of the heart". In: *Cardiovascular Research* 62.3, pp. 489–499.

Horton, Amy C. et al. (2008). "Conservation of linkage and evolution of developmental function within the Tbx2/3/4/5 subfamily of T-box genes: implications for the origin of vertebrate limbs". In: *Development Genes and Evolution* 218.11-12, pp. 613–628.

Hsu, Patrick D, Eric S Lander, and Feng Zhang (2014). "Development and Applications of CRISPR-Cas9 for Genome Engineering". In: *Cell* 157.6, pp. 1262–1278.

Huang, Yongjiu et al. (2020). "TBX3 knockdown suppresses the proliferation of hypopharyngeal carcinoma FaDu cells by inducing G1/S cell cycle arrest and apoptosis". In: *Oncology Letters* 19.1, pp. 113–120.

Humtsoe, Joseph O et al. (2011). "Transcriptional profiling identifies upregulated genes following induction of epithelial-mesenchymal transition in squamous carcinoma cells." In: *Experimental Cell Research* 318.4, pp. 379–90.

Hurlin, Peter J. and Jie Huang (2006). "The MAX-interacting transcription factor network". In: *Seminars in Cancer Biology* 16.4, pp. 265–274.

Hurlin, Peter J., Eirìkur Steingrìmsson, et al. (1999). "Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif". In: *The EMBO Journal* 18.24, pp. 7019–7028.

Ichijo, Ryo, Yui Iizuka, et al. (2017). "Essential roles of Tbx3 in embryonic skin development during epidermal stratification". In: *Genes to Cells* 22.3, pp. 284–292.

Ichijo, Ryo, Hiroki Kobayashi, et al. (2017). "Tbx3-dependent amplifying stem cell progeny drives interfollicular epidermal expansion during pregnancy and regeneration". In: *Nature Communications* 8.1, p. 508.

Imsland, Freyja et al. (2016). "Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses". In: *Nature Genetics* 48.2, pp. 152–158.

Inukai, Sachi, Kian Hong Kock, and Martha L Bulyk (2017). "Transcription factor–DNA binding: beyond binding site motifs". In: *Current Opinion in Genetics & Development* 43, pp. 110–119.

Ismail, Amin and Andrew Bateman (2009). "Expression of TBX2 promotes anchorage-independent growth and survival in the p53-negative SW13 adrenocortical carcinoma". In: *Cancer Letters* 278.2, pp. 230–240.

ITO, A et al. (2005). "Tbx3 expression is related to apoptosis and cell proliferation in rat bladder both hyperplastic epithelial cells and carcinoma cells". In: *Cancer Letters* 219.1, pp. 105–112.

Iwafuchi-Doi, Makiko et al. (2016). "The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation". In: *Molecular Cell* 62.1, pp. 79–91.

Jacobs, JJL, P Keblusek, and Robanus-Maandag E Nature... (2000). "Senescence bypass screen identifies TBX2, which represses Cdkn2a (p19ARF) and is amplified in a subset of human breast cancers". In: *Nature Genetics*.

Janku, Filip, Timothy A. Yap, and Funda Meric-Bernstam (2018). "Targeting the PI3K pathway in cancer: are we making headway?" In: *Nature Reviews Clinical Oncology* 15.5, pp. 273–291.

Jerome-Majewska, Loydie A et al. (2005). "Tbx3 , the ulnar-mammary syndrome gene, and Tbx2 interact in mammary gland development through a p19 Arf /p53-independent pathway". In: *Developmental Dynamics* 234.4, pp. 922–933.

Jiang, Ke, Chunyan Ren, and Venugopalan D Nair (2013). "MicroRNA-137 represses Klf4 and Tbx3 during differentiation of mouse embryonic stem cells." In: *Stem cell research* 11.3, pp. 1299–313.

Jolma, Arttu, Jian Yan, et al. (2013). "DNA-Binding Specificities of Human Transcription Factors". In: *Cell* 152.1-2, pp. 327–339.

Jolma, Arttu, Yimeng Yin, et al. (2015). "DNA-dependent formation of transcription factor pairs alters their binding specificity". In: *Nature* 527.7578, pp. 384–388.

Kaidbey, Kays H. et al. (1979). "Photoprotection by melanin—a comparison of black and Caucasian skin". In: *Journal of the American Academy of Dermatology* 1.3, pp. 249–260.

Kandimalla, Raju et al. (2012). "Genome-wide Analysis of CpG Island Methylation in Bladder Cancer Identified TBX2, TBX3, GATA2, and ZIC4 as pTa-Specific Prognostic Markers". In: *European Urology* 61.6, pp. 1245–1256.

Karoulia, Zoi, Evripidis Gavathiotis, and Poulikos I. Poulikakos (2017). "New perspectives for targeting RAF kinase in human cancer". In: *Nature Reviews Cancer* 17.11, pp. 676–691.

Ke, Minxia et al. (2018). "Leukemia inhibitory factor regulates marmoset induced pluripotent stem cell proliferation via a PI3K/Akt-dependent Tbx-3 activation pathway". In: *International Journal of Molecular Medicine* 42.1, pp. 131–140.

Khan, Saif F et al. (2019). "The roles and regulation of TBX3 in development and disease." In: *Gene* 726, p. 144223.

Kim, Hong et al. (2016). "The Telomerase-Derived Anticancer Peptide Vaccine GV1001 as an Extracellular Heat Shock Protein-Mediated Cell-Penetrating Peptide". In: *International Journal of Molecular Sciences* 17.12, p. 2054.

Kim, Min Seong et al. (2018). "JIB-04, A Small Molecule Histone Demethylase Inhibitor, Selectively Targets Colorectal Cancer Stem Cells by Inhibiting the Wnt/β-Catenin Signaling Pathway". In: *Scientific Reports* 8.1, p. 6611.

Kispert, A. and B.G. Herrmann (1993). "The Brachyury gene encodes a novel DNA binding protein." In: *The EMBO Journal* 12.8, pp. 3211–3220.

Klingenstein, Moritz et al. (2015). "TBX3 Knockdown Decreases Reprogramming Efficiency of Human Cells." In: *Stem Cells International* 2016, p. 6759343.

Krakhmal, N V et al. (2015). "Cancer Invasion: Patterns and Mechanisms." In: *Acta naturae* 7.2, pp. 17–28.

Krstic, Milica, Haider M Hassan, et al. (2019). "Isoform-specific promotion of breast cancer tumorigenicity by TBX3 involves induction of angiogenesis". In: *Laboratory Investigation* 100.3, pp. 400–413.

Krstic, Milica, Bart Kolendowski, et al. (2019). "TBX3 promotes progression of pre-invasive breast cancer cells by inducing EMT and directly up-regulating SLUG". In: *The Journal of Pathology* 248.2, pp. 191–203.

Krstic, Milica, Connor D. Macmillan, et al. (2016). "The transcriptional regulator TBX3 promotes progression from non-invasive to invasive breast cancer". In: *BMC Cancer* 16.1, p. 671.

Kumar, Pavan et al. (2014). "TBX3 regulates splicing in vivo: a novel molecular mechanism for Ulnar-mammary syndrome." In: *PLoS genetics* 10.3, e1004247.

Kumar, Sanjeev et al. (2011). "A pathway for the control of anoikis sensitivity by E-cadherin and epithelial-to-mesenchymal transition". In: *Molecular and cellular biology* 31.19, pp. 4036–4051.

Kumar, Vinit et al. (2016). "The Nature of Myeloid-Derived Suppressor Cells in the Tumor Microenvironment". In: *Trends in Immunology* 37.3, pp. 208–220.

Kunasegaran, Kamini et al. (2014). "Transcriptional Repressor Tbx3 Is Required for the Hormone-Sensing Cell Lineage in Mammary Epithelium". In: *PLoS ONE* 9.10, e110191.

Kwart, Dylan et al. (2017). "Precise and efficient scarless genome editing in stem cells using CORRECT". In: *Nature Protocols* 12.2, pp. 329–354.

Lambert, Samuel A. et al. (2018). "The Human Transcription Factors". In: *Cell* 172.4, pp. 650–665.

Larkin, James et al. (2015). "Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma". In: *The New England Journal of Medicine* 373.1, pp. 23–34.

Lee, Tong Ihn and Richard A. Young (2013). "Transcriptional Regulation and Its Misregulation in Disease". In: *Cell* 152.6, pp. 1237–1251.

Lee, Young-Ju, Jung-Hwan Park, and Sang-Muk Oh (2019). "TOPK promotes epithelial-mesenchymal transition and invasion of breast cancer cells through up-regulation of TBX3 in TGF-β1/Smad signaling". In: *Biochemical and Biophysical Research Communications* 522.1, pp. 270–277.

Li, Jarod, Deeya Ballim, et al. (2014). "The Anti-proliferative Function of the TGF-β1 Signaling Pathway Involves the Repression of the Oncogenic TBX2 by Its Homologue TBX3". In: *Journal of Biological Chemistry* 289.51, pp. 35633–35643.

Li, Jarod, Marc S Weinberg, et al. (2013). "The oncogenic TBX3 is a downstream target and mediator of the TGF-β1 signaling pathway." In: *Molecular Biology of the Cell* 24.22, pp. 3569–76.

Li, Xiaomeng et al. (2018). "TBX3 promotes proliferation of papillary thyroid carcinoma cells through facilitating PRC2-mediated p57KIP2 repression". In: *Oncogene* 37.21, pp. 2773–2792.

Li, Yina et al. (2004). "Sonic hedgehog signaling regulates Gli3 processing, mesenchymal proliferation, and differentiation during mouse lung organogenesis". In: *Developmental Biology* 270.1, pp. 214–231.

Li, Zhian et al. (2018). "Expression of TBX3 in Hepatocellular Carcinoma and Its Clinical Implication." In: *Medical science monitor : international medical journal of experimental and clinical research* 24, pp. 9324–9333.

Lickwar, Colin R. et al. (2012). "Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function". In: *Nature* 484.7393, pp. 251–255.

Lin, Yuxin, Jianxin Xu, and Huiyin Lan (2019). "Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications". In: *Journal of Hematology & Oncology* 12.1, p. 76.

Lingbeek, Merel E, Jacqueline JL Jacobs, and Maarten van Lohuizen (2002). "The T-box Repressors TBX2 and TBX3 Specifically Regulate the Tumor Suppressor Gene p14 ARF via a Variant T-site in the Initiator". In: *Journal of Biological Chemistry* 277.29, pp. 26120–26127.

Liu, Fang et al. (2013). "TBX2 expression is regulated by PAX3 in the melanocyte lineage." In: *Pigment Cell & Melanoma Research* 26.1, pp. 67–77.

Liu, Xingyu et al. (2019). "TBX2 overexpression promotes proliferation and invasion through epithelial-mesenchymal transition and ERK signaling pathway". In: *Experimental and Therapeutic Medicine* 17.1, pp. 723–729.

Liu, Zheng-gang and Delong Jiao (2020). "Necroptosis, tumor necrosis and tumorigenesis". In: *Cell Stress* 4.1, p. 1.

Louphrasitthiphol, Pakavarin et al. (2020). "Tuning Transcription Factor Availability through Acetylation-Mediated Genomic Redistribution". In: *Molecular Cell* 79.3, 472–487.e10.

Lu, Juan et al. (2010). "TBX2 and TBX3: The special value for anticancer drug targets". In: *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1806.2, pp. 268–274.

Lu, Jun et al. (2020). "TBX2 Expression predicts Tumor Recurrence and Adjuvant Chemotherapy Benefits in Gastric Cancer Patients following R0 Resection: a proposed approach for risk stratification". In: *Journal of Cancer* 11.11, pp. 3172–3179.

Lu, Rui, Acong Yang, and Ying Jin (2010). "Dual functions of T-box 3 (Tbx3) in the control of self-renewal and extraembryonic endoderm differentiation in

mouse embryonic stem cells." In: *The Journal of Biological Chemistry* 286.10, pp. 8425–36.

Lüdtke, Timo et al. (2013). "Tbx2 Controls Lung Growth by Direct Repression of the Cell Cycle Inhibitor Genes Cdkn1a and Cdkn1b". In: *PLoS Genetics* 9.1, e1003189.

Lüdtke, Timo H et al. (2016). "Tbx2 and Tbx3 Act Downstream of Shh to Maintain Canonical Wnt Signaling during Branching Morphogenesis of the Murine Lung". In: *Developmental Cell* 39.2, pp. 239–253.

Lüdtke, Timo H -W et al. (2008). "Tbx3 promotes liver bud expansion during mouse development by suppression of cholangiocyte differentiation". In: *Hepatology* 49.3, pp. 969–978.

Lugano, Roberta, Mohanraj Ramachandran, and Anna Dimberg (2020). "Tumor angiogenesis: causes, consequences, challenges and opportunities". In: *Cellular and Molecular Life Sciences* 77.9, pp. 1745–1770.

Luke, Jason J. et al. (2017). "Targeted agents and immunotherapies: optimizing outcomes in melanoma". In: *Nature Reviews Clinical Oncology* 14.8, pp. 463–482.

Maciejowski, John and Titia de Lange (2017). "Telomeres in cancer: tumour suppression and genome instability". In: *Nature Reviews Molecular Cell Biology* 18.3, pp. 175–186.

Martin, Nadine et al. (2012). "Physical and functional interaction between PML and TBX2 in the establishment of cellular senescence". In: *The EMBO Journal* 31.1, pp. 95–109.

McArthur, Grant A et al. (2014). "Safety and efficacy of vemurafenib in BRAF V600E and BRAF V600K mutation-positive melanoma (BRIM-3): extended follow-up of a phase 3, randomised, open-label study". In: *The Lancet Oncology* 15.3, pp. 323–332.

Mendonsa, Alisha M, Tae-Young Na, and Barry M Gumbiner (2018). "E-cadherin in contact inhibition and cancer". In: *Oncogene* 37.35, pp. 4769–4780.

Menzies, Scott W. et al. (2013). "Dermoscopic Evaluation of Nodular Melanoma". In: *JAMA Dermatology* 149.6, pp. 699–709.

Meurette, Olivier and Patrick Mehlen (2018). "Notch Signaling in the Tumor Microenvironment". In: *Cancer Cell* 34.4, pp. 536–548.

Miao, Zhi-Feng et al. (2016). "Tbx3 overexpression in human gastric cancer is correlated with advanced tumor stage and nodal status and promotes cancer cell growth and invasion". In: *Virchows Archiv* 469.5, pp. 505–513.

Michaloglou, Chrysiis et al. (2005). "BRAFE600-associated senescence-like cell cycle arrest of human naevi". In: *Nature* 436.7051, pp. 720–724.

Mohamad, Trefa et al. (2018). "EGR1 interacts with TBX2 and functions as a tumor suppressor in rhabdomyosarcoma". In: *Oncotarget* 9.26, pp. 18084–18098.

Momcilovic, M and D B Shackelford (2015). "Targeting LKB1 in cancer – exposing and exploiting vulnerabilities". In: *British Journal of Cancer* 113.4, pp. 574–584.

Motahari, Zahra et al. (2016). "Tbx3 represses bmp4 expression and, with Pax6, is required and sufficient for retina formation". In: *Development* 143.19, pp. 3560–3572.

Mowla, Shaheen et al. (2011). "PMA-induced up-regulation of TBX3 is mediated by AP-1 and contributes to breast cancer cell migration". In: *Biochemical Journal* 433.1, pp. 145–153.

Mughal, Tariq I and John M Goldman (2006). "Molecularly targeted treatment of chronic myeloid leukemia: beyond the imatinib era". In: *Frontiers in Bioscience* 11.1, p. 209.

Nandana, Srinivas et al. (2017). "Bone metastasis of prostate cancer can be therapeutically targeted at the TBX2-WNT signaling axis." In: *Cancer Research* 77.6, pp. 1331–1344.

Nelson, Joel D, Oleg Denisenko, and Karol Bomsztyk (2006). "Protocol for the fast chromatin immunoprecipitation (ChIP) method". In: *Nature Protocols* 1.1, pp. 179–185.

Niwa, Hitoshi et al. (2009). "A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells". In: *Nature* 460.7251, pp. 118–122.

Nurwidya, Fariz, Fumiyuki Takahashi, and Kazuhisa Takahashi (2016). "Gefitinib in the treatment of nonsmall cell lung cancer with activating epidermal growth factor receptor mutation". In: *Journal of Natural Science, Biology and Medicine* 7.2, pp. 119–123.

O'Bryan, John P. (2018). "Pharmacological Targeting of RAS: Recent Success with Direct Inhibitors". In: *Pharmacological Research* 139, pp. 503–511.

Oh, Teak-Jung et al. (2019). "TBX3 represses TBX2 under the control of the PRC2 complex in skeletal muscle and rhabdomyosarcoma". In: *Oncogenesis* 8.4, p. 27.

Olivier, Magali, Monica Hollstein, and Pierre Hainaut (2010). "TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use". In: *Cold Spring Harbor Perspectives in Biology* 2.1, a001008.

Osterwalder, Marco et al. (2014). "HAND2 targets define a network of transcriptional regulators that compartmentalize the early limb bud mesenchyme." In: *Developmental Cell* 31.3, pp. 345–57.

Pan, Chih-Ming et al. (2020). "MicroRNA-7 targets T-Box 2 to inhibit epithelial-mesenchymal transition and invasiveness in glioblastoma multiforme". In: *Cancer Letters* 493, pp. 133–142.

Papaioannou, Virginia E (2014). "The T-box gene family: emerging roles in development, stem cells and cancer". In: *Development* 141.20, pp. 3819–3833.

Paquet, Dominik et al. (2016). "Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9". In: *Nature* 533.7601, pp. 125–129.

Parrish, Janet K et al. (2018). "The Jumonji-domain histone demethylase inhibitor JIB-04 deregulates oncogenic programs and increases DNA damage in Ewing Sarcoma, resulting in impaired cell proliferation and survival, and reduced tumor growth". In: *Oncotarget* 9.69, pp. 33110–33123.

Pavan, Kumar P et al. (2014). "Coordinated control of senescence by lncRNA and a novel T-box3 co-repressor complex". In: *eLife* 3, e02805.

Pavlova, Natalya N. and Craig B. Thompson (2016). "The Emerging Hallmarks of Cancer Metabolism". In: *Cell Metabolism* 23.1, pp. 27–47.

Paxton, Christian et al. (2002). "Murine Tbx2 contains domains that activate and repress gene transcription". In: *Gene* 283.1-2, pp. 117–124.

Pedersen, Marianne Terndrup et al. (2014). "The Demethylase JMJD2C Localizes to H3K4me3-Positive Transcription Start Sites and Is Dispensable for Embryonic Development". In: *Molecular and Cellular Biology* 34.6, pp. 1031–1045.

Peres, Jade, Emily Davis, et al. (2010). "The Highly Homologous T-Box Transcription Factors, TBX2 and TBX3, Have Distinct Roles in the Oncogenic Process". In: *Genes & Cancer* 1.3, pp. 272–282.

Peres, Jade, Eliza M. Kwesi-Maliepaard, et al. (2017). "The tumour suppressor, miR-137, inhibits malignant melanoma migration by targetting the TBX3 transcription factor". In: *Cancer Letters* 405, pp. 111–119.

Peres, Jade, Shaheen Mowla, and Sharon Prince (2014). "The T-box transcription factor, TBX3, is a key substrate of AKT3 in melanomagenesis". In: *Oncotarget* 6.3, pp. 1821–1833.

Peres, Jade and Sharon Prince (2013). "The T-box transcription factor, TBX3, is sufficient to promote melanoma formation and invasion". In: *Molecular Cancer* 12.1, pp. 1–5.

Perkhofer, Lukas et al. (2016). "Tbx3 fosters pancreatic cancer growth by increased angiogenesis and activin/nodal-dependent induction of stemness". In: *Stem Cell Research* 17.2, pp. 367–378.

Piovesan, Allison et al. (2019). "Human protein-coding genes and gene feature statistics in 2019". In: *BMC Research Notes* 12.1, p. 315.

Platonova, Natalia et al. (2007). "TBX3, the gene mutated in ulnar-mammary syndrome, promotes growth of mammary epithelial cells via repression of p19ARF, independently of p53". In: *Cell and Tissue Research* 328.2, pp. 301–316.

Prince, Sharon et al. (2004). "Tbx2 directly represses the expression of the p21(WAF1) cyclin-dependent kinase inhibitor." In: *Cancer Research* 64.5, pp. 1669–74.

Qu, Xueping et al. (2003). "Promotion of tumorigenesis by heterozygous disruption of the beclin 1 autophagy gene". In: *Journal of Clinical Investigation* 112.12, pp. 1809–1820.

Quarta, Carmelo et al. (2019). "Functional identity of hypothalamic melanocortin neurons depends on Tbx3." In: *Nature metabolism* 1.2, pp. 222–235.

Rechsteiner, Martin and Scott W Rogers (1996). "PEST sequences and regulation by proteolysis". In: *Trends in Biochemical Sciences* 21.7, pp. 267–271.

Redmond, KL et al. (2010). "T-box 2 represses NDRG1 through an EGR1-dependent mechanism to drive the proliferation of breast cancer cells". In: *Oncogene* 29.22, p. 3252.

Reich, Shoshana et al. (2020). "Tbx2 mediates dorsal patterning and germ layer suppression through inhibition of BMP/GDF and Activin/Nodal signaling". In: *BMC Molecular and Cell Biology* 21.1, p. 39.

Reid, A.L. et al. (2013). "Markers of circulating tumour cells in the peripheral blood of patients with melanoma correlate with disease recurrence and progression". In: *British Journal of Dermatology* 168.1, pp. 85–92.

Reinhardt, Sven et al. (2019). "Transcriptional repression of the ectodomain sheddase ADAM10 by TBX2 and potential implication for Alzheimer's disease". In: *Cellular and Molecular Life Sciences* 76.5, pp. 1005–1025.

Renard, Claire-Angélique et al. (2007). "Tbx3 Is a Downstream Target of the Wnt/β-Catenin Pathway and a Critical Mediator of β-Catenin Survival Functions in Liver Cancer". In: *Cancer Research* 67.3, pp. 901–910.

Renkawitz, Rainer (2006). "Transcription Factors and Regulation of Gene Expression". In: *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1886–1890. URL: https://doi.org/10.1007/3-540-29623-9_2520.

Ribas, Antoni et al. (2014). "Combination of vemurafenib and cobimetinib in patients with advanced BRAF V600-mutated melanoma: a phase 1b study". In: *The Lancet Oncology* 15.9, pp. 954–965.

Robert, Caroline et al. (2011). "Ipilimumab plus Dacarbazine for Previously Untreated Metastatic Melanoma". In: *The New England Journal of Medicine* 364.26, pp. 2517–2526.

Rodriguez, Mercedes et al. (2008). "Tbx3 Represses E-Cadherin Expression and Enhances Melanoma Invasiveness". In: *Cancer Research* 68.19, pp. 7872–7881.

Rosenkranz, A. A. et al. (2013). "Malignant melanoma and melanocortin 1 receptor". In: *Biochemistry (Moscow)* 78.11, pp. 1228–1237.

Rowland, Benjamin D et al. (2002). "E2F transcriptional repressor complexes are critical downstream targets of p19ARF/p53-induced proliferative arrest". In: *Cancer Cell* 2.1, pp. 55–65.

Roworth, A Poppy, Fatemeh Ghari, and Nicholas B La Thangue (2015). "To live or let die – complexity within the E2F1 pathway". In: *Molecular & Cellular Oncology* 2.1, e970480.

Russell, Ronan et al. (2015). "A Dynamic Role of TBX3 in the Pluripotency Circuitry." In: *Stem Cell Reports* 5.6, pp. 1155–70.

Saadi, Irfan et al. (2013). "Msx1 and Tbx2 antagonistically regulate Bmp4 expression during the bud-to-cap stage transition in tooth development". In: *Development* 140.13, pp. 2697–2702.

Sahm, Bianca Del B. et al. (2020). "Targeting the Oncogenic TBX2 Transcription Factor With Chromomycins". In: *Frontiers in Chemistry* 8, p. 110.

Sainathan, Satheesh et al. (2015). "Histone Demethylases in Cancer". In: *Current Pharmacology Reports* 1.4, pp. 234–244.

Sakabe, M et al. (2011). "Ectopic retinoic acid signaling affects outflow tract cushion development through suppression of the myocardial Tbx2-Tgf 2 pathway". In: *Development* 139.2, pp. 385–395.

Sample, Ashley and Yu-Ying He (2018). "Mechanisms and prevention of UV-induced melanoma". In: *Photodermatology, Photoimmunology & Photomedicine* 34.1, pp. 13–24.

Sanborn, J. Zachary et al. (2015). "Phylogenetic analyses of melanoma reveal complex patterns of metastatic dissemination". In: *Proceedings of the National Academy of Sciences* 112.35, pp. 10995–11000.

Sander, Jeffry D and Keith J Joung (2014). "CRISPR-Cas systems for editing, regulating and targeting genomes". In: *Nature Biotechnology* 32.4, pp. 347–355.

Sardar, Samra et al. (2019). "The oncoprotein TBX3 is controlling severity in experimental arthritis." In: *Arthritis research & therapy* 21.1, p. 16.

Schadendorf, Dirk et al. (2015). "Melanoma". In: *Nature Reviews Disease Primers* 1.1, p. 15003.

Schmid-Burgk, Jonathan L et al. (2016). "CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism". In: *Nature Communications* 7, p. 12338.

Serra, Francesco et al. (2019). "Palbociclib in metastatic breast cancer: current evidence and real-life data". In: *Drugs in Context* 8, pp. 1–16.

Shain, A. Hunter and Boris C. Bastian (2016). "From melanocytes to melanomas". In: *Nature Reviews Cancer* 16.6, pp. 345–358.

Shain, A. Hunter, Iwei Yeh, et al. (2015). "The Genetic Evolution of Melanoma from Precursor Lesions". In: *The New England Journal of Medicine* 373.20, pp. 1926–1936.

Shan, Ze-Zhi et al. (2014). "Overexpression of Tbx3 is correlated with Epithelial-Mesenchymal Transition phenotype and predicts poor prognosis of colorectal cancer." In: *American journal of cancer research* 5.1, pp. 344–53.

Sharpless, Norman E and Charles J Sherr (2015). "Forging a signature of in vivo senescence". In: *Nature Reviews Cancer* 15.7, pp. 397–408.

Shay, Jerry W. and Woodring E. Wright (2000). "Hayflick, his limit, and cellular ageing". In: *Nature Reviews Molecular Cell Biology* 1.1, pp. 72–76.

Shklover, Jeny et al. (2007). "MyoD uses overlapping but distinct elements to bind E-box and tetraplex structures of regulatory sequences of muscle-specific genes". In: *Nucleic Acids Research* 35.21, pp. 7087–7095.

Shors, A.R. et al. (2006). "Dysplastic naevi with moderate to severe histological dysplasia: a risk factor for melanoma". In: *British Journal of Dermatology* 155.5, pp. 988–993.

Siu, Ka Tat, Marsha Rich Rosner, and Alex C. Minella (2014). "An integrated view of cyclin E function and regulation". In: *Cell Cycle* 11.1, pp. 57–64.

Slattery, Matthew et al. (2011). "Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins". In: *Cell* 147.6, pp. 1270–1282.

Stewart-Ornstein, Jacob and Galit Lahav (2016). "Dynamics of CDKN1A in Single Cells Defined by an Endogenous Fluorescent Tagging Toolkit". In: *Cell Reports* 14.7, pp. 1800–1811.

Suzuki, A et al. (2008). "Tbx3 controls the fate of hepatic progenitor cells in liver development by suppressing p19ARF expression". In: *Development* 135.9, pp. 1589–1595.

Symington, Lorraine S (2016). "Mechanism and regulation of DNA end resection in eukaryotes." In: *Critical Reviews in Biochemistry and Molecular biology* 51.3, pp. 195–212.

Talve, Lauri et al. (1997). "Loss of expression of the p16INK4/CDKN2 gene in cutaneous malignant melanoma correlates with tumor cell proliferation and invasive stage". In: *International Journal of Cancer* 74.3, pp. 255–259.

Tasaka, Reiko et al. (2018). "TBX2 expression is associated with platinum-sensitivity of ovarian serous carcinoma". In: *Oncology Letters* 15.3, pp. 3085–3090.

Tomaz, Rute A. et al. (2017). "Jmjd2c/Kdm4c facilitates the assembly of essential enhancer-protein complexes at the onset of embryonic stem cell differentiation". In: *Development* 144.4, dev.142489.

Topalian, Suzanne L. et al. (2014). "Survival, Durable Tumor Remission, and Long-Term Safety in Patients With Advanced Melanoma Receiving Nivolumab". In: *Journal of Clinical Oncology* 32.10, pp. 1020–1030.

Tzatsos, Alexandros, Polina Paskaleva, et al. (2011). "Lysine-specific demethylase 2B (KDM2B)-let-7-enhancer of zester homolog 2 (EZH2) pathway regulates cell cycle progression and senescence in primary cells." In: *The Journal of Biological Chemistry* 286.38, pp. 33061–9.

Tzatsos, Alexandros, Raymond Pfau, et al. (2009). "Ndy1/KDM2B immortalizes mouse embryonic fibroblasts by repressing the Ink4a/Arf locus." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.8, pp. 2641–6.

Ugurel, Selma, Annette Paschen, and Jürgen C. Becker (2013). "Dacarbazine in Melanoma: From a Chemotherapeutic Drug to an Immunomodulating Agent". In: *Journal of Investigative Dermatology* 133.2, pp. 289–292.

Vance, Keith W, Suzanne Carreira, et al. (2005). "Tbx2 Is Overexpressed and Plays an Important Role in Maintaining Proliferation and Suppression of Senescence in Melanomas". In: *Cancer Research* 65.6, pp. 2260–2268.

Vance, Keith W, Heather M Shaw, et al. (2010). "The Retinoblastoma Protein Modulates Tbx2 Functional Specificity". In: *Molecular Biology of the Cell* 21.15, pp. 2770–2779.

Vaquerizas, Juan M. et al. (2009). "A census of human transcription factors: function, expression and evolution". In: *Nature Reviews Genetics* 10.4, pp. 252–263.

Vivas-García, Yurena et al. (2020). "Lineage-Restricted Regulation of SCD and Fatty Acid Saturation by MITF Controls Melanoma Phenotypic Plasticity". In: *Molecular Cell* 77.1, 120–137.e9.

Waghray, Avinash et al. (2015). "Tbx3 Controls Dppa3 Levels and Exit from Pluripotency toward Mesoderm". In: *Stem Cell Reports* 5.1, pp. 97–110.

Waldman, Alex D., Jill M. Fritz, and Michael J. Lenardo (2020). "A guide to cancer immunotherapy: from T cell basic science to clinical practice". In: *Nature Reviews Immunology*, pp. 1–18.

Wang, Bin et al. (2012). "The T Box Transcription Factor TBX2 Promotes Epithelial-Mesenchymal Transition and Invasion of Normal and Malignant Breast Epithelial Cells". In: *PLoS ONE* 7.7, e41355.

Wang, Lei et al. (2013). "A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth". In: *Nature Communications* 4.1, p. 2035.

Wang, Xiaofeng, Jeffrey R. Haswell, and Charles W.M. Roberts (2014). "Molecular Pathways: SWI/SNF (BAF) Complexes Are Frequently Mutated in Cancer—Mechanisms and Potential Therapeutic Insights". In: *Clinical Cancer Research* 20.1, pp. 21–27.

Wang, Yifan (2018). "Expression level of TBX3 gene in renal carcinoma and its clinical significance". In: *Oncology Letters* 15.4, pp. 4235–4240.

Wansleben, Sabina et al. (2014). "T-box transcription factors in cancer biology". In: *Biochimica et Biophysica Acta* 1846.2.

Wensing, Lislaine A and Alexandre H Campos (2014). "TBX3, a downstream target of TGF-β1, inhibits mesangial cell apoptosis." In: *Experimental Cell Research* 328.2, pp. 340–50.

Williams, Patricia F. et al. (2011). "Melanocortin 1 receptor and risk of cutaneous melanoma: A meta-analysis and estimates of population burden". In: *International Journal of Cancer* 129.7, pp. 1730–1740.

Willmer, T et al. (2016). "The T-box transcription factor 3 is a promising biomarker and a key regulator of the oncogenic phenotype of a diverse range of sarcoma subtypes." In: *Oncogenesis* 5.2, e199.

Willmer, Tarryn, Shannagh Hare, et al. (2016). "The T-box transcription factor TBX3 drives proliferation by direct repression of the p21WAF1 cyclin-dependent kinase inhibitor". In: *Cell Division* 11.1, p. 6.

Willmer, Tarryn, Jade Peres, et al. (2015). "The T-Box factor TBX3 is important in S-phase and is regulated by c-Myc and cyclin A-CDK2". In: *Cell Cycle* 14.19, pp. 3173–3183.

Wojahn, Irina et al. (2019). "TBX2-positive cells represent a multi-potent mesenchymal progenitor pool in the developing lung". In: *Respiratory Research* 20.1, p. 292.

Wolchok, Jedd D. et al. (2013). "Nivolumab plus Ipilimumab in Advanced Melanoma". In: *The New England Journal of Medicine* 369.2, pp. 122–133.

Xiao, Yufei et al. (2014). "A novel significance score for gene selection and ranking". In: *Bioinformatics* 30.6, pp. 801–807.

Xie, Huilin et al. (2018). "Identification of TBX2 and TBX3 variants in patients with conotruncal heart defects by target sequencing." In: *Human genomics* 12.1, p. 44.

Yamada, Miho et al. (2000). "Expression of Chick Tbx-2, Tbx-3, and Tbx-5 Genes during Early Heart Development: Evidence for BMP2 Induction of Tbx2". In: *Developmental Biology* 228.1, pp. 95–105.

Yang, Lei et al. (2006). "Isl1Cre reveals a common Bmp pathway in heart and limb development". In: *Development* 133.8, pp. 1575–1585.

Yano, Tomoki et al. (2011). "Tara up-regulates E-cadherin transcription by binding to the Trio RhoGEF and inhibiting Rac signaling." In: *The Journal of Cell Biology* 193.2, pp. 319–32.

Yao, Chen et al. (2014). "The Transcription Factor T-box 3 Regulates Colony-stimulating Factor 1-dependent Jun Dimerization Protein 2 Expression and Plays an Important Role in Osteoclastogenesis". In: *Journal of Biological Chemistry* 289.10, pp. 6775–6790.

Yarosh, Will et al. (2008). "TBX3 Is Overexpressed in Breast Cancer and Represses p14ARF by Interacting with Histone Deacetylases". In: *Cancer Research* 68.3, pp. 693–699.

Yeh, Iwei, Andreas von Deimling, and Boris C. Bastian (2013). "Clonal BRAF Mutations in Melanocytic Nevi and Initiating Role of BRAF in Melanocytic

Neoplasia". In: *JNCI: Journal of the National Cancer Institute* 105.12, pp. 917–919.

Yi, Ming et al. (2019). "Advances and perspectives of PARP inhibitors". In: *Experimental Hematology & Oncology* 8.1, p. 29.

Yin, Chengqian et al. (2019). "Pharmacological Targeting of STK19 Inhibits Oncogenic NRAS-Driven Melanomagenesis." In: *Cell* 176.5, 1113–1127.e16.

Yu, Huajian et al. (2020). "TBX2 Identified as a Potential Predictor of Bone Metastasis in Lung Adenocarcinoma via Integrated Bioinformatics Analyses and Verification of Functional Assay". In: *Journal of Cancer* 11.2, pp. 388–402.

Yu, Weina et al. (2019). "Cadherin Signaling in Cancer: Its Functions and Role as a Therapeutic Target". In: *Frontiers in Oncology* 9, p. 989.

Yu, Yong et al. (2018). "Targeting the Senescence-Overriding Cooperative Activity of Structurally Unrelated H3K9 Demethylases in Melanoma." In: *Cancer Cell* 33.2, 322–336.e8.

Yu, Zhanjiang et al. (2020). "Transcription box-3 protects human umbilical vein endothelial cells in a high-glucose environment through sirtuin 1/AKT signaling". In: *Molecular Medicine Reports* 22.2, pp. 1145–1154.

Zhang, Ran-ran et al. (2019). "A regulatory variant in TBX2 promoter is related to the decreased susceptibility of congenital heart disease in the Han Chinese population". In: *Molecular Genetics & Genomic Medicine* 7.2, e00530.

Zhang, Wensheng et al. (2019). "The BAF and PRC2 Complex Subunits Dpf2 and Eed Antagonistically Converge on Tbx3 to Control ESC Differentiation". In: *Cell Stem Cell* 24.1, 138–152.e8.

Zhang, Zeda et al. (2020). "Loss of CHD1 Promotes Heterogeneous Mechanisms of Resistance to AR-Targeted Therapy via Chromatin Dysregulation." In: *Cancer cell*.

Zhang, Zhen et al. (2011). "The microRNA-processing enzyme Dicer is dispensable for somite segmentation but essential for limb bud positioning." In: *Developmental Biology* 351.2, pp. 254–65.

Zhao, Hongyi et al. (2020). "Overexpression of TBX3 in human induced pluripotent stem cells (hiPSCs) increases their differentiation into cardiac pacemaker-like cells." In: *Biomedicine & Pharmacotherapy* 130, p. 110612.

Zhu, B et al. (2015). "TBX2 represses PTEN in rhabdomyosarcoma and skeletal muscle". In: *Oncogene* 35.32, pp. 4212–4224.

Zhu, Bo et al. (2014). "TBX2 blocks myogenesis and promotes proliferation in rhabdomyosarcoma cells". In: *International Journal of Cancer* 135.4, pp. 785–797.

Zhu, Fangjie et al. (2018). "The interaction landscape between transcription factors and the nucleosome". In: *Nature* 562.7725, pp. 76–81.

Zimmerli, Dario et al. (2020). "TBX3 acts as tissue-specific component of the Wnt/β-catenin transcriptional complex". In: *eLife* 9, e58123.

Zolghadr, Kourosh, Ulrich Rothbauer, and Heinrich Leonhardt (2012). "The Fluorescent Two-Hybrid (F2H) Assay for Direct Analysis of Protein–Protein Interactions in Living Cells". In: *Two Hybrid Technologies: Methods and Protocols*. Ed. by Bernhard Suter and Erich E. Wanker. Totowa, NJ: Humana Press, pp. 275–282.