The Role of FOXM1 and Co-factors in Breast Cancer – Endoplasmic Reticulum Stress, Epigenetics, Endocrine Therapy and Chemotherapy

By

Jay Sze Yong

A thesis submitted for the degree Doctor of Philosophy

Department of Surgery and Cancer

Imperial College London

Statement of Originality

I declare that otherwise acknowledged below or stated in the text, this thesis is the result of my own work. I am grateful for the ATAC-sequencing experiments conducted as a collaborated work between **Dr. Raffaella Lucciola**, a then PhD student of Professor Jan Brosens from the University of Warwick, and the members of Professor Eric Lam's research group – **Dr. Stefania Zona** and I, under the support of our laboratory and resources. Additionally, the MCF-7-Y537S mutant cell line was a kind contribution of **Dr. Alison Harrod** from Professor Simak Ali's research team from Department of Surgery and Cancer, Imperial College London.

I am thankful for the helpful guidance of **Dr. Lukasz Boryn** and **Dr. Chun Fui Lai** during the initial set-up of my bioinformatics analyses. I am appreciative of the Master's students whom I supervised during my PhD study, **Mr. Yuliang He**, who contributed to Chapter 4 on the study of DDIT3; and **Miss Dilyana Doncheva,** who have contributed to Chapter 5 on the study of DNA methyltransferases.

Copyright Declaration

'The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.'

Abstract

Breast cancer is an umbrella term for heterogeneous diseases that originated from breast tissues. To date, five main treatment strategies are available for breast cancer treatment – comprising of surgery, radiotherapy, chemotherapy, hormone therapy, and biological therapy – and these treatments are not categorically exclusive but are used in combinations. The treatment arrangement is also dependent on the type, size, location, and stage of cancer. Nevertheless, the interpretation and complete removal of the disease based on sequencing results remain a challenge. In order to better tackle the disease, different classifiers have been established to categorise breast cancer cases into buckets with their own distinct features. Particularly, molecular subtyping separates breast cancer cases based on the presence of specific markers such as ESR1, PR, HER2, and Ki-67 with diverse prognostic outcomes and treatment responses in patients. FOX proteins belong to a transcription factor superfamily that drives not only important molecular mechanisms in normal cells, but the aberrant regulation of these members leads to oncogenic outcomes and cancer progression. Particularly, *FOXM1* is an oncogene that is associated with the hallmarks of cancer. It has been shown that FOXM1 cistrome predicts breast cancer metastatic outcomes better than FOXM1 expression levels. Therefore, clarity about FOXM1 cistrome and interactome needs to be prioritised to improve breast cancer treatment and reduce treatment resistance. In this dissertation, I built on existing knowledge of the role of FOXM1 in breast cancer cell lines by providing new preliminary insights that pave new roads for future studies in hormonal and chemotherapeutic response, unfolded protein response, and epigenetic mechanism. FOXM1 inhibitors are attractive agents to constitute combinatorial treatment with fulvestrant for constitutively activate ESR1-positive breast cancer with lower toxicity to cells. Additionally, FOXM1 serves a role in unfolded protein response in breast cancer cells by a potential axis with endoplasmic reticulum stress effector, DDIT3. Lastly, FOXM1 differentially interacts with DNA methyltransferases – specific interaction of FOXM1 and DNMT3B was detected in MCF-7 and MCF7-Epi^{R} cell lines. Collectively, this study extended our understandings of the molecular diversity of FOXM1 and its cofactors in breast cancer cell lines and may serve as a stepping stone to identify prognostic markers and treatment targets in breast cancer.

Acknowledgment

Throughout the course of my PhD study, I have received a great deal of support and assistance from many people, without whom I would not have been able to complete this research and made it through the PhD.

Firstly, I am indebted to my family. Family always comes first. I am not the best daughter and sister, but you have never failed to accommodate me, act as my backbone, and provide me unlimited reasons to call you my home proudly. I hope I can make you feel the same to any extent. To my dearest brothers, thank you for providing me so much happiness in life. To my parents, you are the reason for my being, and I owe everything to you! As Alfred Adler quoted: “A lucky person’s lifetime is healed by childhood, while an unlucky person’s lifetime is spent healing childhood.” I am indeed extremely lucky – you nurtured me with unconditional love and empathy and thanks to that I grew up to be happy, kind, fearless, and ambitious.

Secondly, I would like to acknowledge the whole Professor Eric Lam research group, present and past, for their collaboration and friendship during my time in the laboratory. Thank you for your empathetic ears and advice even when the road got tough. A special thank you to the colleagues I have met on the first floor of Imperial Centre for Translation and Experimental Medicine (ICTEM), whom I can now proudly call my friends – those long laboratory days and late nights would have been unbearable if it wasn’t for your company and support.

Thirdly, I would like to thank Imperial College London for awarding me a scholarship to undertake my PhD. Thank you for allowing me the opportunity to embark on this eye-opening, once in a lifetime learning journey. As Nelson Mandella said: “Education is the most powerful weapon we can use to change the world.” I am blessed to have equipped myself with the best weapon for the path forward.

Fourthly, thank you to my supervisor, Dr. Ernesto Yague, the Director of Postgraduate Studies, Professor Michael Seckl, and Postgraduate Education Administrators from the Department of Surgery and Cancer, Jennifer Simeon and Raleigh Mangsat, for supporting me through the last stretch of this seemingly endless PhD journey then and pushing me to the finish line.

Finally, I am beyond grateful for the beautiful friendships and relationships that I have developed in London. Shout out to Alyssa Rachel Kong, Yvonne Goh, Richard La, Eva Woon, Wen Yee Kok, Min Hua Chia, and Yue Yang Huang - I would not be the unique person who knows how to be confident in my own skin and love myself without your great care. I have had many acquaintances but when I met you I understood the meaning of true connections and relationships. I will always try to provide you with the same. Thank you for motivating me to want to be a better person and yet love me for who I am.

Contents

Statement of Originality 2
Copyright Declaration 2
Abstract 3
Acknowledgment 4
List of Figures 11
List of Tables 12
List of Appendix 12
List of Abbreviations 13
1 Introduction 19
1.1 Breast Cancer 19
1.1.1 Breast Cancer Heterogeneity and Subtypes 20
1.2 ESR1 21
1.2.1 Mutations in ESR1s in breast cancer tumours 21
1.2.2 Biology of ESR1 LBD Mutations: Y537 and D538 mutation “hotspots” 23
1.3 Mechanism of Action of Estrogen and ESR1 24
1.4 ER Ligands and Endocrine Therapies 27
1.4.1 Tamoxifen, anti-estrogen 27
1.4.2 Aromatase Inhibitors (AIs) 28
1.4.3 Fulvestrant, the ‘pure’ anti-estrogens 29
1.5 Breast Cancer Treatments 30
1.5.1 Endocrine Therapy Resistance 31
1.5.2 ESR1-negative Breast Cancer Treatments 32
1.6 Forkhead Transcription Factor Superfamily 33
1.6.1 Protein domains, expression patterns and post-translational modifications 33
1.6.2 FOXA family 36
1.6.3 FOXM1 37
1.6.4 FOXO factors 38
1.6.5 Overlapping and Co-regulatory Roles for FOX TFs 39
1.6.6 Molecular Profile of FOXM1 in Breast Cancer Subtypes 40
1.6.7 The Autoregulatory Loop of FOXM1 41
1.7 Endoplasmic Reticulum 44

1.7.1 ER Stress 44
1.7.2 UPR 45
1.7.3 Role of Protein Sensors 46
1.7.3.1 DDIT3 in UPR-mediated Apoptosis 46
1.8 DNA Methyltransferases (DNMTs) and Methylation of CpG islands 49
1.8.1 DNA methylation and breast cancer 50
1.8.2 Breast Cancer Treatment Through Regulation of Methylation 51
1.9 Breast Cancer Cell Lines 53
1.10 Overall Aims and Thesis Outline 54
2 Materials and Methods 56
2.1 Consumables 56
2.2 Cell Culture 56
2.2.1 Cell Lines 56
2.2.2 Cell Storage 57
2.2.3 Chemicals 57
2.3 Protein Analysis 58
2.3.1 Preparation of Total Protein Lysates 58
2.3.2 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blotting 58
2.3.2.1 Antibodies 59
2.4 RNA Analysis 59
2.4.1 RNA Extraction and Reverse Transcription 59
2.4.2 RT-qPCR Primer Design 60
2.4.3 RT-qPCR 63
2.5 DNA Analysis 63
2.5.1 Plasmid Constructs and Expression Vectors 63
2.5.2 DNA Quantification 64
2.5.3 Subcloning 64
2.5.3.1 Extraction and Purification of DNA 64
2.5.3.2 Bacteria Transformation and Culture 64
2.5.3.3 Plasmid Miniprep and Maxiprep 64
2.5.3.4 Polymerase Chain Reaction (PCR) 65
2.6 Transfection of Plasmid DNA and Expression Vectors 65
2.7 RNA interference (RNAi) 65

2.8 ChIP – qPCR 66
2.9 Co-immunoprecipitation (Co-IP) 67
2.10 Assay for transposase accessible chromatin with high throughput sequencing (ATAC–Seq) 68
2.11 Bioinformatics 68
2.11.1 Data analysis pipeline and visualisation 68
2.11.2 ChIP-Sequencing data collection 69
2.11.2.1 Cistrome Data Browser (Cistrome DB) Data Mining 69
2.11.3 RNA-Sequencing 70
2.11.4 ChIP-Sequencing Data Processing 70
2.11.5 ChIP-Seq Peaks Calling and Overlapping 70
2.11.6 Genomic Distribution of ChIP-Seq peaks 71
2.11.7 Motif Enrichment Analysis 71
2.11.8 Gene Ontology Analysis 71
2.11.9 Comparative Binding Analysis 72
2.11.10 ATAC-Seq Data Analysis Pipeline 72
2.11.11 ATAC-Seq Pre-Analysis: Quality Control and Alignment 72
2.11.11.1 Pre-alignment quality control 72
2.11.11.2 Alignment 73
2.11.11.3 Post-alignment processing and quality control 73
2.11.12 ATAC-Seq Core Analysis: Peak Calling 73
2.11.13 ATAC-Seq Further Analysis 73
2.11.14 BOX Plots 74
2.11.15 Genomic Distributions 74
2.11.16 Core Promoter Motif Analysis 74
2.11.17 Statistical Analysis 74
3 The functional differences between the FOXM1/ESR1 common transcriptional targets and FOXM1 specific targets in MCF-7 cell lines 76
3.1 Introduction 76
3.2 Aims and Objectives 79
3.3 Results 80
3.3.1 ChIP and ATAC sequencing to assess transcriptional profile and chromatin accessibility of target regions 80

3.3.2 FOXM1 and ESR1 share similar transcriptional binding profile in MCF-7 cell line, and there was little overlap for FOXM1 binding profile between MCF-7 and MDA-MB-231 82
3.3.3 FOXM1/ESR1 co-binding regions are cell line specific, whereas FOXM1/ESR1 negative regions are conserved across various cancer cell lines 83
3.3.4 FOXM1/ESR1 positive and ESR1 negative binding regions display varied genome-wide distribution traits 85
3.3.5 ChIP-seq profiles of epigenetic and enhancer markers in FOXM1/ESR1 positive and negative binding regions hint different genomic insights 86
3.3.6 ESR1 binding discriminates the biological pathways of FOXM1/ESR1 peaks and ESR1 negative peaks 89
3.3.7 GO enrichment analyses show fundamental biological process differences between FOXM1/ESR1 positive and ESR1 negative binding regions 92
3.3.8 FOXM1 knockdown impacts chromatin accessibility in MCF-7 cell line 95
3.3.9 Recruitment of FOXM1 to estrogen-responsive targets are dependent on ESR1 binding 97
3.3.10 Thiostrepton has greater effects in inhibiting ESR1 and ESR1-target gene expression compared to FDI-6 100
3.3.11 FDI-6 inhibits the expressions of ESR1 and ESR1-target genes in ESR1 mutant MCF-7 cell line 101
3.3.12 Combinatorial treatment of FDI-6 and fulvestrant on MCF-7 and MCF7-Y537S cell lines show promising results 103
3.4 Discussion 106
4 The relationship between FOXM1 and DDIT3 in breast cancer cell lines 115
4.1 Introduction 115
4.1.1 Breast cancer and UPR 115
4.1.2 Abnormal UPR Signalling in Breast Cancer 116
4.1.3 DDIT3 116
4.2 Aims and Objectives 118
4.3 Results 119
4.3.1 FOXM1 and DDIT3 bind to the promoter regions of *GRP78* and *GRP94* in MCF-7 cell line 119
4.3.2 Epirubicin treatment downregulates the expression of FOXM1 and DDIT3 121
4.3.3 FOXO3 activation represses FOXM1 and DDIT3 expression 123

4.3.4 The correlation of FOXM1 and DDIT3 with GRP78 and GRP94 were observed in breast cancer patient samples 125
4.3.5 The transcriptional binding profile of FOXM1 and DDIT3 in cancer cell lines 126
4.4 Discussion 131
5 The relationship between FOXM1 and DNMTs in breast cancer cell lines 135
5.1 Introduction 135
5.1.1 Epigenetics and Breast Cancers 135
5.1.2 FOXM1 and epigenetic regulation in breast cancer 135
5.2 Aims and Objectives 137
5.3 Results 138
5.3.1 FOXM1 and DNMT3 form protein complex in MCF-7 and MCF7-Epi^{R} breast cancer cell lines 138
5.3.2 FOXM1 does not interact with DNMT1 in both breast cancer MCF-7 and MCF7-Epi^{R} cell lines 139
5.3.3 Downregulation of FOXM1 does not affect the protein expression level of DNMTs in MCF-7 and MCF7-EpiR cell lines 141
5.3.4 The mRNA expression of *GATA3* and *HELLS* were impacted by the aberrant expression of FOXM1 in MCF-7 and MCF7-Epi^{R} cell lines 143
5.3.5 FOXM1 and DNMTs share similarities in transcriptional binding profiles 144
5.3.6 The known biological function of FOXM1 and DNMTs 146
5.3.7 The common binding targets and potential co-regulatory role of FOXM1 and DNMTs 148
5.3.8 FOXM1 in breast cancer basal subtype MDA-MB-231 cells and DNMTs have common binding targets and potential co-regulatory roles 151
5.4 Discussion 154
6 Final Discussion 159
6.1 Overview 159
6.2 Key limitations 160
6.2.1 Cancer cell lines model 160
6.2.2 Real-time qPCR 160
6.2.3 Western blotting 161
6.2.4 Co-immunoprecipitation 161
6.2.5 ChIP 162

6.2.6 Experimental replicates 162
6.3 Overall future directions 164
7 References 165
8 Appendix 175

List of Figures

Figure 1.1 ESR1 structural/functional domains and the distribution of the somatic LBD point mutations identified in ESR1-positive breast tumours. 22
Figure 1.2 Mechanisms of ESR1 action in breast cancer cells. 26
Figure 1.3 Modes of action of ESR1 and ESR1 ligands – oestradiol, tamoxifen and fulvestrant. 30
Figure 1.4 Structural differences of Forkhead transcription factors – FOXO, FOXM, FOXA, FOXC and FOXP factors. 35
Figure 1.5 Proposed mechanism of the antagonistic relationship between FOXO3A and FOXM1. 40
Figure 1.6 Mechanisms of FOX transcriptional factors activation. 42
Figure 1.7 FOX transcriptional factor regulated repression. 43
Figure 1.8 ER stress induced apoptosis. 48
Figure 1.9 Structural illustration and transcriptional binding of mammalian DNA 52
Figure 3.1 (A) ChIP-Sequencing (ChIP-Seq) and (B) ATAC-Sequencing (ATAC-Seq) workflow 81
Figure 3.2 The overlap between ChIP-Seq peaks of FOXM1 and ESR1. 82
Figure 3.3 Conservation of activity of FOXM1/ESR1-negative regions in the absence 84
Figure 3.4 Genome-wide distribution of Chip-seq FOXM1 peaks. 86
Figure 3.5 Chromatin landscape of FOXM1 bound regions. 88
Figure 3.6 ESR1 binding discriminates between the biological pathways of FOXM1/ESR1 peaks and FOXM1 peaks. 90
Figure 3.7 GO enrichment analysis shows housekeeping roles for FOXM1 specific target regions whereas FOXM1/ESR1 regions displays developmental roles. 94
Figure 3.8 The chromatin accessibility impact on FOXM1 and ESR1 bound regions upon FOXM1 silencing in MCF-7 cell line. 96
Figure 3.9 FOXM1 binding to the regions that are absent of ESR1 is not affected. 98
Figure 3.10 Recruitment of FOXM1 to estrogen-responsive targets are dependent on ESR1 binding. 99
Figure 3.11 Thiostrepton has greater effects in inhibiting ESR1 and ESR1-target gene expression compared to FDI-6. 101
Figure 3.12 FDI-6 inhibits the expressions of ESR1 and ESR1-target genes in ESR1 mutant MCF-7 cell line. 102
Figure 3.13 Effects of FDI-6 on growth in ESR1 mutant cells in combination with fulvestrant. 104
Figure 3.14 Proposed simplified models for enhancer- and promoter- regulatory properties for (left) FOXM1/ESR1 negative and (right) ESR1 positive target regions in MCF-7 cell line. 108
Figure 4.1 Differential binding effects seen for FOXM1 and DDIT3 binding on the promoter regions of GRP78 and GRP94. 120
Figure 4.2 The overexpression or downregulation of FOXM1 impacts the expression levels of GRP78 and GRP94 in the absence and presence of epirubicin. 122
Figure 4.3 The relationship between FOX TFs and DDIT3, GRP78 and GRP94. 124
Figure 4.4 The correlation between expression level of TFs, FOXM1 and DDIT3, and chaperone proteins, GRP78 and GRP94. 125
Figure 4.5 Microarray data of FOXM1 knockdown. 126

Figure 4.6 The potential transcriptional co-regulatory role of FOXM1 and DDIT3 in breast cancer subtypes. 130
Figure 5.1 FOXM1 interacts with DNMT3A/B at the protein-protein level shown by co-immunoprecipitation (co-IP) in native MCF-7 and epirubicin-resistant counterpart MCF7-EpiR cell lines. 139
Figure 5.2 FOXM1 and DNMT1 does not interact at the protein level from co–IP experiments in MCF7 and MCF7-Epi^{R} cell lines. 140
Figure 5.3 Expression analyses after downregulation of FOXM1 in MCF-7 and MCF7-Epi^{R} cell lines. 142
Figure 5.4 Genome-wide binding regions of FOXM1, DNMT1, DNMT3A, DNMT3B in cancer cell lines. 145
Figure 5.5 The transcriptional role of DNMTs on target genes. 147
Figure 5.6 The potential co-regulatory role of common targets between FOXM1 in breast carcinoma MCF-7 cells and DNMTs. 149
Figure 5.7 The GO analyses of the common binding targets between FOXM1 in MCF-7 cells and DNMTs. 150
Figure 5.8 The potential co-regulatory role of common targets between FOXM1 in breast 152
Figure 5.9 The GO analyses of the common binding targets between 153
Figure 5.10 FOXM1 differentially binds on DNMT1, 3A and 3B. 157

List of Tables

Table 2.1 Cell cultures used in this research. 57
Table 2.2 List of antibodies used for western blots. 59
Table 2.3 Oligonucleotides qPCR primer pair sequences. 60
Table 2.4 ChIP qPCR primer pair sequences. 62
Table 2.5 List of ChIP-seq data downloaded from Cistrome DB. 69

List of Appendix

Appendix 8.1 Recipes for gels and buffers of SDS-PAGE. 175
Appendix 8.2 Buffer recipes. 176

List of Abbreviations

4-OHT	4-hydroxytamoxifen
AF	Activating function
AI	Aromatase inhibitor
AKT	Protein kinase B
AMPK	AMP-activated protein kinase
AR	Androgen receptor
ATAC-seq	Assay for transposase accessible chromatin with high throughput sequencing
ATF6	Activating Transcription Factor 6
AURKA	Aurora Kinase A
BAG5	Bcl-2-associated anthanogene 5
BCL2	B-cell lymphoma-2
BSA	Bovine serum albumin
CDK1/2	Cyclin-dependent kinase 1/2
cDNA	Complementary DNA
C. elegans	Caenorhabditis elegans
CDS	Coding DNA sequence
ChIP	Chromatin immunoprecipitation
CHK2	Checkpoint kinase 2
CHOP	C/EBP homologous protein
Cistrome DB	Cistrome Data Browser
CML	chronic myelogenous leukemia
Co-IP	Co-immunoprecipitation
CTCF	CCCTC-binding factor
DBD	DNA binding domain

DDIT3	DNA damage-inducible transcript 3
DFS	Disease-free survival
DMEM	Dulbecco’s Modified Eagle Media
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA cytosine methyltransferase
DPE	Downstream promoter element
DSS	Double charcoal stripped serum
DTT	Dithiothreitol
E2	Estradiol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
eIF2	eukaryotic initiation factor 2
Epi	Epirubicin
ESR1	Estrogen receptor alpha
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ERE	Estrogen responsive element
ERK	Extracellular signal-regulated kinase
FCS	Foetal calf serum
FDR	False discovery rate
FGF	Fibroblast Growth Factor
FHRE	Forkhead responsive element
FOX	Forkhead box
Ful	Fulvestrant

G1/S	First growth phase/DNA synthesis
G2/M	Second growth phase/Mitosis
GADD153	DNA damage-inducible gene 153
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GEO	Gene Expression Omnibus
GO	Gene Ontology
GR	Glucocorticoid receptor
GRP78	glucose-regulated protein 78 kDa
h	Hour
H3K27Ac	Histone H3 lysine 27 acetylation
H3K4me1	Histone H3 lysine 4 methylation
HAT	Histone acetyltransferase
HCL	Hydrochloric acid
HDAC	Histone deacetylase
H12	Helix 12 subdomain
HELLS	SWI/SNF helicase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2	Human epidermal growth factor receptor 2
HNSCC	Head and neck squamous cell carcinomas
ICI 182,780 /Faslodex^{TM}	Fulvestrant
IGF	Insulin-like growth factor
IgG	Immunoglobulin G
IKKß	Inhibitor of nuclear factor kß
INR	Initiator
IRE1	Inositol Requiring Enzyme 1
JNK	c-Jun N-terminal Kinase

L19	*Ribosomal* protein L19
LB	Lysogeny broth
LBD	Ligand binding domain
LINE	Long interspersed nuclear element
LZ	Leucine zipper
MACS2	Model-based Analysis of ChIP-Seq 2
MAPK	Mitogen-activated protein kinases
MBC	Metastatic breast cancer
MCF7-Epi^{R}	Epirubicin resistant MCF-7 cell line
MCF7-Ful^{R}	Fulvestrant resistant MCF-7 cell line
ME	Mobile element
MEME	Multiple Expression motifs for Motif Elicitation
min	Minute
MMP	Matrix metalloproteases
mRNA	Messenger RNA
MTE	Motif Ten Element
$NaCl_2$	Sodium chloride
NCOA3	Nuclear receptor coactivator 3
NCoR	Nuclear receptor co-repressor
NaF	Sodium fluoride
Na_3VO_4	Sodium orthovanadate
ncExon	Non-coding exon
NES	Nuclear export sequence
NLS	Nuclear localisation sequence
NSC	Non-specific control
OSC	Ovarian somatic cell

PAGE	Polyacrylamide gel electrophoresis
PARP	Poly-ADP ribose polymerase
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline-tween
PDX	Patient driven xenograft
PERK/EIF2AK3	PKR-like ER Kinase
PI3K	Phosphoinositide-dependent kinase 1
PMSF	Phenylmethylsulfonyl fluoride
PR	Progesterone Receptor
PWM	Position-weight matrices
RNA Pol II	RNA Polymerase II
RASSF1A	Ras association domain family 1 isoform A
RNA	Ribonucleic acid
RNAi	RNA interference
RNA Pol II	RNA Polymerase II
RT	Room temperature
RT-qPCR	real time quantitative polymerase chain reaction
SA	Satellite repeats
SDS	Sodium dodecyl sulphate
sec	second
SERD	Selective estrogen degrader
SEM	standard error of the mean
SERM	Selective estrogen modulator
SINE	Short interspersed element
siRNA	Short interfering RNA
SMRT	Silencing mediator of retinoid and thyroid receptor

SR	Simple repeats
STI	Signal transduction inhibitor
TBS	Tris buffered saline
TBST	Tris buffered saline-tween
TCGS	The Cancer Genome Atlas
TDM1	Trastuzumab Emtansine
TET	Ten-eleven translocation enzyme
TF	Transcription factor
TGFα	Transforming growth factor α
TNBC	Triple negative breast cancer
TRIB3	Tribbles-related protein 3
TRD	Transactivation repressor domain
TRIS	Tris(hydroxymethyl)aminomethane
TSG	Tumour suppressor gene
TSS	Transcription start site
UPR	Unfolded protein response
UTR	Untranslated region
VEGF	Vascular endothelial growth factor
WB	Western blotting

1 Introduction

1.1 Breast Cancer

According to World Health Organization in 2021, the overall number of people diagnosed with cancer nearly doubled in the past two decades, from an estimated 10 million in 2000 to 19.3 million in 2020. Today, one in 5 people globally will develop cancer during their lifetime. Forecasts show that the number of people diagnosed with cancer will further increase in the future years and will be approximately 50% higher in 2040 than in 2020. The number of deaths from cancer has also increased, from 6.2 million in 2000 to 10 million in 2020. More than one out of every six deaths are due to cancer. Statistics released by the International Agency for Research on Cancer (IARC) in December 2020 illustrated that breast cancer (2,261,419 new cases in 2020, 11.7% of total new cases) has now overtaken lung cancer (2,206,771 new cases in 2020, 11.4% of total new cases) as the world's most commonly diagnosed cancer (1).

While modern lifestyles, such as unhealthy diets, insufficient physical activity, use of tobacco and abuse of alcohol consumption, are all attributes to the increasing cancer burden, a significant part can also be contributed by increasing longevity, as the risk of developing cancer increases with age. This reinforces the need to invest in both cancer prevention and cancer control, focusing on actionable cancers like breast, cervical, and childhood cancers as they all have high chance of cure if diagnosed early and treated appropriately. This also explains though breast cancer is the world's most diagnosed cancer, cancer death rate remains the highest for lung cancer. In contrast, with a high survival rate of 90%, breast cancer stays as the fifth highest cancer death after lung, colorectum, liver, and stomach (1).

Breast cancer arises when uncontrolled cell proliferation originated from breast tissues and some of these mutations are activated in response to estrogen (2). Similar to most cancer initiation and progression, breast cancer progression is a multistep process that involves genetic and epigenetic alterations that lead to the activation or silencing of a range of genes. Unfortunately, breast cancer is not a homogeneous disease, each individual, even at the same breast cancer stage, has different outcomes from the same treatment. Thus, it is incredibly challenging to predict the clinical behaviour of the tumour. Identifying novel prognostic genetic targets in breast cancer processes, such as initiation, progression, and drug resistance, gives rise to early detection for

accurate and effective treatments. Therefore, extensive research has been conducted to advance current diagnostic and therapeutic approaches to improve the quality of life and extend the lifespan of breast cancer patients.

1.1.1 Breast Cancer Heterogeneity and Subtypes

As mentioned, breast cancer is a heterogeneous set of diseases with distinct clinical characteristics, disease progressions, and treatment responses. There are ample classifiers determined by extrinsic and intrinsic factors to distinguish breast cancer subtypes (3-9). Intrinsic factors include classical histological, immunopathological classifications, and molecular subtypes; the extrinsic factor includes tumour microenvironmental features. Established molecular portraits for breast cancer subtypes are essential for designing prognostic genomic tools to the administration of treatment options, plus to predict overall and relapse-free survival and the development of therapy resistance. In this study, we will focus on molecular subtypes that is mainly separated into five subtypes, including luminal A (HER2-, low levels of Ki67, ESR1+ and/or PR+), luminal B (HER-, high levels of Ki67, ESR1+ and/or PR+), basal-like breast cancer (HER2-, ESR1-, PR-), HER2 overexpression (HER2+, ESR1-, PR-) and normal-like (HER2-, low level of Ki67, ESR1+, PR+).

Breast cancer molecular subtypes can be characterised by the presence of hormone receptors that are responsive to ligands, such as estrogen and progesterone (or both), namely the estrogen receptor (ESR1)-positive or progesterone receptor (PR)-positive breast cancer respectively. ESR1s are expressed in 75% of breast cancers overall and is the main driver of most breast cancers. It is frequently associated with post-menopausal women and there is a 99% survival rate over ten years. Hormone-sensitive breast cancer has a strong correlation with lower tumour grade; lower levels of amplification of the human epidermal growth factor receptor 2 (*HER2*) oncogene and loss of *TP53* tumour suppressor gene; expression of PR, soft tissue and bone metastases and slower rates of disease recurrence. The most common receptors that are over-expressed in breast cancers are members of the epidermal growth factor receptor (EGFR) family of receptor tyrosine kinases: EGFR and HER2 are over-expressed in roughly 40% and 25% of breast cancers accordingly and are responsible for aggressive tumour phenotype and poor prognostic outcome. Lastly, triple-negative

breast cancer (TNBC) is characterised by lack of expression of both ESR1 and PR plus HER2 protein and is frequently related to the negative prognostic outcome as no treatment is available as of now for TNBC.

1.2 ESR1

ESR1 is an important transcription factor (TF) for the development and function of female reproductive organs, including the mammary gland, bone homeostasis, cardiovascular remodelling and brain functions (10). However, deregulation of estrogen and ESR1 signalling pathways also promote mammary, ovarian and endometrial tumorigenesis (10).

The ESR1 is capable of binding a diverse set of ligands that are broadly categorized as agonists or antagonists, depending on their abilities to activate or suppress transcriptional responsiveness (11). Structurally, ESR1 comprises a central DNA binding domain (DBD) flanked by two autonomous transcriptional activation domains, AF-1 and AF-2 (Figure 1.1). Particularly, AF-2 is in the ligand-binding domain (LBD) and is ligand-dependent. AF-2 domain also modulates interactions with co-regulators that mediate ESR1 transcriptional activity. When ESR1 bind to the agonistic estrogen, ligand binding triggers helix 12 subdomain (H12) to fold itself over the carboxy-terminal hydrophobic pocket. However, when ESR1 binds to partial agonists or antagonists such as tamoxifen, fulvestrant, or raloxifene, the “bulky” side chain of the molecules hinders H12 from positioning itself in an agonist bound position over the pocket, therefore antagonising co-activator binding to the ESR1s. Tamoxifen-regulated inhibition on ESR1 allow the recruitment of co-repressors such as nuclear receptor co-repressor (NCoR) and silencing mediator of retinoid and thyroid receptor (SMRT).

1.2.1 Mutations in ESR1s in breast cancer tumours

ESR1 is the main biomarker of response to endocrine therapy in hormonal cancers. Based on the ESR1 status, breast tumours can be classified as ESR1-positive and ESR1-negative. ESR1-positive tumours are responsive to endocrine therapies, also sensitive to CDK4/6 inhibitors (10). However, ESR1-negative tumours are more

aggressive and metastatic. The absence of a specific therapeutic target poses a challenge to treat ESR1-negative breast cancer.

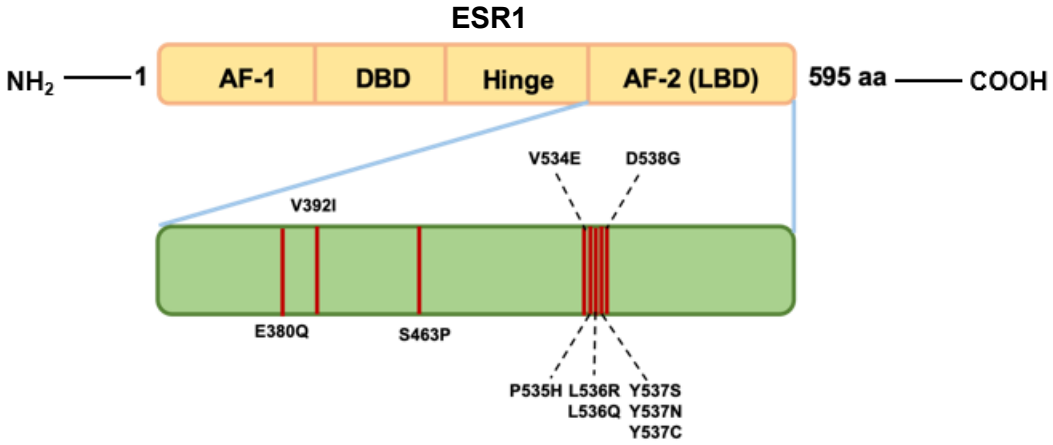


Figure 1.1 ESR1 structural/functional domains and the distribution of the somatic LBD point mutations identified in ESR1-positive breast tumours.

Each mutation is represented by a red bar at the specific ER protein position. AF-1 and AF-2, Activating Factor 1 and 2. DBD, DNA Binding Domain. LBD, Ligand Binding Domain.

Additionally, ESR1 expression in breast cancer is dynamic and reversible (10). Studies indicated the loss of ESR1 expression or ESR1 mutations as potential explanations for the development of acquired resistance (12) (Figure 1.1). However, loss of ESR1 expression happens in only 15 – 20% of resistant breast cancer cases (13) and less than 1% of ESR1-positive tumours have ESR1 mutations (14).

Complete deprivation of ESR1 is rare and observed in only a minority of 10 – 20% of metastatic breast cancers (15). However, frequent ESR1 mutations in metastatic tumours after several lines of endocrine therapy has led to studies with deep genomic sequencing of ESR1 in breast cancer cell lines with acquired resistance (16). Such mutations are infrequent in primary tumours prior to cancer therapy, but they are common in metastatic breast tumours in which it has spread to secondary locations (17). These mutations in ESR1 can highly impair the effectiveness of standard endocrine therapeutic drugs (18). The observation that many tumours with acquired endocrine resistance are still responsive to additional lines of endocrine therapies indicates that ESR1 remains to play an important role, even though a deregulated one

in these tumours. Recurrent activating missense mutations clustered within the ESR1 LBD have been indicated by several studies and identified in an ample proportion of endocrine-resistant metastatic ESR1-positive breast cancers (Figure 1.1).

Targeted next-generation sequencing found 14 patients with ESR1 LBD mutation among the 80 patients studied (17.5%). Moreover, sequencing of 183 primary tumours of a subset of the BOLERO-2 patients with metastatic disease detected mutations in six of these primary tumours (3.3%) (19). Nonetheless, in another study, 76 metastatic ESR1-positive tumours were sequenced and 11 ESR1 LBD mutations were identified (14.4%) (20). Lastly, a different study has detected four ESR1 LBD mutations among 37 ESR1-positive matched metastatic and primary tumour samples (10.8%) (20). In the limited available matched primary tumours without matched metastatic samples, the ESR1 mutations were not identified – 203 primary tumours were screened by dd-PCR and a relatively high percentage of these tumour samples gained LBD ESR1 mutations (the Y537N mutation was detected in 12% of the tumours, whereas Y537S in 5%, and the D538G mutation in 2%), another study has detected ESR1 mutations in 2.5% of 270 primary tumours and 20% of 55 ESR1-positive metastatic tissue samples.

1.2.2 Biology of ESR1 LBD Mutations: Y537 and D538 mutation “hotspots”

Recurrent point mutations of ESR1 that are associated with metastasis with ESR1-positive breast cancer origin were all detected within the LBD of the receptor – usually in Y537 and D538 amino acid residues within H12 subdomain, a key structural location of AF2 (15) (Figure 1.1). These mutations alter the conformation of the receptor, consequently limiting the therapeutic lifespan of standard cancer treatments and leading to treatment resistance (18). Therefore, endocrine resistance is commonly observed and poses a critical clinical concern in metastatic breast cancer as it is incurable to this date, with a median 5-year survival rates of less than 25% (16).

In breast cancer cells with wild type ESR1, when estrogen binds to the receptor, the H12 subdomain folds and holds the hormone in the ligand-binding pocket of the receptor. This active conformational change reveals an interacting surface for coactivators, such as Nuclear Receptor Coactivator 3 (NCOA3), which are proteins that cooperate with the receptor to activate gene transcription (21). Y537S and D538G

mutations stabilise the unliganded receptors, alter the H12 to consistently adopt an active conformation, hence making ESR1 activated even in the absence of estrogen (18). Typically, the binding of selective estrogen receptor modulators (SERMs) modifies the shape of the H12 in a way that inhibits its interaction with a coactivator and encourages co-repressor binding, leading to the suppression of ESR1 downstream transcriptional activity.

The LBD ESR1 mutants in human metastatic breast cancer tumours harness a growth advantage in estrogen-independent circumstances and relative resistance to tamoxifen and fulvestrant. Early structural studies have indicated that the Y537S mutation in ESR1 stabilises H12 in the agonistic conformation – like wild-type ESR1 when bound to estrogen. D538G also stabilises H12 in the agonistic conformation in the absence of estrogen, although less potently in relative to Y537S mutation (18). Notably, antagonistic ligand tamoxifen and fulvestrant bound to D538G and Y537S mutants showed that these mutations resulted in an aberrant agonistic conformation, contributing to resistance in the endocrine drug. Additionally, these mutants are less dependent on estrogen. Studies also showed upregulated co-activator binding to mutant ESR1 when compared to wild-type ESR1 under ligand-independent circumstances or in the presence of tamoxifen or fulvestrant (17). Additionally, breast cancer cells that carry the D538G mutation showed an increased migratory capacity, indicating an enhanced invasive phenotype, which could elucidate the selective growth advantage of metastatic cells with these mutations. As mentioned, gene expression and signalling network analysis showed that the mutations activate a distinctive set of genes and heighten the crosstalk with Insulin-like growth factor (IGF) signalling (22).

1.3 Mechanism of Action of Estrogen and ESR1

ESR1 activity and signalling are mediated by a plethora of pathways which could also lead to resistance to ESR1-targeted therapies, particularly when the pathways initiate aberrant activity in the context of cancer (12, 23). In Figure 1.2, there are four main ESR1 signalling pathways for ESR1-positive breast cancer mechanisms. In Pathway 1, it is the classic function of ESR1 in its nuclear/genomic role to regulate the expression of genes important for normal cellular function, tumour growth, and tumour

survival. ESR1 mediates the expression of hundreds of genes, some by upregulation of expression and the rest by downregulation (24, 25). Upon binding to estrogen, ESR1 dimerises and forms the correct conformation for a complex of co-activators and co-repressors to specific sites on DNA. ESR1 can also bind to other TFs such as AP-1 and SP-1 on their specific sites on DNA, thereby serving as a co-regulator (26). Co-regulators provide as a fine-tuning mechanism by enhancing or depleting the transcriptional activity of the receptor (21). Among the notable co-regulators identified in multiple cancer types, *AIB1* (also known as *SRC-3*) is a gene overexpressed in two-thirds of all breast cancer subtypes (27). Additionally, *AIB1* gene amplification has been associated with tamoxifen resistance (28). In Pathway 2, another branch of an ESR1 signalling pathway is regulated by membrane receptor tyrosine kinases, such as EGFR, HER2, and IGF1-R (29, 30). These membrane kinases trigger signalling pathways that eventually lead to phosphorylation of ESR1, as well as its co-activators and co-repressors at several sites to affect their specific functions. This activation of ESR1 by growth factor receptor signalling is occasionally defined as the ligand-independent receptor activation. Crosstalk between the growth factor receptor and ESR1 pathways has been established through multiple different mechanisms as well. Estrogen can increase the expression of ligands including transforming growth factor α (TGFα) and IGF1 which result in the activation of the growth factor receptor pathway. Moreover, estrogen signalling decreases the expression of EGFR and HER2 while upregulates the expression of IGF1-R. Activation of the Phosphoinositide-dependent kinase 1/Protein kinase B (PI3K)/AKT and the p42/44 mitogen-activated protein kinases (MAPK) pathways by these receptors result in the downregulation of the expression of ESR1 and PR. Therefore, while these receptor tyrosine kinases can activate the transcriptional function of ESR1, they may also reduce estrogen dependence by decreasing the expression of ESR1, probably contributing to the relative resistance to endocrine therapies in tumours amplified for *HER2* (31, 32).

In Pathway 3, studies also implicate that ESR1 may work by non-nuclear/non-genomic mechanism. Low levels of ESR1 have been detected outside the nucleus in the membrane, cytoplasm, or even in the mitochondria (33), however, the exact location for this receptor remains debatable. Part of the non-genomic role of estrogen appears to be too rapid for a transcriptional effect to activate growth factor receptor signalling, covering the PI3K/AKT and the Ras/p42,44 MAPK pathways. Therefore, ESR1 –

through this non-genomic role – can manipulate the expression of genes usually modulated by growth factors.

Lastly, Pathway 4 shows the stress kinase pathway via p38 and c-Jun N-terminal Kinase (JNK) can also regulate ESR1 function by phosphorylation of ESR1 and its co-regulators, resulting in increased invasiveness in breast tumours (34). Additionally, tissue microenvironment and its related integrin signalling may play a similar role (35).

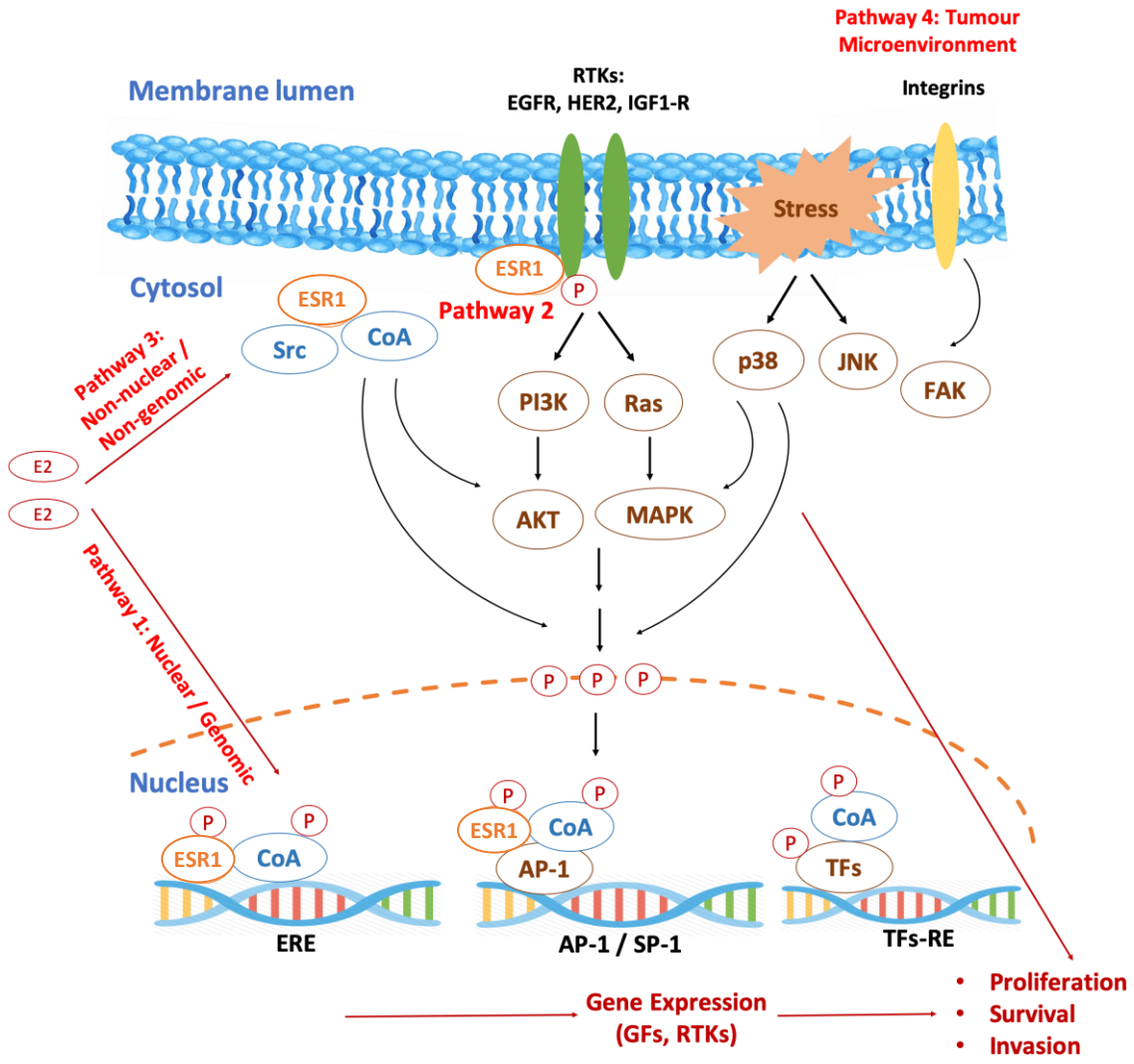


Figure 1.2 Mechanisms of ESR1 action in breast cancer cells.

Four known pathways of ESR1, genomic and non-genomic, in the regulation of estrogen response elements (EREs) on target gene promoter regions.

1.4 ER Ligands and Endocrine Therapies

ESR1 belongs to the nuclear hormone receptor family of TFs, the same as the glucocorticoid receptor (GR), PR and androgen receptor (AR). Particularly, oestradiol (E2) is the default ligand of the ESR1 that binds and stimulates receptor dimerisation and consequent activation of two activation domains (AF-1 and AF-2) (Figure 1.1). The activated receptor complex undergoes nuclear localisation and then binds to specific DNA sequences – the estrogen response elements (EREs) – situated in the upstream regulatory regions of estrogen-controlled genes (Figure 1.3). Upon binding to these sequences, transcriptional co-activators or co-repressors are recruited to the complex, causing the activation or inhibition of gene expression (12, 36). The management of both early and advanced breast cancers has mainly focused on manipulating the response of hormone-sensitive, ESR1 positive tumours to circulating estrogens. Theoretically, this regulation of estrogen-activated process can be conducted at a variety of levels: adjusting circulating levels of estrogens, altering ligand binding to the ESR1 and shifting downstream responses as an outcome of ligand binding to the ESR1. These different approaches have all led to the development of a variety of well-known and commonly used agents as shown in Figure 1.3.

1.4.1 Tamoxifen, anti-estrogen

Estrogen and ESR1 are major drivers in breast cancer progression; therefore estrogen/ESR1 pathway has been targeted for decades to treat patients with ESR1-positive breast cancer. The non-steroidal SERMs have been used to partially suppress tumour growth in ESR1-positive breast cancers. The net effect of these SERM agents is a balance between agonist and antagonist activity. The most notable drug in this category is tamoxifen (Nolvadex^{TM}). Tamoxifen was the first type of SERMs approved for ESR1-positive metastatic breast cancer-preventing recurrences by roughly 40 – 50% as stated on BreastCancer.org. Tamoxifen has been around for about 60 years and it started off as the first-choice endocrine therapy for the clinical management of postmenopausal women with hormone-sensitive early and advanced breast cancer. As opposed to estrogen, tamoxifen binds to ESR1 and dimerisation occurs but is unable to activate the AF2 domain (Figure 1.3). Eventually, this causes incomplete

inhibition of transcription and therefore, while tamoxifen releases anti-estrogen activity, the remaining active AF1 domain allows a degree of gene expression to occur, relaying tamoxifen's partial agonist activity.

Despite the significant reductions to both tumour recurrence rates and overall death rates in patients treated with tamoxifen, there exists multiple significant clinical challenges that may restrict the long-term use of tamoxifen. The issues revolving around the use of tamoxifen mainly stemmed from the partial agonist activity tamoxifen treatment induces. It can be classified as either an agonist or an antagonist depending on the location it is metabolised, the species studied, the target organ evaluated, and the endpoint measured (2). Factors like binding partners, the gene promoter responsive unit, and the different molecules that come into contact will predict the effect of tamoxifen. There is a wide spectrum of side-effects that comes with this agonist activity of tamoxifen treatment, such as the benefits towards bone mineral and blood lipid profiles and the undesirable side-effects on the endometrium and uterus plus the increase of thromboembolic complication risk in some patients. In women who displayed high risks for breast cancer and were managed under tamoxifen treatment, their chances of acquiring the disease decreased by 45%, but they obtained a higher risk of endometrial cancer and a decrease in bone fractures. Finally, the development of tamoxifen resistance also hinders the prolonged efficacy of the drug in breast cancer patients (15).

1.4.2 Aromatase Inhibitors (AIs)

To tackle the negative agonistic trophic effects (side-effects that arise from hormones that activate the release of other hormones, usually have a hypertrophic property), while refining or maintaining the advantages that tamoxifen offers, the third-generation AIs, such as anastrozole (Arimidex^{TM}), letrozole, and exemestane, exercise their clinical effects by inhibiting the peripheral production of estrogens and therefore significantly impact on total circulating estrogens. Aromatase enzyme converts androgen substrates to estrogen, and AIs inhibit estrogen circulation in tumours as they prevent the biosynthesis of androgens to estrogens through suppression of the aromatase enzyme. Understandably, AIs cannot abrogate the trophic actions of dietary or environmental estrogens.

1.4.3 Fulvestrant, the ‘pure’ anti-estrogens

The third class of agents under the limelight are known as selective ESR1 downregulators (SERDs), covering ICI 182,780, known as fulvestrant (Faslodex^{TM}), which works by completely restricting the effects of estrogens. While similar as tamoxifen, fulvestrant binds to the ESR1, however, the receptor dimerisation is impaired, receptor degradation is accelerated by the increase of receptor protein turnover. Additionally, both AF1 and AF2 remain inactive (Figure 1.3). Disrupted nuclear localisation of ligand-ESR1 complexes and increased proteolytic degradation of the ESR1 lead to complete attenuation of ESR1-regulated transcription, causing complete inhibition of the expression of estrogen-dependent genes. Therefore, fulvestrant is termed as a ‘pure’ anti-estrogen for its lack of agonist effects, as opposed to SERMs that may act as antagonist in one tissue type but as an agonist in another (37). It is predicted that fulvestrant offers a longer duration of treatment response with a decreased risk of unwanted complications that have been associated with tamoxifen treatment in certain patients. Pre-clinical trials have indicated the utilisation of fulvestrant after a breast cancer patient develops resistance to tamoxifen (12).

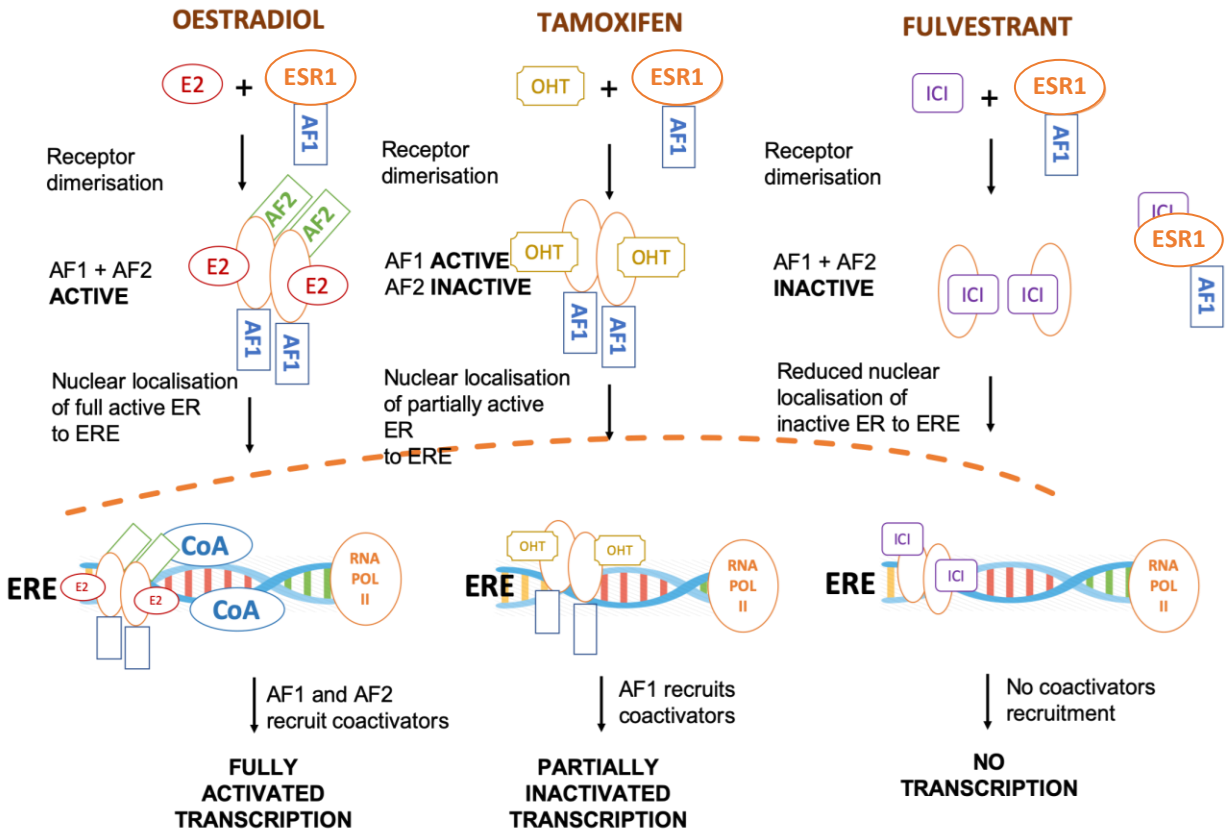


Figure 1.3 Modes of action of ESR1 and ESR1 ligands – oestradiol, tamoxifen and fulvestrant.

Each anti-estrogen differs in the mechanism at the level of the E**S**R**1**. Molecular events that typically happen include receptor dimerisation, and association with distinct strogen response elements in the estrogen-targeted genes, followed by recruitment of other proteins to form a transcriptional complex, resulting in either transcriptional activation or inhibition. Oestradiol (E2) binds to E**S**R**1** and leads to receptor dimerization and fully activated transcriptions. Tamoxifen (OHT) leads to partial activation of AF1 and inactivation of AF2 and results in impaired gene transcriptions. Fulvestrant (ICI) binds to E**S**R**1** and restricts receptor dimerization, therefore inhibits gene transcription. E**S**R**1**, estrogen receptor alpha. E**RE**, estrogen response elements. AF, activating function. RNA Pol II, RNA Polymerase II. E2, oestradiol. OHT, tamoxifen. ICI, fulvestrant.

1.5 Breast Cancer Treatments

The distinctive heterogeneity in breast cancer poses significant challenges to effective treatment for the disease. Therefore, it is imaginable that breast cancer patients would benefit drastically from personalised and targeted therapies to treat their specific types of cancer. The advancement in whole-genomic sequencing has improved the process of identifying specific mutations in cancer of any individuals. Nevertheless, the

interpretation and complete removal of the disease from sequencing results remains a challenge. Five main types of treatment strategies are available for breast cancer, comprising of surgery, radiotherapy, chemotherapy, biological therapy, and hormone/endocrine therapy (2). Most of the time, these treatment strategies are not categorically exclusive but are used in combinations. The treatment arrangement is dependent on type, size, location, and stage of cancer. Additionally, personal factors like age, health, and if female, the menopausal status are also considered. Surgery is the chief treatment for breast cancer and eradicates the mass. The type of surgery varies from breast-conserving to total mastectomy, depending on the location and size of the tumour, and other factors stated above. Ultimately, it is down to the patient to decide what option to go for. Radiotherapy utilises high energy beams to remove cancer cells. It is frequently combined with lumpectomy to eliminate potential hidden cancer cells. It can also be used pre-surgery to reduce tumour size. Additionally, chemotherapy agents are typically used in breast cancer treatments, they include alkylating agents like cyclophosphamide, anthracyclines such as doxorubicin, anti-metabolites such as methotrexate and fluorouracil, and taxanes like paclitaxel and docetaxel (38). Not forgetting, biological therapy, including the widely popular Herceptin and hormone/endocrine therapy, is often administered.

1.5.1 Endocrine Therapy Resistance

Resistance to hormone therapies is a tricky challenge within hormone-sensitive breast cancers even though ESR1- and PR-targeted therapies have proven to be very effective, improving the quality of life of hormone-sensitive breast cancer patients. There are multiple pathways associated with endocrine resistance mechanisms: HER tyrosine kinase receptor family; receptors for insulin/IGF1, Fibroblast Growth Factor (FGF), Vascular endothelial growth factor (VEGF), SRC, AKT, stress-related kinases – each play a decisive role in contributing to endocrine therapy resistance when their cognate ligands are amplified or over-expressed. Crosstalk between the ESR1 and growth factor receptor signalling with hyperactivation of the PI3K pathway has also been connected to endocrine resistance, as illustrated in Pathway 2 of Figure 1.2. Additionally, nuclear receptors and the ARs may also act as alternative growth triggers by post-translational modification, allowing the bypass of ESR1 inhibition. Co-targeting

the EGFR and HER2 pathways seems to be the most hopeful way forward in preventing endocrine resistance.

1.5.2 ESR1-negative Breast Cancer Treatments

For hormone independent breast cancers, biology therapy options are available. The HER2 protein is the most common over-expressed receptor in breast cancer and is recognised as a relevant biomarker for treatment. The recombinant antibody trastuzumab (Herceptin^{TM}) antagonises HER2 and was the first drug approved by the FDA in 1998 for the treatment of HER2-positive breast cancer. Other similar agents developed are pertuzumab and lapatinib with frequent development of resistance with significant side-effects in patients. The conjugated monoclonal antibody trastuzumab emtansine (TDM1) may be used in HER2-positive breast cancers as trastuzumab effectively delivers the DM1 drug, a microtubule inhibitor, directly into the breast cancer cells to suppress growth (28).

For triple negative breast cancers (TNBCs), they are notoriously challenging to treat as they lack hormone receptors and HER2 – fortunately, they may respond to agents like Poly (ADP-Ribose) Polymerase 1 (PARP1) inhibitors and may have HER1 as a potential target (39, 40). Additionally, some subtypes of TNBC may be EGFR inhibition sensitive. The typical routine of TNBC treatment is taxol derivatives and anthracycline chemotherapy until more “druggable” targets are developed. Targeted therapies against VEGF have also been developed, for example, bevacizumab has proven effective in the treatment of advanced metastatic breast cancer when used in combination with paclitaxel or docetaxel. Other than that, inhibitors of downstream signalling pathways like PI3K/AKT/mTOR and RAS/MEK/ERK are also available as treatment options as well as agents directed against other tyrosine kinases like SRC, IGF/IGF-receptor, PARP inhibitors and matrix metalloproteases (MMPs) which are associated with cancer cell invasion and metastasis. Compensatory survival pathways, for example, increased PI3K signalling, receptor tyrosine kinase signalling outside the ErbB/HER family and involvement of other HER receptors, may all be involved in the development of alternative molecular pathways responsible for the development of therapeutic resistance in breast cancer cells.

1.6 Forkhead Transcription Factor Superfamily

The first *forkhead* gene was initially identified in a random mutagenesis screen performed in *Drosophila melanogangster* (41). In *Drosophila*, it was shown that forkhead is indispensable for normal gut development and that its absence leads to a characteristic 'forked head' appearance resulted from the homeotic transformation of the foregut into a head structure, hence this group of TFs are termed FOX proteins and were later identified in a range of organisms, covering from yeasts to humans.

The number of *FOX* genes currently identified varies extensively among different organisms. Human and mouse both have 44, *Drosophila* 11, *Caenorhabditis elegans* 15, and *Xenopus* 45. *Xenopus* models contributed vastly to the initial discovery of *Fox* expression patterns in early embryogenesis. In mammals, FOX TFs are bucketed in subclasses based on sequence homology within and outside of the forkhead box. Often, homozygous deletion of one *FOX* gene may lead to embryonic or perinatal lethality and, in humans, mutations or abnormalities of *FOX* gene regulation are connected to developmental disorders and diseases such as cancer, Parkinson's disease, autism spectrum disorder, ocular abnormalities, deficiency in immune regulation and function, and defect in language acquisition.

1.6.1 Protein domains, expression patterns and post-translational modifications

FOX TFs bind to identical DNA sequences, though with different affinities, owing to their highly conserved DNA-binding domain. However, members of this large gene family have separate roles – how? The different sequences outside of the conserved DNA-binding domain probably distinguish the function of these proteins, as do their diverse temporal and spatial gene activation patterns. For instance, the binding domains of FOXA TFs have structural homology to linker histones H1 and H5 (42-44). This characteristic permits FOXA TFs to access closed chromatin, thereby resulting in the recruitment of alternative histones and assisting the downstream binding of other TFs at nearby sites.

FOXM1 is typically only expressed during the cell cycle, which explains it is famously known for its crucial role in the regulation of cell cycle (45-47). Its forkhead DNA-binding domain shares merely 39% sequence homology with that of FOXA1 and FOXM1 has a lower affinity for the forkhead consensus binding site than the other

FOX proteins (48). There are two mechanisms for FOXM1 transcriptional regulation. Firstly, FOXM1 binds to proteins in the DREAM complex – a larger assembly of proteins that restricts the transcription of proliferative target genes in the inactive state but activates the expression of similar target genes during the cell cycle rather than directly interacting with the promoter region of its proliferative target genes. Secondly, Sanders et al. have shown the direct binding of FOXM1 to its target genes (49). The activity of FOXM1 is also tightly regulated by several post-transcriptional modifications (50, 51): (a) FOXM1 is typically secluded from the nucleus, but phosphorylation by MAPK triggers its translocation; (b) an intramolecular repressive domain suppresses the binding of FOXM1 to DNA region until phosphorylation by CHK2; (c) FOXM1 only recruits co-factors for gene *trans*-activation when phosphorylated by CDK1/2; and (d) FOXM1 is targeted to the proteasome for degradation when phosphorylated by Extracellular signal-regulated kinase (ERK) 1/2 (Figure 1.4).

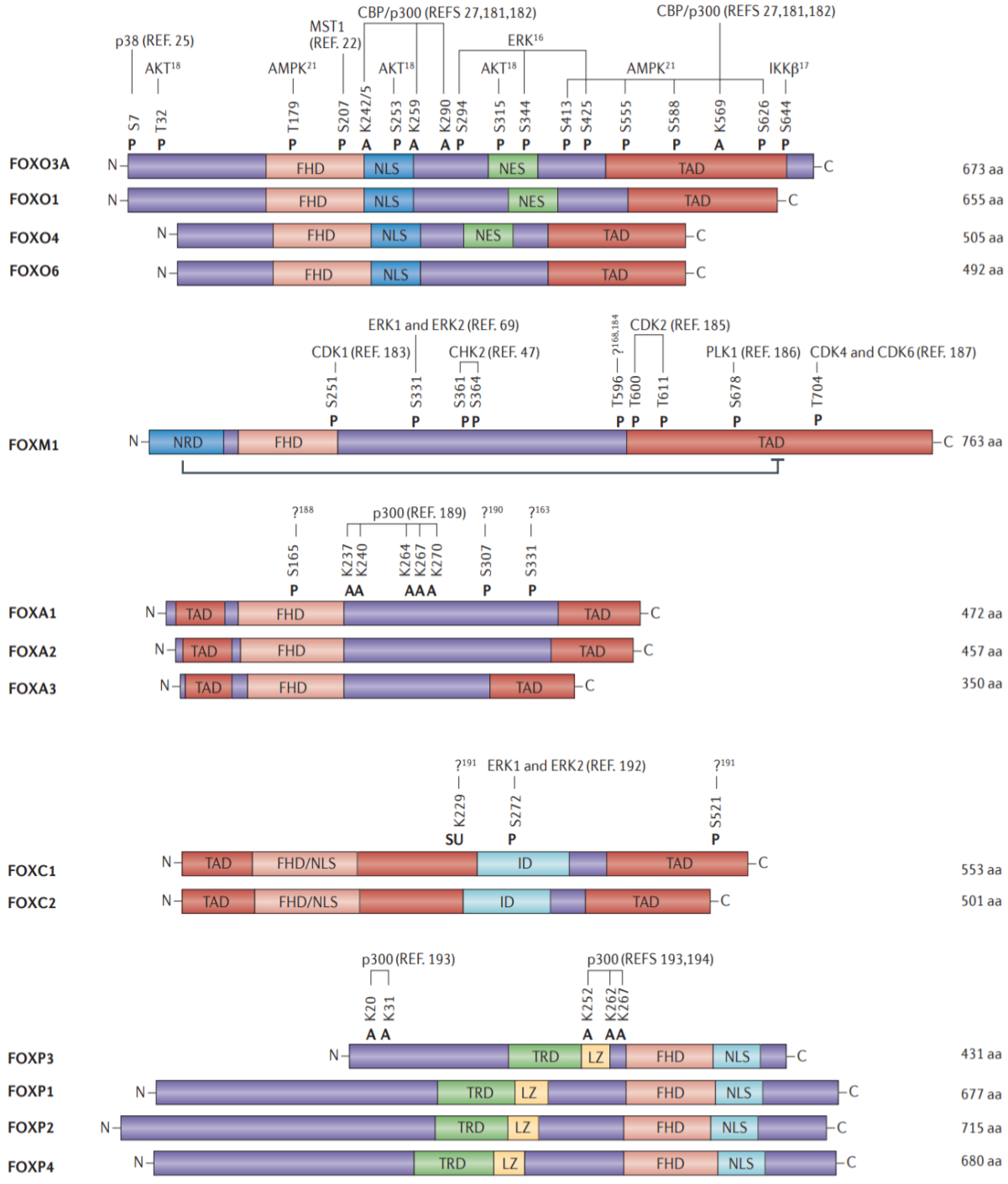


Figure 1.4 Structural differences of Forkhead transcription factors – FOXO, FOXM, FOXA, FOXC and FOXP factors.

Illustration showing the different domains that characterise forkhead box (FOX) proteins. All forkhead TFs share a highly homologous forkhead DNA-binding domain which binds to conserved sequences in the target genes. This domain can be localised to the amino-terminal and carboxy-terminal positions. The post-translational modifications are also shown here. A, acetylation site; aa, amino acids; AMPK, AMP-activated protein kinase; CDK, cyclin-dependent kinase; ID, inhibitory domain; IKKß, inhibitor of nuclear factor kß; LZ, leucine zipper; MST1, mammalian

Ste20-like kinase; NES, nuclear export sequence; NLS, nuclear localisation sequence; NRD, N-terminal repressor domain, P, phosphorylation site; PLK1, polo-like kinase 1; SU, sumoylation site; TAD, transactivation domain; TRD, transcriptional repressor domain. Adapted from Lam et al. 2013 (52).

The activity of FOXO proteins – a group of proteins that suppresses the cell cycle and regulates lifespan and metabolism – is regulated post-translationally. Phosphorylation by AKT leads to the nuclear exclusion of FOXO1 and isolation away from the DNA region by the protein chaperone 14-3-3. Contrasting to other FOX TFs, FOXO factors carry both a nuclear export signal (NES) and a nuclear localisation signal (NLS). FOXO proteins facilitate the activation of the transcription of their target genes by recruiting the SWI/SNF chromatin remodelling complex. The transcriptional target genes of FOXO are associated with antioxidants, cell cycle inhibitors, and metabolic genes; Generally, FOXO proteins act oppositely to FOXM1 (Figure 1.5).

Most TFs are transcriptional activators; however, FOXPs, have dual activator and repressor roles. FOXP family members bind as homo- or heterodimers though their leucine zipper and zinc finger domains, and this interaction is necessary for FOXP group of members to inhibit their target genes, probably through a conformational change that occurs upon binding that allows the recruitment of co-repressors such as histone deacetylases (HDACs), lysine acetyltransferase TIP60 (Kat5), and the SMRT complex. Like FOXM1, the FOXP group has been indicated a role in indirect target gene binding though co-factors to DNA.

1.6.2 FOXA family

In mammalian cells, FOXA1, -2, and -3 display overlapping but different expression patterns in a range of developing and mature tissues (100). FOXA2 mutant study indicates FOXA2 on its own is not essential for normal liver differentiation, though defects in lung morphogenesis are observed. Additionally, serious defects in pancreatic islet formation and alpha and beta-cell maturation that results in changes in glucose homeostasis are investigated (53). Notably, mice expressing an activated form of Foxa2 in neurons demonstrate a stimulated appetite, implying a potential of FOXA2 in metabolic regulation, such as food intake and energy output.

Conditional deletion of genes encoding FOXA TFs showed little specificity in the role of each FOXA family member in the liver (50, 53, 54). Gross pancreatic morphology was also unimpacted when only a single FOXA TF was absent, suggesting that either FOXA family members are replaceable in these organs or, alternatively, that they can compensate for each other. Both FOXA1 and FOXA2 must be inhibited to eliminate the expression of Muc2, which encodes a protein crucial for intestinal function, and both FOXA1 and FOXA2 are upregulated in FOXA3 null livers.

1.6.3 FOXM1

FOXM1 was identified nearly at the same time by three groups, therefore it was originally named three different terms: Trident, HNF3, and WIN (50, 55). All three groups found FOXM1 by homology to FOXA proteins. As previously stated, FOXM1 is vastly known for its role in promoting cell cycle. Its direct transcriptional targets are involved in both the G1/S and G2/M phases, as well as karyokinesis and cytokinesis. Accordingly, FOXM1 expression is largely kept within actively replicating cells; its expression is, therefore, wide in embryos and decreases in differentiated tissue as well as with aging (46).

The role of FOXM1 in *Drosophila* and *C. elegans* might be substituted by other forkhead TFs, usually through the increased replications in the absence of FOXO homologues (52). In humans, splice variants of *FOXM1* produce four different FOXM1 proteins: FOXM1A, FOXM1B, FOXM1C and FOXM1D (56). FOXM1A is a transcriptional repressor whereas the other isoforms are all transcriptional activators. In mice, only one FOXM1 protein splice variant is present; it is almost identical to human FOXM1B and is likewise a *trans*-activator.

FOXM1 is one of the most widely researched forkhead TFs because of its prominence in cancer (57). It is over-expressed in a range of tumour types and is linked to increased tumour aggressiveness. Increased expression of FOXM1B does not randomly initiate tumours in mice models. However, over-expression of FOXM1B results in larger, more vascularised, and more metastatic tumours. Increased expression of Foxm1b in lung accelerates tumour growth in mice with activated Kras expression while the absence of Foxm1b suppresses lung tumorigenesis regulated by Kras (58).

FOXM1 also seems to be associated with aging (59). The level of FOXM1 decreases with age in several tissues and this decline is connected to a decrease in both baseline proliferation and replicative ability. The re-activation of FOXM1 activity in aged hepatocytes and pancreatic beta cells was able to replenish replication in these tissues. As is the case for other proteins involved in the cell cycle, the impacts of activating or inhibiting FOXM1 activity must be carefully evaluated. Drug treatments are currently not targeted to any specific tissue, activating FOXM1 endogenously to replenish the replicative potential in a tissue of interest without aiming it to a specific cell type could potentially cause existing tumours to become more prone to growth or metastasis. Similarly, suppressing FOXM1 in cancer therapy could have destructive outcomes in tissues that are prone to need FOXM1 activity and display high cellular turnover.

1.6.4 FOXO factors

FOXO proteins have diverse and crucial roles: they are involved in cell cycle, apoptosis, metabolism, and lifespan (60-63). Interest in FOXO proteins largely lies in their functions in regulating insulin and insulin-like signalling (64). In brief, insulin binding to the insulin receptor results in the phosphorylation of insulin receptor substrate 1/2 (IRS1/2), leading to a phosphorylation signalling cascade that involves PI3K and AKT, and eventually results in the phosphorylation of FOXO. Without insulin signalling, FOXO proteins reside in the nucleus; however, active insulin signalling causes them to translocate to the cytoplasm, where they are unable to bind or trans-activate their target genes. Moreover, AKT-phosphorylated FOXO proteins are stabilised in the cytoplasm though interactions with protein chaperone 14-3-3. Noteworthy, this stabilisation can be abolished by the phosphorylation of 14-3-3 by Jun kinase and dephosphorylation of FOXO by protein phosphatase 2A. FOXO protein stability is also regulated by ERK1/2, which phosphorylate FOXO proteins and target them for degradation.

Mammalian FOXOs have four variants: FOXO1, FOXO3, FOXO4, and FOXO6. FOXO6 shares the least sequence homology with other FOXO family members (62). FOXO1 is the most detailed of the family however, FOXO3 and FOXO4 have roles in regulating lifespan and metabolism and in regulating cell cycle arrest (60). To date, the mechanisms by which FOXO3 or other FOXO TFs regulate lifespan, if any, in

humans is still unclear. Available evidence shows that FOXO plays a role in several molecular processes, such as stress resistance, cell cycle resistance and apoptosis – which could affect healthy aging.

1.6.5 Overlapping and Co-regulatory Roles for FOX TFs

Albeit the distinct expression patterns of FOX TFs and their separate roles in the regulation of biological processes, multiple TFs of the family frequently have similar roles (50). FOXA and FOXO TFs are involved in glucose homeostasis and metabolism. Nonetheless, FOXM1, FOXP1, FOXP2, and FOXP4 were implicated to affect these processes as well. The FOX family also acts as pioneer factors, the famously known FOXA group, as well as both FOXO and FOXE, to regulate chromatin accessibility.

Particularly, several FOX TFs have been associated with cell proliferation. FOXM1 and FOXO, have long been implicated to have antagonistic roles in cell cycle and tumorigenesis (52, 65). FOXC2, FOXE3, FOXF1, FOXF2, FOXN1, and FOXN3, also regulate proliferation. FOXA2 acts as a tumour suppressor by maintaining a differentiated state and decreasing metastasis in liver, lung and breast tissue, while FOXA1 acts as an oncogene in breast and prostate cancer (53). In the liver, FOXA proteins act as both repressors and activators of hepatic carcinogenesis, based on the sex of the organisms (54, 66). Interestingly, while male mice lacking both FOXA1 and FOXA2 develop lesser tumours, it is the opposite in female mice, showing the key role of FOXA factors in modulating sex steroid hormone receptors.

Distinct FOX TFs also interact with ESR1 to regulate several aspects of tumorigenesis (52). In the human breast cancer MCF-7 cell line, FOXA1 plays as a pioneer TF and loosens up chromatin, permitting the nearby binding of ESR1. At cell cycle regions, FOXM1 can displace FOXA1 and trans-activate cell cycle genes. FOXM1 binds 70% of similar genes as FOXA1 in MCF-7 (67). Particularly, FOXO TFs act as ESR1 status-dependent tumour suppressors in breast cancer. In ESR1-positive breast cancer, FOXO3 is linked to less tumour invasiveness, whereas in ESR1-negative tumours, FOXO3 is linked to more invasiveness (68).

Most of the FOX TFs are essential for life; they evidently are unable to compensate entirely for each other. Nonetheless, their highly related consensus binding sites indicate that they can replace each other to a certain extent (Figure 1.4). It is expected

that the overlapping roles of FOX proteins echo differences in expression patterns, such that one FOX protein might act the same role in the gut but another member act in breast tissues. Another façade of this regulation is different FOX proteins likely act in diverse pathways on the same gene regions, for instance by recruiting distinct co-factors (69) (Figure 1.6; Figure 1.7).

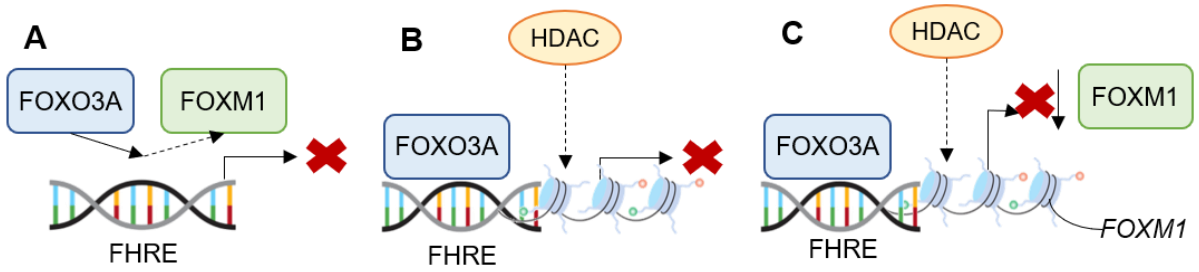


Figure 1.5 Proposed mechanism of the antagonistic relationship between FOXO3A and FOXM1.

Diagram showing the various mechanisms by which FOXO3A antagonises FOXM1 activity. A. Activated FOXO3A displaces FOXM1 from the forkhead response element (FHRE) of target gene promoter region. B. FOXO3A recruits chromatin remodelling proteins, including HDACs, to promote DNA condensation, thus restricting the binding of TFs. C. *FOXM1* itself is a negative transcriptional target of FOXO3A. Thus, activated FOXO3A displaces FOXM1 on the FHRE of the *FOXM1* promoter, leading to sustained suppression of FOXM1 expression. Dashed arrows represent the output effects.

1.6.6 Molecular Profile of FOXM1 in Breast Cancer Subtypes

Considering the role of FOXM1 in carcinogenesis and the distinctive molecular profile of FOXM1 target genes in different breast cancer subtypes, the development of targeted therapeutics for high-risk patients is emphasised. Hyperactivation of FOXM1 was reported in TNBC patients, enabling enhanced expression of subtype-specific cell proliferation gene signatures (6). Lu et al. and Park et al. (70) have also reported higher *FOXM1* mRNA levels in TNBC patients. Overexpression of *FOXM1* was also identified in the luminal B subtype. Lu et al. has shown through clinical datasets that increased expression of FOXM1 in patients with ER-positive breast cancer subtypes, luminal A and B, predicts a shorter distant metastasis-free survival.

In the MDA-MB-231 cell line, a highly metastatic TNBC basal subtype, FOXM1 upregulation was observed, thus reinforcing the association between increased FOXM1 levels and cancer progression in certain breast cancers. Additionally,

immunohistochemistry and real-time quantitative PCR (RT-qPCR) analyses showed an over-expression of *FOXM1* mRNA and protein in breast cancer tissues relative to healthy tissues. Therefore, a better understanding of the role of FOXM1 and the associated signalling pathways could be utilised to improve breast cancer prognosis and contribute to the development of effective therapeutic targets for breast cancer treatment with a lower rate of recurrence and chemotherapy resistance.

1.6.7 The Autoregulatory Loop of FOXM1

The transcriptional inhibitors of FOXM1, siomycin A and thiostrepton, downregulated both the transcriptional activity and expression levels of FOXM1, indicating the presence of an autoregulatory loop for FOXM1 (71). Another study reported that -745/-738bp site is necessary for the autoregulatory activation of *FOXM1* promoter (56). Additionally, nuclear AURKA can be recruited by FOXM1 as a co-factor to *trans*-activate FOXM1 target genes in a kinase-independent manner (72). AURKA and FOXM1 are involved in a tightly coupled positive feedback loop to enhance breast cancer stem cell phenotype. AURKA can effectively *trans*-activate the *FOXM1* promoter though a forkhead response element, whereas FOXM1 can activate *AURKA* expression at the transcriptional level.

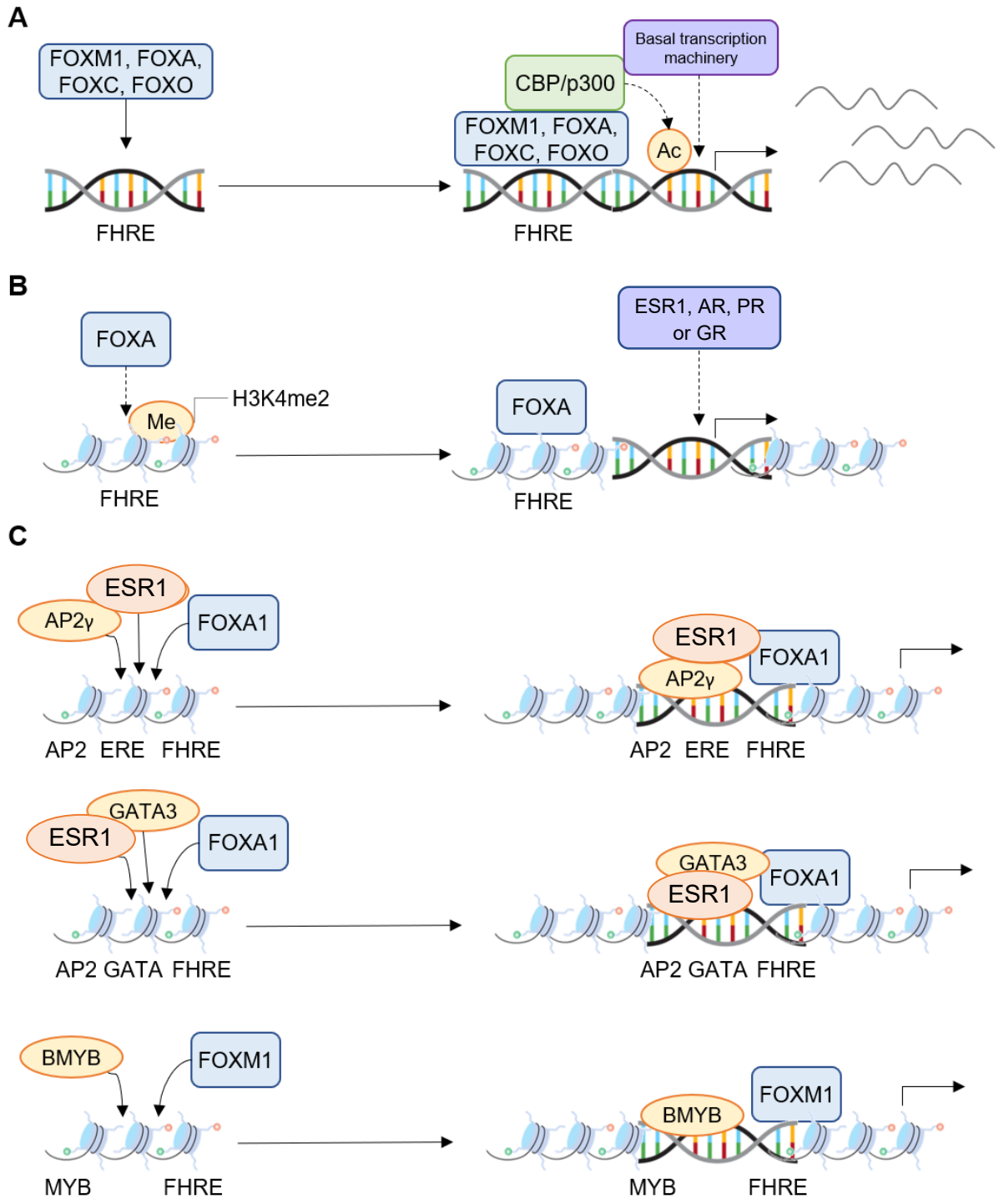


Figure 1.6 Mechanisms of FOX transcriptional factors activation.

Various mechanisms by which FOX proteins activate transcription of target genes. **A**. FOX proteins can function as classic TFs that recruit co-activators such as CBP/p300 that promote the recruitment and assembly of components of the general transcription machinery, such as transcription initiation factors TFIID and TFIIB, and the RNAPII. **B**. FOX proteins can also act as pioneer factors that open condensed chromatin and facilitate binding of other TFs. **C**. They may also be cofactors that actively recruit other TFs into complex. Notably, CBP/p300 interacts simultaneously with several TFs and thereby promotes transcriptional synergy. It also contains a histone acetyltransferases (HAT) domain that is important for

chromatin acetylation (Ac) and decondensation. Representative examples of FOX proteins and their cooperating partners are shown: activating enhancer-binding protein 2γ (AP2γ), oestrogen receptor-α (ESR1) and FOXA1 interact directly to mediate oestrogen-mediated transcription; ESR1, FOXA1 and GATA3 form a functional enhanceosome to drive the transcription of the ERα-regulated gene in breast cancer cells; and BMYB and FOXM1 cooperate to promote mitotic gene expression. Dashed arrows represent the resulting effects. AR, androgen receptor; ERE, oestrogen response element; FHRE, forkhead response element; GR, glucocorticoid receptor; Me, methylation; PR, progesterone receptor.

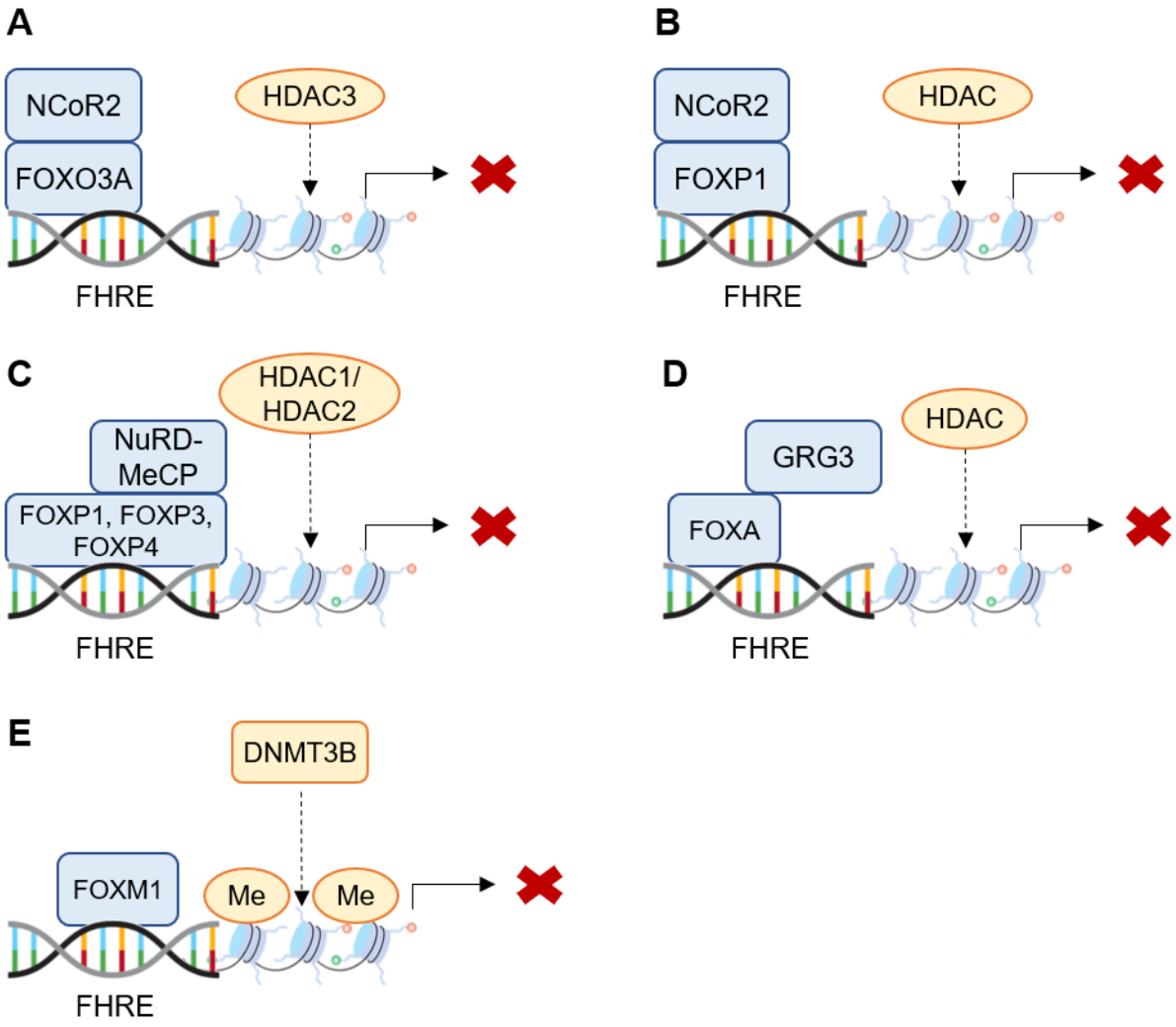


Figure 1.7 FOX transcriptional factor regulated repression.

Diagram showing various mechanisms suppressing target genes by FOX proteins. FOX proteins, their cofactors and chromatin or DNA modifiers regulated target gene transcriptional suppression are illustrated. **A**. FOXO3A forms a complex with nuclear receptor co-repressor 2 (NCoR2) and HDAC3 to compact chromatin and restrict gene activation. **B**. FOXP1 also interacts with the NCoR complex, which recruits HDACs to suppress target genes. **C**. FOXP1, FOXP3 and FOXP4 bind the NuRD-MeCP1 chromatin remodelling complex and recruit HDAC1 and HDAC2 to suppress gene expression. **D**. FOXA1 and FOXA2 also recruit GRG3, resulting in binding of HDACs to suppress gene expression. **E**.

FOXM1 can suppress gene expression, such as GATA3, through recruitment of DNA methyltransferase DNMT3B.

1.7 Endoplasmic Reticulum

The endoplasmic reticulum (ER) is a eukaryotic intracellular organelle that functions as a powerhouse for the production and folding of cellular proteins, thus conferring it unwavering responsibility in the maintenance of cellular homeostasis and the delicate balance between cell survival and death (73). ER is associated with synthesis, maturation, folding, quality control, post-translation modifications and degradation of protein secretome and trans-membrane proteins, these unfolded proteins enter the ER and exit as either correctly folded proteins in transit to their target organelles or as misfolded proteins destined for degradation (74).

1.7.1 ER Stress

In cancer tumours, the fine balance to maintain protein folding homeostasis in the ER gets threatened when cells face adverse conditions. Intrinsic stress factors include the increased rate of oncoprotein synthesis, while extrinsic factors cover toxins, ischemia, hypoxia, acidosis, viral infection, pH changes, lack of energy supply and nutrients (63, 74). This over demand for protein folding can lead to disrupted protein folding or accumulation of unfolded protein folding in the ER lumen. When ER protein folding capacity is swamped with cellular demand and/or cellular energy availability is insufficient to properly fold proteins produced in ER, this homeostatic imbalance can cause misfolded proteins to accumulate within the ER lumen. This phenomenon is known as ER stress and causes a transient activation of a cellular stress response named as the unfolded protein response (UPR). While the UPR primarily works to decrease excess unfolded proteins and restore ER function, severe or sustained UPR signals can activate cell death. The pro-survival arm of UPR tries to counter ER stress through translational attenuation, upregulation of ER chaperones and related proteins, and degradation of misfolded or unfolded proteins by quality control. However, when the ER is too overwhelmed by protein folding demands, the organelle releases

apoptotic signals. Expectably, many tumours display consistent UPR activation which allows them to acclimatise and thrive under extreme conditions (20).

1.7.2 UPR

UPR signalling originates from three protein sensors that span across the ER membrane – Inositol Requiring Enzyme 1 (IRE1, or ERN1), Activating Transcription Factor 6 (ATF6), and PKR-like ER Kinase (PERK, or EIF2AK3) (Figure 1.8). Under normal conditions, all three sensors are bound by a protein chaperone known as glucose-regulated protein 78 kDa (GRP78; also known as BIP or HSPA5) which renders them in an inactive state. However, upon stressful conditions, ER stress occurs and GRP78 dissociates from the protein sensors, leading to their activation. The UPR supports cell survival by promoting the expansion of the ER, increasing the abundance of protein chaperones to facilitate protein folding and driving protein degradation systems, including autophagy and ER-associated degradation (ERAD). Nonetheless, if these mechanisms fail to restore homeostasis in ER, the UPR activates protein chaperones/heat shock proteins, such as DNA damage-inducible transcript 3 (DDIT3), JNK, and ER-associated caspase-12, and this triggers cell death. Misfolded and abnormal proteins are degraded through a systemic pathway: aberrant proteins are labelled by protein chaperones, poly-ubiquitinated, and degraded in the 26S subunit of the proteasome, directing the cells towards adaptive UPR. Afterward, if the cell can overcome the ER stress condition, pro-survival events are activated, however, if stress condition is prolonged and persistent, UPR triggers pro-death signalling pathways. Temporal regulation of IRE1, PERK, and ATF6 manage the switch between pro-survival and pro-death UPR signalling. This process is manipulated by signals that crosstalk between plasma membrane, ER, cytosol, mitochondria, and nucleus, switching between cellular metabolism essential to enable activation or suppression of apoptosis and/or autophagy. It is essential to note that human cancers, such as hepatocellular, lung, pancreatic, and breast cancer, are characterised by chronic UPR and this leads to tumour growth and therapy resistance.

1.7.3 Role of Protein Sensors

As mentioned above, UPR has two arms, one leading to survival and another leading to apoptosis. Initially, cells prioritise the pro-survival arm to overcome adverse conditions and maintain ER homeostasis. To do that, this organelle goes through three ER stress sensors: IRE1, ATF6, and PERK. Firstly, UPR reacts with translational attenuation. Activated PERK phosphorylates its target protein eukaryotic initiation factor 2 (eIF2) during ER stress and inhibits protein synthesis, which stops further influx of ER client proteins. Secondly, UPR triggers the upregulation of the expression of molecular chaperones by ATF6, such as GRP78 and GRP94 to increase folding capacity in the ER. Additionally, components of ERAD, such as IRE1, are transcriptionally activated to remove misfolded proteins from the ER by the ubiquitin-proteasome system. Thirdly, the association with Ca^{2+}-NFκB pathway which is induced by accumulation of membrane proteins in the ER. For the pro-apoptosis arm, cells react to a severely impaired ER with overloaded protein folding demand by activating relevant caspases. The essential downstream proteins include DDIT3, JNK, and ER-associated caspase-12.

1.7.3.1 DDIT3 in UPR-mediated Apoptosis

One of the critical downstream proteins of the UPR-mediated apoptosis is DNA damage-inducible transcript 3 (DDIT3). It is also known as C/EBP homologous protein (CHOP) or growth arrest and DNA damage-inducible gene 153 (GADD153). DDIT3 is one of the six identified members of C/EBP trans-acting family of TFs that bind to the CCAAT box motif in several gene promoters, which is implicated in the regulation of biological processes in cellular proliferation, differentiation, and expression, and energy metabolism of cell type-specific genes (75). This 27 kDa DDIT3 protein has been receiving an increasing amount of attention due to its role of in apoptosis during ER stress and in tumours development (75, 76).

Normally, the expression of *DDIT3* is extremely low. However, when ER stress gets out of control, apoptosis can be initiated (Figure 1.8). The expression of *DDIT3* and its accumulation in the nucleus are upregulated when the pro-apoptosis arm of ER stress-induced signalling pathway is activated. *DDIT3* was also induced by nutrient starvation such as glucose deprivation and amino acid depletion. Overexpression of *DDIT3* has

been reported to lead to cell cycle arrest and/or apoptosis (77). Cells lacking DDIT3's major dimer C/EBβ are resistant to ER stress-induced apoptosis, suggesting that DDIT3 works as a transcriptional factor that regulates genes involved in either pro- or anti- apoptotic genes, including *BCL2* (B-cell lymphoma-2), *TRIB3* (tribbles-related protein 3), *GADD34* (78-80), and other undescribed genes referred to as *DOC*s (downstream targets of DDIT3) (81).

The other widely accepted mechanism of DDIT3-induced apoptosis is the suppression of the pro-survival protein BCL2. Under ER stress, DDIT3 downregulates the expression of *BCL2* and sensitises cells to apoptosis. *Bim* and subsequently, *Bax*, is also upregulated during ER stress in cardiomyocyte models, indicating that *Ddit3*-deficient mice have less apoptotic cell death and lower caspase-3 activation related to a decrease of Bax levels (82). Bcl-2-associated anthanogene 5 (*BAG5*) is overexpressed in prostate cancer and inhibits ER stress-induced apoptosis (83). BAG5 overexpression led to a decrease in *DDIT3* and *BAX* and an increase in anti-apoptotic *BCL-2* gene expression. DDIT3 has been reported to regulate the expression of BH3-only proteins by interacting with FOXO3 in neuronal cells treated with tunicamycin and AP-1 complex protein c-Jun leading to its phosphorylation in saturated fatty acid-treated hepatocytes (80). The regulation of mitochondrial permeabilization through BAX and BAK is a major pro-apoptotic action via the DDIT3-ERO1α-IP3R-calcium-calcium/calmodulin-dependent protein kinase II (CaMKII) pathway (84). *GADD34*, another pro-apoptotic downstream target of sustained DDIT3 expression, promotes the dephosphorylation of phosphorylated eIF2 and hence restores protein translation. In a study utilising cells inserted with tamoxifen-inducible *DDIT3* expression constructs, distinct gene expression profiles regulated in cytoplasm and nucleus were illustrated (85). More than two-thirds of the 175 target genes were downregulated, suggesting that DDIT3 mainly functions as a dominant-negative regulator in either cytoplasmic or nuclear sequestration of dimer forming TF partners.

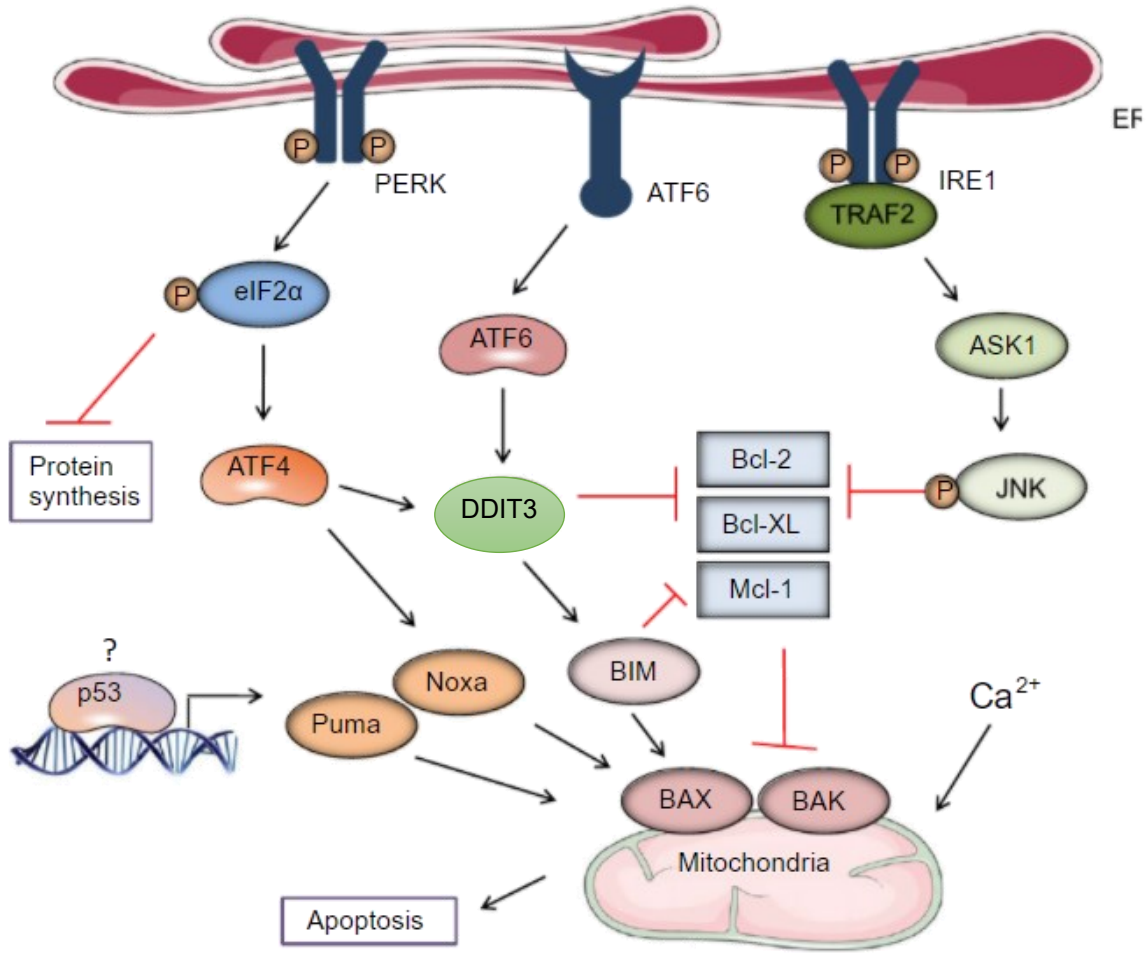


Figure 1.8 ER stress induced apoptosis.

ER stress activates PERK, which in turn phosphorylates and suppresses eIF2a. This leads to the induction of ATF4 and its downstream target proteins DDIT3 and NOXA, resulting in apoptosis. Secondly, DDIT3 can also be activated by ATF6, leading to activation of BIM and suppresses BCL-2. The involvement of p53 activation upon ER stress and resulting activation of NOXA and PUMA is as yet unclear. Thirdly, IRE1a can recruit TRAF2, which activates ASK1 and its downstream target JNK. JNK can activate apoptosis by suppressing anti-apoptotic proteins such as BCL-2 and BCL-XL. Figure adapted from Lurlaro and Munoz-Pinedo 2015 (86).

Hypoxia is a physiologically important trait that triggers cell death in the context of the ER, especially in solid tumours, where DDIT3 has been implicated in the activation of tumour cell death. *DDIT3* expression is correlated to prolonged disease-free survival (DFS) in patients administered with DNA-damaging chemotherapy drugs, the anthracycline class including doxorubicin (87). Upregulated levels of *DDIT3* are reported in decreasing ER stress-induced resistance to doxorubicin in human hepatocellular carcinoma cells. Tumour cells with high *DDIT3* expression may be more

susceptible to anthracycline cytotoxicity and predictive of an improved DFS. Patients with GRP78-positive and DDIT3-negative breast tumours had comparatively worse clinical outcomes than other patient populations when treated with anthracycline, suggesting that *DDIT3* may be useful to predict responsiveness to chemotherapy. DDIT3-mediated *DR5* expression is involved in caspase 8-mediated apoptotic pathway. Asparagine was showed as a DDIT3 inhibitor with anti-apoptotic function, inhibition of asparagine synthetase may trigger DDIT3-activated apoptosis and help with the treatment of solid tumours (88). Moreover, DDIT3 forms a complex with C/EBPβ and decreases C/EBPβ-dependent *ALDH1A3* expression in chemoresistant cells in the butein-induced enhancement of chemoresistant cell apoptosis and benefits the treatment of non-small cell lung cancer with garcinol (89). Furthermore, DDIT3 downregulated anti-apoptotic *p21* in cancer cells treated with tunicamycin and consequently enhanced chemotherapeutic drug efficacy (90) . During radiotherapy with high-LET carbon ions, chloroquine co-treatment promotes apoptosis through the IRE1-DDIT3 signalling pathway in-vitro and in-vivo (91).

1.8 DNA Methyltransferases (DNMTs) and Methylation of CpG islands

Methylation is now considered as the main player involved in gene expression regulation. Unlike transcription profiles that represent snapshots of expressions of molecules and transcription activities at a specific time, DNA methylation signature is more permanent therefore represent longer-lasting effects thus carries higher predictive values of molecular state and cancer predisposition of a cell (92).

In mammalian cells, DNA methylation is a reversible process that happens most commonly with an addition of a methyl group in the fifth position of the pyrimidine ring of cytosine on CpG sites within the genome, known as CpG islands, in roughly 60% of genes. Methylation of promoter regions upstream of CpG islands, known as CpG island shores, as well as intragenic sequences, are essential for the regulation of gene expression and can be involved in disease initiation and progression. Generally, in normal cells, gene promoters are unmethylated, while gene bodies and intergenic regions are methylated.

Aberrant methylation in these CpG islands may lead to the silencing of genes associated with key proliferation and apoptosis pathways such as tumour suppressor

genes (TSGs), DNA repair and hormone receptor genes, as well as genes that suppress angiogenesis. Human DNA methylation is introduced into the sequence of nucleotides by DNA cytosine methyltransferases (DNMT). The DNMT family includes 5 enzymes – DNMT1, DNMT2, DNMT3A, DNMT3B, and DNMTL – however, only DNMT1, DNMT3A, and DNMT3B are associated with active methyltransferase activity (93) (Figure 1.9). DNMT1 methylates newly synthesised strands prior to chromatin packaging and is localised to the replication foci during the S-phase. DNMT1 is typically highly expressed for maintenance of the methylation pattern during replication. Overexpression of DNMT1 is observed in many types of cancers and this increase in expression is associated with cellular transformation, while decreased DNMT1 expression levels seem to be related to a protective effect. DNMT3A and DNMT3B proteins allow *de novo* DNA methylation activity in vitro, without differentiation between unmethylated and hemimethylated DNA. Studies reported an interaction between DNMT1 and DNMT3B may be crucial for DNA methylation pattern in the transformation of human colon carcinoma (93). Collectively, the levels of DNMTs, particularly DNMT3A and DNMT3B, are typically upregulated in many types of cancer, often contributed to the hypermethylation of CpG islands in promoter regions of tumour suppressor genes in tumours.

1.8.1 DNA methylation and breast cancer

Many studies have validated that abnormal DNA methylation footprint is a hallmark of cancer and biomarker in various cancers, including breast cancer (94). The expression of oncogenes and tumour suppressor genes in the process of breast tumorigenesis are affected by hypo- and hypermethylation, respectively (9, 95). Additionally, the expression of DNMTs represents a prognostic marker in breast cancer cells (96, 97). Clark et al. (98) have demonstrated (a) ESR1-negative breast cancer is related to synchronised hypermethylation, and they classified TNBC into 3 unique methylation clusters associated with prognostic outcomes, with 17 differentially methylated regions identified as significantly correlated with survival outcomes; (b) an organised hypermethylation of key gene loci is associated with compromised binding of important transcriptional regulators and the initiation of breast cancer, and (c) the methylation state of enhancers is involved in ESR1-positive breast cancer and therapeutic sensitivity. Furthermore, abnormal DNA methylation correlates with clinical and

pathological features of breast cancer, such as *TP53* status, tumour and histological stage, gene expression changes, and prognosis (92). Additionally, hypermethylation of *BRCA1* is correlated to ESR1-negative breast cancer and poor clinical outcomes. DNA hypermethylation was linked to tumour hypoxia, a study showed that tumour hypoxia decreases the demethylation activity of ten-eleven translocation enzyme (TET), consequently inducing methylation at the promoters of tumour suppressors.

Particularly, the expression of DNMT3B is upregulated in human cancer types, while its suppression leads to tumour cell apoptosis (93). Inactivation of the TSG Ras association domain family 1 isoform A (*RASSF1A*), through promoter hypermethylation induced by DNMT3B, is a common event in several cancers. Additionally, DNMT3B plays a predominant role over DNMT3A and DNMT1 in breast tumorigenesis, given its overexpression in breast cancer tissues compared to DNMT1 and DNMT3A (99). Overexpression of DNMT3B protein significantly contributes to the upregulated DNMT activity observed in breast cancer cell lines (99). Though DNMT3L has no methyltransferase catalytic activity on its own, this enzyme supports the activity of DNMT3A and DNMT3B by promoting their binding capacity for the methyl group donor, S-adenosyl-L-methionine.

1.8.2 Breast Cancer Treatment Through Regulation of Methylation

DNA methylation pattern modifications are one of the earliest alterations happening in the process of cancer development, hence they can be helpful for early detection (94). Early stage detection (stages I and II) and breast cancer treatment improve the 5-year survival rate to more than 93% as opposed to a late-stage (stage IV) diagnosis for patients with metastatic cancer with barely 22% survival rate (8). Several TSGs, including DNA repair, apoptosis, hormone receptor, cell cycle, and TF genes, have been recognised as being differentially methylated in breast cancer and may be hopeful therapeutic targets. Furthermore, breast cancer treatment and biomarkers through modulation of molecules involved in methylation processes are also favourable (6).

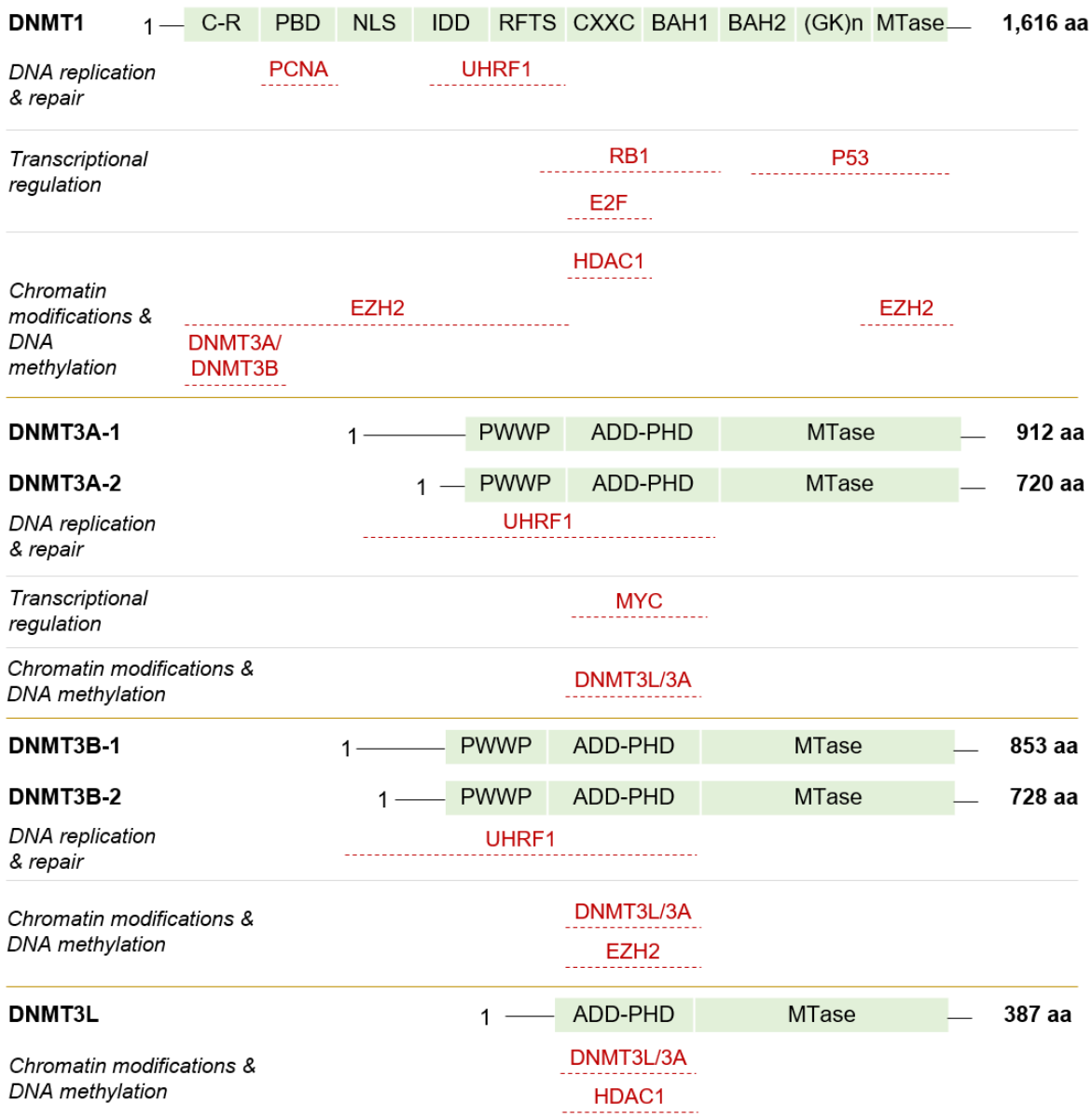


Figure 1.9 Structural illustration and transcriptional binding of mammalian DNA methyltransferases (DNMTs).

Human DNMT1 harbours a charge-rich (C-R) domain containing the proliferating cell nuclear antigen binding domain (PBD), an intrinsically disordered domain (IDD) with a nuclear localisation sequence (NLS), a replication foci target sequence (RFTS), a zinc finger domain (CXXC), and two bromo-adjacent homology domains (BAH 1 / 2). The catalytic and regulatory domains are linked by a series of Glycine-Lysine repeats (GK)n. DNMT3A-1 and DNMT3A-2 are transcriptional isoforms due to alternative promoters, likewise for DNMT3B-1 and DNMT3B-2. The catalytic methyltransferase domain (MTase) harbours highly conserved motifs. PWWP domain and an ATRX-DNMT3B-DNMT3L and Plant-homeodomain (ADD-PHD) are also present in DNMT3B, DNMT3L, and DNMT3C. Interacting RNAs are depicted in red. Functional relationships were referenced from (100).

Vijayaraghavalu et al. noted a significant increase in doxorubicin treatment efficacy in MCF-7, MDA-MB-231 and BT-459 cells when combined with a demethylating agent, decitabine (5-aza-2'-deoxycytidine) (101). This combinatorial treatment induced cell cycle arrest in more than 90% of cells and restored sensitivity to doxorubicin through upregulation of p21 and silencing of DNMT1, leading to the expression of tumour suppressor genes and depletion of oncogenes to overcome drug resistance of these breast cancer cells. A treatment combining methyltransferase inhibitor (5-azacytidine) administration and suppression of *MBD2* expression using short interfering RNA (siRNA) led to re-activation of invasion and metastatic processes in breast cancer cells (102). The same strategy including DNMT inhibition and *LSD2* knockdown was also effective for cell growth suppression in breast cancer MDA-MB-231 and MCF-7 cells. This treatment induced expression of epigenetically silenced genes including those encoding PR and ESR1.

1.9 Breast Cancer Cell Lines

MCF-7, a breast cancer cell line isolated in 1970, is a well-characterised human ESR1-positive breast cancer cell line model that is estrogen-responsive *in vitro*, estrogen-dependent *in vivo*, anti-estrogens sensitive, and usually easily manipulated with plasmid vectors transiently overexpressing a number of ligand growth factors and respective growth factor receptors to study the effects of constitutive overexpression of these factors on breast cancer phenotypes (103). *In vitro*, MCF-7 cells are typically used to investigate the effects and mechanisms of hormonal therapies or endocrine mechanisms because they contain large masses of ESR1-positive cells and are therefore easily manipulated with anti-estrogens or estrogen/oestradiol (2, 104). MCF-7 (103), T47D (105), and MDA-MB-231 (39) contribute to more than two-thirds of *in vitro* studies for breast cancer discoveries. Both MCF-7 and T47D cell lines belong to ESR1-positive, luminal A subtype but T47D may be a better model to study progesterone-specific effects, whereas MDA-MB-231 cell line is a cell line representative for TNBC.

1.10 Overall Aims and Thesis Outline

Forkhead transcription factor superfamily houses various proteins that are involved in diverse essential biological processes by refining the spatial and temporal expression of extensive number of target genes in developing cells and in adult tissues. This ability to regulate essential signalling pathways has rendered them attractive targets for tumorigenesis. Indeed, Forkhead transcription factors are strongly implicated in cancer processes, from initiation, maintenance, progression to drug resistance.

In normal cells, the Forkhead transcription factor FOXM1 is an essential regulator of cell cycle. In pan-cancer cells, overexpression of FOXM1 is commonly correlated to uncontrolled proliferation, migration, metastasis, angiogenesis, drug resistance, and poor prognostic outcome. Naturally, FOXM1 was presumed to be an attractive therapeutic target for cancer treatment. However, as is the case for other cell cycle proteins, the impacts of activating or inhibiting FOXM1 expression or activity must be carefully evaluated. Drug treatments are currently not targeted to specific tissues, activating FOXM1 endogenously to replenish replicative potential in a tissue of interest without aiming it to a specific cell type could potentially cause existing tumours to become prone to growth or even metastasis. Similarly, suppressing FOXM1 in cancer therapy could have destructive outcomes in cells that require FOXM1 activity to drive fundamental biological processes such as mitotic spindle assembly or correct chromosome segregation, especially cells with high cellular turnovers. Therefore, the activation or inhibition of FOXM1 directly in treatments are often not a viable option.

In **Introduction Section 1.6.5 Overlapping and Co-regulatory Roles for FOX TFs**, I have stated the overlapping but distinct roles of FOX TFs. Essentially, most of the FOX TFs are essential for life, they evidently are unable to compensate entirely for each other. Nonetheless, their highly related consensus binding sites indicate that they can replace each other to a certain extent. It is expected that the overlapping roles of FOX proteins echo differences in expression patterns, such that one FOX protein might act the same role in the gut but another member act in breast tissues. Another façade of this regulation is different FOX proteins likely act in diverse pathways on the same gene regions, for instance by recruiting distinct co-factors for either activation or suppression of gene expressions. Additionally, FOXM1, as an oncogenic TF, exerts its activity through regulating the expression of target genes. These FOXM1 signalling

networks open up new routes for its suppression as an alternative to direct suppression of FOXM1.

My study focuses primarily on clarifying the relationship between FOXM1 and its putative co-factors, and understanding the genetic and epigenetic relevance of these interactions, if present, in essential signalling pathways involved in endocrine resistance, chemotherapy, and UPR. I focused on the identification of these potential additions to FOXM1 cistrome in a breast cancer cell line-oriented manner.

In **Chapter 3**, I explored the functional differences between the FOXM1/ESR1 common regulatory regions and FOXM1 specific (ESR1 negative) target regions. In **Chapter 4**, I studied the relationship between FOXM1 and ER stress downstream effector, DDIT3 in breast cancer cell lines. Lastly, in **Chapter 5**, I investigated the relationship between FOXM1 and different DNMTs, namely DNMT1, 3A, and 3B, in breast cancer models.

FOXM1 cistrome and cancer cell line-specific co-factors regulating FOXM1 cistrome have indicated better prognostic outcome over FOXM1 expression itself in breast cancer. Collectively, there is a need to identify the targets, mode of actions, resistance mechanisms regulated by FOXM1 to facilitate a better understanding of breast cancer development, progression, treatment, and therapy resistance.

2 Materials and Methods

2.1 Consumables

All pipette tips were from Starlab (Milton Keynes, UK) and VWR (Radnor, PA, USA). Cell scrapers, centrifuge tubes, tissue culture dishes, tissue culture flasks, and sterile serological pipettes were from VWR (Radnor, PA, USA). Aspirator Pipettes were from Sarstedt (Nümbrecht, Germany). Cloning cylinders were from Sigma-Aldrich (St. Louis, MO, USA). Microcentrifuge tubes and all remaining plastic wares were from Starlab (Milton Keynes, UK). The nitrocellulose membrane 0.2 μM was purchased from Thermo Fisher Scientific (Waltham, MA, USA).

2.2 Cell Culture

2.2.1 Cell Lines

The human breast carcinoma cell lines MCF-7 and MDA-MB-231 originated from the American Type Culture Collection (Manassas, VA, USA) and were purchased from the Cell Culture Service from Cancer Research UK, where they were authenticated. The cell lines were cultured in Dulbecco's modified eagle medium (DMEM) (Sigma Aldrich; Cat. No.: D5546) with 10 % foetal calf serum (FCS) (First Link; Cat. No.: 02 – 00 – 850) 100 U/ml penicillin/streptomycin and 2mM glutamine (Sigma Aldrich; Cat. No.: G6784), and maintained in a humidified incubator with 10 % carbon dioxide (CO_2) at 37 °C.

The epirubicin-resistant MCF7-Epi^R cell line was derived from the parental MCF-7 cell line. It was generated previously in the laboratory and were cultured as detailed above and supplemented with 17 μM of epirubicin.

MCF7-Y537S cell line was obtained from Dr Alison Harrods (104) (Imperial Centre for Translational and Experimental Medicine, London, UK).

MCF7-FulR is a fulvestrant resistant breast cancer cell line. The parental MCF-7 cell line was treated long-term with 100 nM fulvestrant and the surviving clone was expanded into a resistant cell line. To maintain a high level of resistance, the cells are

cultured in DMEM/F12 medium (no phenol red) supplemented with 1% fetal bovine serum, 2.5 mM GlutaMAX, 6 ng/mL insulin, and 100 nM fulvestrant.

TABLE 2.1 Cell cultures used in this research.

Cell line	Type	Medium
MCF-7	Human breast carcinoma	DMEM
MCF7-EpiR	Human breast carcinoma, epirubicin resistant	DMEM, supplemented with epirubicin
MCF7-Y537S	Human breast carcinoma with CRISPR engineered Y537S mutation	DMEM
MCF7-FulR	Human breast carcinoma, fulvestrant resistant	DMEM supplemented with fulvestrant
MDA-MB-231	Human breast carcinoma	DMEM

2.2.2 Cell Storage

Cells were detached from monolayer surface with trypsin, washed in complete DMEM media and spun down at 1,200rpm for 5 min. The cell pellet was then resuspended in FCS with 10% Dimethyl sulfoxide (DMSO). 1ml aliquots were transferred to cryotubes and slowly frozen in a Mr. Frosty Nalgene® freezing container (VWR International) at -80°C overnight before transferring to long term storage in vapour-phase liquid nitrogen tanks. When thawing cells, the cryotubes were defrosted rapidly in a waterbath (~37°C) washed in complete DMEM, then spun down at 1,200rpm for 3 min to remove any DMSO residue media and the pellet was resuspended in 10ml of fresh supplemented DMEM in a 25cm^2 culture flask.

2.2.3 Chemicals

Epirubicin (2mg/ml in 0.9% sodium chloride) was obtained from Imperial College Healthcare, UK and stored in solution at 4°C. **17ß-estradiol (E2)** and **Faslodex™ (fulvestrant)** (Sigma-Aldrich, Dorset, UK) were prepared in ethanol.

2.3 Protein Analysis

2.3.1 Preparation of Total Protein Lysates

Whole cell protein extracts were prepared by harvesting cells and washing the cell pellets in PBS buffer before centrifuging at 2,000rpm for 5 min. Supernatant was discarded and pellets were stored at -80°C until lysis was conducted. Frozen pellets were lysed in 2 volumes of NP40 lysis buffer [(1% (v/v) Nonidet P-40, 150mM $NaCl_2$, 50mM TRIS-HCl (pH 7.6), 5mM EDTA, 1mM DTT, 1mM NaF, 2mM PMSF, 1mM sodium orthovanadate and cOmplete and Mini EDTA-free protease inhibitor cocktail, as described by the manufacturer's protocol (Roche Diagnostics, Burgess Hill, West Sussex, UK)] at 4°C for 20 min. Insoluble lysate material was removed by centrifugation at 13,000rpm for 10 min at 4°C. Supernatant was transferred to a new and clean Eppendorf tube where the protein concentration was measured. Protein concentration was measured using the Bio-Rad D_c protein assay (Bio-Rad Laboratories, CA, USA), as described by the manufacturer's protocol. Absorbance was read at 750nm for each of the experimental samples as well as the non-protein negative control. Protein concentrations were measured by the equation absorbance x 25 = µg/µl.

2.3.2 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blotting

Protein samples were analysed by SDS-PAGE using whole-cell extracts. The samples were heat-denatured at 95°C for 5 min and briefly spun at 13,500g. With an estimation, 25 – 50µg of protein for each sample were loaded and resolved by 6 – 15% SDS-PAGE at 110V for 1.5 hrs in TRIS-glycine running buffer. Proteins were transferred from the gels onto 0.2µM nitrocellulose membranes using wet-transfer (Tris-glycine running buffer with 20% (v/v) methanol) at 90V for 1.5 hrs. The membranes were blocked in albumin for 1 h at room temperature (RT) and immune-blotted with the respective primary antibodies at 4°C overnight. Membranes were washed 3X with PBST (0.05% Tween-20) for 10 min, before applying HP-conjugated secondary antibodies. Membranes were washed 3X again and immunoreactivity tested using Pierce™ ECL (enhanced chemiluminescence) western blotting substrate for 1 min. Targeted proteins were visualised under quantitative Fusion Solo Chemiluminescence

Imager and image analysis was completed using ImageJ software (National Institute of Health, Bethesda, Maryland, U.S.; Laboratory for Optical and Computational Instrumentation, University of Wisconsin, Madison, Wisconsin, USA). Antibodies against Tubulin or Actin were used as a loading control. The recipe for buffers and gels used in SDS-PAGE is stated in Appendix.

2.3.2.1 Antibodies

Selected antibodies were used to probe for target proteins on Western Blot as following:

TABLE 2.2 List of antibodies used for western blots.

Antibody Name	Company
Beta-tubulin	Cell Signalling Technology, #2146S
DDIT3 / CHOP	Cell Signalling Technology, #2895S
DNMT1	Abcam, ab13537
DNMT3A/B	Abcam, Ab13604
FOXA1	Abcam, ab23738
FOXM1	Santa Cruz Biotechnology, sc-502
FOXO3	Abcam, ab23683
Anti-mouse or anti-rabbit secondary horseradish peroxidase-conjugated secondary antibodies	Dako, Glostrup, Denmark

2.4 RNA Analysis

2.4.1 RNA Extraction and Reverse Transcription

Total RNA was extracted using a kit according to the manufacturer's protocol. In short, cells were lysed with lysis buffer containing 1% β-mercaptoethanol and homogenised. Nucleic acids binding to the membrane of a spinning cartridge provided was assisted by ethanol. Proteins, salts and remaining ethanol were eliminated during the washing steps and the RNA was eluted in RNase-free ddH_2O at RT. The concentration of extracted RNA was measured using a NanoDrop 1000 spectrophotometer and diluted

to a final concentration at 50ng/μl. The High-Capacity cDNA Reverse Transcription kit with random primers was utilised to convert the RNA into cDNA according to the manufacturer's protocols (Thermo Fisher Scientific).

2.4.2 RT-qPCR Primer Design

Primers were designed based on human genes, non-pseudogenes, non-RNA transcripts using http://www.ensembl.org/index.html. The optimum protein-coding sequence was selected and primers were designed based on the selected exon sequences via Primer3plus (http://primer3plus.com/). Primers generated were within 20 – 30 base pairs long, with melting temperatures spanning 55 °C – 60 °C and similar GC% content of less than 60%. Primer pairs for chromatin immunoprecipitation (ChIP) were designed to span the 5' UTR promoter region of target genes to measure gene expression level.

TABLE 2.3 Oligonucleotides qPCR primer pair sequences.

Target Name	Oligonucleotide sequence (5' – 3')	Oligonucleotide sequence (3' – 5')
Actin	AGCGCGGCTACAGCTTCA	CGTAGCACAGCTTCTCCTTAATGT
ADCY1	CGACAATGTGAGCATCCTGTTTG	GCTAATTCATCGAACTTGCCGAA
BIM	GGCCCCTACCTCCCTACA	GGGGTTTGTGTTGATTTGTCA
BRCA1	GCGGACTCCCAGCACAGA	GCTTATTCCATTCTTTTCTCTCACACA
CCNB1	CGCGATCGCCCTGGAAACGCA	CCCAGCAGAAACCAACAGCCGT
CCND1	TGCCACACACCAGTGACTTT	ACAGCCAGAAGCTCCAAAAA
CDC25	AAGAGCCCATCAGTTCCGCTTG	CCCATTTTACAGACCTGGACGC
CDK1	AGGGGTTCCTAGTACTGCAATTC	TGAGATATAACCTGGAATCCTGCA
CDKN3	ATCATCATCCAATCGCAGATGGA	CCAAGTCCTCCATAGCAGTGTAT
CNTRL	TTCTCCATCCCCTTCATCCCTAT	ACCTTTCTTCCCGCTTCTTTGAT
DDIT3	GGTATGAGGACCTGCAAGA	AAGCAGGAAATCGAGCG
DNMT1	AATTCGCGCACCTACTCCAA	CACCTCGATGTTGGGCAGAT
DNMT3A	TGAAGCCTCAAGAGCAGTGG	TCCTGCAATGACCTTGGCTT

DNMT3B	GGAGGCAGACAGTGGAGATG	GTCGGGACCTCTGAGTAACC
ESR1	TGATCAGGTCCACCTTCTAGAATG	CGCCAGACGAGACCAATCAT
ESR1	CCAACCAGTGCACCATTGATAAA	TTCAACATTCTCCCTCCTCTTCG
FOXA1	GCTGGACTTCAAGGCATACGA	GGCAACGTAGAGCCGTAAGG
FOXM1	TGCAGCTAGGGATGTGAATCTTC	GGAGCCCAGTCCATCAGAACT
FOXO3	GCGTGCCCTACTTCAAGGAT	GCTCTTGCCAGTTCCCTCAT
FSIP1	CTTAGGAACACCAAAGAGCAACG	TCAGAGAGACATGATGTGGACTC
FZR1	GGCTCAAGTGGTCCACAGAC	AGGTGCTCCGTGTACTGCTG
GATA3	CCGAGGGCTGGTTTCCTTGA	GACTCTGCAATTCTGCGAGC
GMDS	TACCAAGCCTCAACAAGTGAACT	AGTTCACCACAATCCAATAGGCA
GREB1	GAAGGGCAGAGCTGATAAC	GACCCAGTTGCCACACTTTT
GREB1	GAGGAAGAGGAAGAAGAGGGAGA	TGGCAAAACCCGTCTGTGGTA
GRP94	ACTGTTGAGGAGCCCATGGAGG	CAAAGGGCGCTCTGAGAAGTCG
GTSE1	AGTTTTCTATTGGTGACTCCCCG	CAGATGAGCGTCTAACCGGG
HELLS	TTTGACCCAGTCCGGAAGTG	TGAAGATCCGACTGGGGGTT
HPRT	TATGGCGACCCGCAGCCCT	CATCTCGAGCAAGACGTTCAG
IGSF1	CCTGAGGATGAAGGGGTTTACAT	TGTCAAAGTTGGTTTGGGGTAGA
L19	GCGGAAGGGTACAGCCAAT	GCAGCCGGCGCAAA
MYC	GATTCTCTGCTCTCCTCGACG	CTTCTTGTTCCTCCTCAGAGTCG
p16	AGCCTTCGGCTGACTGGCTGG	CTGCCCATCATCATGACCTGGA
p53	GAGCTGAATGAGGCCTTGGA	CTGAGTCAGGCC CTTCTGTCT T
PLK1	CCAGAGGGAGAAGATGTCCA	GTCGTTGTCCTCGAAAAAGC
PLOD2	ATGTCATATTTGCTGGTGGTCCA	TGCTAGTCTTTTATCTGGCCACA
PUMA	GACGACCTCAACGCACAGTA	CTAATTGGGCTCCATCTCG
PVT1	AACTTCCTTTGGGTCTCCCTATG	CAGGGAGTATGGTCAGCTCAAG
RAD21	CTGTTTTCTTTACCTGCTCAGCC	TTATCTGCCTCTCCTCCTTTCCT
SOGA1	AGGGGCTCTCTAAGGAGATTCTT	TTAGAGTCCAAGTCACCAGTGC
SPTB	ACCAGGAGATGAAGACCTTGATG	GTGTTCAGATGGTCCTGGTACTG
TFF1	AAACGCTGGCAACGACCTGT	ATGCTCTGCGCGGGCTAC

TRIB3	CAGATCGTGCAACTGCTGTG	GCTTCCTGGACGGGGTACA
XBP1	ATACTTGGCAGCCTGTGACC	GGTCCACAAAGCAGGAAAA

TABLE 2.4 ChIP qPCR primer pair sequences.

Target Name	Oligonucleotide sequence (5' – 3')	Oligonucleotide sequence (3' – 5')
FR01 (EIF2AK4)	AATGACATGCCAACTTCATTCCT	AAAGGCTGGGTTTAGATTCCAAC
FR02 (ENOX2)	GTTAAGCAAAGTTGCTCCAAGTG	TGATCTCACTGGGTCCTCTTAGG
FR03 (ADCY1)	GGGGTTATGCAAAGCATTTAGG	TGAGCAGGTCAACTGGGATTAAG
FR04 (-)	GCCTAAAGCGGGAGTCAGTAAAT	CATATTGTCTGTGGTGGCTTTTG
FR05 (RP11-317J19.1)	AACTGCAAATTGCTTTTGTGTCA	AACAAAGGGGAATTTGATCTGGT
FR06 (ESR1)	ACGTTAACTGCAGCCCACACT	CCTGAAACTCGGTTTTATTGACA
FR07 (TPRG1)	CACCCAGGACAAGAAACTACAGG	CACACGACTGGAATAGCATGAAG
FR08 (GMDS)	TCTTCCTGTCAAAGGAGGGAAAT	TGCTAAGCACTGAGTCACCTGAT
FR09 (-)	TTTGAGCAAAAGCCACAAAGTAG	TTTGGTTGCTTCTAGGTCAGAATG
FR10 (-)	CTCTGGGAGCTCTTGAAATGTCT	CCCAGACAACTTTGCAAGTAAGG
FR11 (-)	ATCTACCAGAGCCATTTCTGCAA	CGCGATTCAAATCAATACAGGAA
FR12 (EXT1)	CGCCTCCTTTCCGATTCACT	CAGGCGCACAGGTACCATTTT
FR13 (-)	CCCTGGAAACGCATTCTCTG	GAGCCAGCCTAGCCTCAGATTTA
FR14 (GCH1)	AGCCAATCAACGTCCCACAG	CAGATCCGCTTCTGGTTTCTTAC
FR15 (CEP250)	AGGTGTCCATCCAGCTCAGG	GTTTCTGTTGGGGTTTGAATTAGC
FR16 (CERK)	GCTTCTTACCAGCGTCCTGACTA	CTCCGGCACGGATTTAAACG
FR17 (ZNF365)	CTTTTTCTCTAGCCGCCCTTTC	TTCAAACTCACCGCGCTAAAG
FR18 (ANKS1B)	TGTGTGAAAATGGCTGTGATCTT	GGGTCTTTTGTTGTTCTCTGCAT
FR19 (CTD-2509G16.5)	CTCCCCAGTGCTAGAGAAGATTG	ATTCATCTCTGTGCCTCATGTTT
FR20 (DENND1A)	CGCTAGAGGGAACGAGCTGTAGT	ATCCAATCACCAGCAGCCTCT
NC01 (-)	TAGGGAACTAGCAGAAGGGCAAA	TTGTTTGGGAGTGATGGAAAACG
NC02 (-)	GATGCGAGAGTGATGGTAGAGAA	CATTTTCACTTTGGCTCTGAGGC

NC03 (-)	CATGCCATCAGTTTTGCTTTCCT	TCCATCTTTTCCATCATCTTTTCCA
NC04 (-)	CAACCTATGTGCTTTGAAGGAGG	ACTCCGCCTTCACTTCAAAGTC
NC05 (-)	TTAAGCTCCCTCTTTGCTCAGTT	CACTATGTTAGGTGTGCCCTTGT

2.4.3 RT-qPCR

RT-qPCR was performed using the FAST SYBR Green Master Mix (Applied Biosystems), 2 ng/μl cDNA and 0.4 μM of each specific primer pair (Applied Biosystems). qPCR was completed on an ABI PRISM 7900HT (Applied Biosystems) using the following parameters: initial activation of 20 secs at 95 °C, 40 cycles; denaturation for 1 min at 95 °C; annealing/extension for 20 secs at 60 °C; final melting curve for 15 secs at 95 °C and later for 15 secs at 60 °C. Quantification of mRNA levels was analysed using Excel®. As a control the established housekeeping genes GAPDH or L19 were used to normalise gene expression between samples. The amount of mRNA was calculated using normalisation method. All measurements were performed in triplicates. All primers were diluted in RNase-free water to a final concentration of 1 μM and stored in -20°C according to the manufacturer's manual.

2.5 DNA Analysis

2.5.1 Plasmid Constructs and Expression Vectors

FuGENE6 transfection reagent (Roche Diagnostics, West Sussex, UK) was used to perform overexpression according to the manufacturer's protocol. Briefly, plasmid DNA was diluted in 1:3. 24 hours after transfection, cells were reseeded in T75 flasks and another 24 hours of incubation has waited.

The constitutively active form of FOXM1, pcDNA3-ΔN-FOXM1, was generated by deletion of the FOXM1 N-terminal repressor domain (106) and was cloned by our laboratory.

2.5.2 DNA Quantification

The concentration of DNA was measured using a ND-1000 spectrophotometer (NanoDrop Technologies) which measures real-time the DNA concentration upon the additional of 1µl of sample.

2.5.3 Subcloning

2.5.3.1 Extraction and Purification of DNA

DNA was cleaned with PCR purification kit or Reaction cleanup kit (Qiagen, Crawley, West Sussex, UK) according to manufacturer's protocol. Ethanol precipitation was used where the concentration of DNA permitted; 0.1 volumes of 3M sodium acetate (pH 5.2) and 2.5 volumes of 100% ethanol were added to the DNA solution. The mixture was stored at -80°C for 20 min and centrifuged at 14,000rpm for 30 min at 4°C. The pellets were washed with 70% of ethanol, re-centrifuged for 5 min, air-dried and re-suspended in 10mM TRIS buffer or sterile water.

2.5.3.2 Bacteria Transformation and Culture

1µl of DNA (<10ng) and 20-50µl competent cells were incubated on ice for 30 min. OneShot® TOP10 Chemical competent cells (Thermo Fisher Scientific, Waltham, MA, USA) were used for subcloning ligated DNA. The cells were then heat shocked at 42°C for 45 sec and returned to ice for 2 min. The transformed cells were then spread on agar plates with appropriate selection antibiotics (ampicillin 100µg/ml in water or kanamycin 50µg/ml in water).

A single colony was selected from a plain LB agar plate and inoculated in 5ml of LB broth (Sigma) with selection antibiotics, at 37°C for a minimum of 6 h in a shaking incubator. 1ml of this culture was then added to 250ml LB in a 500ml flask and left to shake in the 37°C shaking incubator overnight.

2.5.3.3 Plasmid Miniprep and Maxiprep

DNA was purified from the overnight grown cultures using a Miniprep Plasmid Purification Kit (Qiagen). For isolation of larger amounts of DNA, a Maxiprep Plasmid

Purification Kit (Qiagen) was used instead. Both kits were utilised according to manufacturer's protocol.

2.5.3.4 Polymerase Chain Reaction (PCR)

Reaction mixtures contained 1µg DNA template, 0.5µM primers, 200µM dNTP, 0.5U Taq DNA polymerase in a 25µl reaction. PCR amplification was performed using a Taq DNA polymerase kit (Qiagen) with a Gene Amp® PCR system 9700 (Applied Biosystems) using the following program: Denaturation for 1 min at 94°C, annealing for 1 – 2 min at 55 – 65°C depending on the oligonucleotides, extension at 72°C for 1 min for 33 cycles followed by a final extension for 10 min at 72°C. Optimisation of the primer annealing temperatures was completed by a gradient of increasing temperatures from 50 – 65°C for 35 cycles, followed by change of magnesium concentration.

2.6 Transfection of Plasmid DNA and Expression Vectors

Cells were seeded into six well plates or $10cm^3$ cell dishes to a confluency of roughly 60%. Plasmid DNA was transfected using a ratio of Fugene 6 (µl) to DNA (µg) of 3:1. After 15 min of incubation, the Fugene/DNA solution was mixed directly to each well or dish and the cells were incubated for 6h at 37°C, then cells were washed in warm PBS and fresh media was added. 24 h post transfection, cells were either directly harvested for western blot and RT-qPCR analysis or re-plated for other conditional experiments.

2.7 RNA interference (RNAi)

Seeded cells were transfected after reaching a minimum of 60% confluency in culture flasks. RNAi transfections were done using Oligofectamine™ according to the manufacturer's protocols (Thermo Fisher Scientific). Briefly, cells were seeded at a density of either 1.5 x 10^5 in 6-well plates for WB and RT-qPCR a day before transfection. For single knockdowns, the RNAi concentration was 20 nM and for double knockdowns (silencing of two different targets) the total concentration was 40

nM (20 nM of each target siRNA). The efficiency of siRNA transfection was evaluated either by RT-qPCR and/or Western Blotting. As a control, all RNAi experiments included a non-targeting siRNA. All siRNA oligos were purchased from Dharmacon ON-TARGET*plus* SMARTpool siRNA catalogue and prepared according to manufacturer's protocol. Briefly, siRNA pools were resuspended in 1X resuspension buffer for a final concentration of 20 μM. Twenty-four hours after transfection, cells were reseeded in other flasks for another 24 hours incubation.

2.8 ChIP – qPCR

Six million cells in 150 mm Petri dish were washed twice with cold PBS and fixed with 1 % formaldehyde (FA) in PBS for 10 minutes. 1 % FA was removed, the plate was washed with cold PBS twice and the remaining FA was quenched with 0.25 M Glycine solution in water for 5 minutes. Glycine solution was then removed, plates washed twice with cold PBS and layered cells harvested using scraping buffer (100 mM TRIS-HCl, pH = 9.4, 0.1 % SDS) with protease inhibitor complex (PIC) (Roche; Cat. No.: 11697498001). Cells were extracted and washed 3 times with following buffers containing PIC: 1) PBS, 2) Buffer 1 (0.25 % Triton X – 100, 10 mM EDTA, 0.5 mM EGTA, 10 mM HEPES, pH = 6.5), 3) Buffer 2 (200 mM NaCl, 10 mM EDTA, 0.5 mM EGTA, 10 mM HEPES, pH = 6.5).

Washed cell pellets were then lysed using lysis buffer (1 % SDS, 10 mM EDTA, 50 mM TRIS-HCl [pH = 8.1]) with PIC for 10 minutes on ice. The lysates were then sonicated using Bioruptor under the following conditions: 16 cycles of High Duty, 30 seconds on and 30 seconds off). Sonicated lysates were spun down and supernatants were diluted with equal volume of Buffer D (1 % Triton X – 100, 2 mM EDTA, 20 mM TRIS-HCl [pH = 8.1], 150 mM NaCl) with PIC. An aliquot was saved for the inputs and remaining lysates were equally split into two vials. One vial was probed with pure Dynabeads Protein A (ThermoFisher Scientific; Cat. No.: 1002D), the other was probed with an identical amount of Dynabeads Protein A conjugated with 5 μg of FOXM1 antibody (Novus Biologicals; Cat. No.: NBP1 – 30961). Both vials were incubated overnight on a rotator at 4 °C.

After 24 hours, bead mixes were washed 3 times with the following buffers with added PIC: 1) TSE I (0.1 % SDS, 1 % Triton X – 100, 2 mM EDTA, 20 mM TRIS – HCl [pH = 8.1], 150 mM NaCl), 2) TSE II, 3) TE (1 mM EDTA, 10 mM TRIS – HCl [pH = 8.0]). Washed DNA was eluted from both beads and input samples with Elution Buffer (0.1 M $NaHCO_3$, 1 % SDS) at 65 °C overnight. Eluted samples were treated with RNase A (Qiagen; Cat. No.: 19101) for 1 hr at 37 °C and Proteinase K (New England Biolabs; Cat. No.: P8107S) for 2 hrs at 65 °C. All samples were purified with SPRI select beads (Beckman Couler; Cat. No.: B23318) at 1.0 ratio (DNA: Beads) according to manufacturer's protocol or Wizard SV Gel and PCR Clean-Up System (Promega) according to manufacturer's instructions, then eluted with endonuclease free water. Such DNA samples were used as templates for qPCR with stated run condition: 95 °C for 10 mins, 50 cycles of 15 s each at 95 °C, 60 °C for 30 s, 72 °C at 30 s) with 2x SYBR Green PCR Master Mix (ThermoFisher Scientific; Cat. No.: 4309155). To confirm target gene binding, primers designed to span across specific gene promoter regions were used to run RT-qPCR which sequences are listed above.

2.9 Co-immunoprecipitation (Co-IP)

Co-immunoprecipitation was performed to investigate protein-protein interactions. Protein lysates were extracted using a co-IP lysis buffer (50mM HEPES pH 7.5, 150mM NaCl, 2mM EDTA, 10% glycerol (v/v), 0.5% Triton-X (v/v), 10mM NaF, 10mM β-Glycerophosphate pH 7.3, 0.1mM Na_3VO_4, 1mM PMSF) to precipitate target protein and co-precipitate binding protein complexes (15µls. The lysates were pre-cleared with agarose beads (15µl per sample) for 30 min. Then, lysates were centrifuged, and supernatant was transferred to a new Eppendorf tube. Protein concentration was then measured and 0.3-0.5mg of protein extract was incubated with selected antibody and beads for 4 h rotating at 4°C. The antibodies used in co-IP were the same as the ones selected for ChIP. Immunoprecipitated solutions were then washed/centrifuged 3 times with the same lysis buffer and samples were used for separation on SDS-PAGE gel.

2.10 Assay for transposase accessible chromatin with high throughput sequencing (ATAC–Seq)

Ten million MCF7 cells were transfected with siNSC control (Dharmacon; Cat. No.: LQ – 009762 – 00 – 0002) or siFOXM1 (Dharmacon; Cat. No.: LQ – 009762 – 00 – 0002) using Oligofectamine™ (Thermofisher Scientific; Cat. No.: 11668019) according to manufacturer's instructions. Cells were cultured for further 48 hours and ATAC-Seq libraries were prepared as described in literature[227]. 50,000 cells were harvested, and nuclei extracted using Nuclei EZ kit (Sigma Aldrich; Cat. No.: NUC – 101), chromatin was tag using Nextra DNA Sample Preparation Kit (Illumina Cat. No.: FC – 121 – 1030). DNA was purified using columns (Zymoresearch; Cat. No.: D4013) and PCR amplified using supplied barcoded primers. Libraries were paired end deep sequenced on Illumina Hiseq2500 platform. Paired-end ATAC-Seq raw profiles were processed using the ChIP-Seq pipeline with left-mate pair and right-mate pair reads merged. Normalised read density profiles were produced using deeptools applying the following parameters: --binSize 10 –normalizeTo1x 2451960000. ATAC-Seq signal in respective regions was extracted and plotted.

2.11 Bioinformatics

2.11.1 Data analysis pipeline and visualisation

All sequencing data were assembled using hg19 or hg38 reference genomes, remapping was performed when swapping between difference versions of reference genome for downstream analyses.

Bioinformatic analyses were performed on Python 3 for Mac OS X and Rstudio Version 1.1.463 for Mac OS X unless otherwise stated. Venn diagrams to show the overlapped and independent targets between different protein molecules from ChIP-seq analyses were plotted using Venny 2.1 (107). All ChIP-seq data was downloaded and reads mapped to the hg38 reference genome and identified statistically significant peaks (108). Bioconda package 'bedtools' was used to sort bedfiles and 'bedops' was used to intersect and identify overlapped peaks of different TFs. R Bioconductor package 'ChIPseeker' (109) was used to annotate the genome-wide distribution of ChIP-seq binding peaks, 'ClusterProfiler' (110) was used for pathway enrichment analysis of identified genes and GO term enrichment was plotted using 'enrichGO' in the dot plot.

All ChIP-seq peaks are aligned and annotated with human genome hg38 'TxDb.Hsapiens.UCSC.hg38.knownGene'. BEDOPS, an open-source command-line toolkit was used to intersect peaks (111, 112) (Seattle, Washington, USA). Peaks visualisation were performed using BigWig files on the Integrative Genomics Viewer IGV browser using human genome hg38 (Broad Institute, Cambridge, Massachusetts, USA). Heatmaps were generated with GENE-E ((Broad Institute, Cambridge, Massachusetts, USA).

2.11.2 ChIP-Sequencing data collection

Raw ChIP-Sequencing profiles and input DNA controls were gathered from the Gene Expression Omnibus (GEO) database detailed with cell line, target proteins and GEO series accession number. MCF7 and MDA-MB-231 FOXM1 profiles: GSE40762. MCF7 ESR1 profiles: GSE32222. All MCF7: RAD21, STAG1: GSE25021. CTCF, FOXA1: GSE25710. EP300: GSE29073. GREB1: GSE41561. H3K27Ac: GSE31755. HepG2, HeLa-S3 and K562 H3K27Ac: GSE29611. MDA-MB-231 H3K27Ac: GSE49651. All MCF7: EGR1, ELF1, FOSL2, GABPA, GATA3, JUND, MAX, PML, SRF, TAF1, TCF12, TEAD4, SIN3A and TCF7L2: GSE32465. MCF7 MYC: GSE33213.

2.11.2.1 Cistrome Data Browser (Cistrome DB) Data Mining

ChIP-sequencing data was mined from dataset browser. ChIP-seq analyses of TFs of interest were filtered, and unique dataset ID specified as follow:

TABLE 2.5 List of ChIP-seq data downloaded from Cistrome DB.

Antibody	Cell Line	GEO / ENCODE	CistromeDB	References
FOXM1	MCF-7	GSM1000995	39443	(113)
FOXM1	MDA-MB-231	GSM1000999	39460	(113)
DDIT3	MLS402	GSM3027229	103939	(114)
ESR1	MCF-7	GSM2071221	71858	(104)

ESR1	MCF-7 (treated with oestrogen)	GSM2071222	71857	(104)
H3K27AC	MCF-7	GSM1382473	54476	(115)
H3K27AC	MCF-7 (treated with oestrogen)	GSM1382483	54486	(115)

2.11.3 RNA-Sequencing

Tamoxifen and fulvestrant resistant MCF7 RNA-Seq data were obtained from the European Nucleotide archive using the accession: PRJEB11448. Paired-end RNA-seq profiles in FASTQ format had poor quality reads and adapters removed using Trimmomatic. The transcripts were then mapped to the UCSC hg19 genome using tophat2 generating BAM files. Cufflinks were used to generate FPKM libraries for each sample. These samples were then overlapped with FOXM1 genes identified from MCF7 FOXM1 ChIP-Sequencing peaks within the gene body and 25 kb upstream. Table S1 displays the complete list of all differentially expressed genes selected by FOXM1 peaks from the top 200 ranked RNA-Sequencing genes.

2.11.4 ChIP-Sequencing Data Processing

Raw ChIP – Seq profiles were converted into fastq format using Fastq-dump from the SRA toolkit. The adapters were trimmed and poor-quality reads removed using Trimmomatic. The trimmed reads were mapped to the UCSC hg19 human genome using Bowtie2. Mapped reads were sorted using Samtools. Picard tools were used to remove multiple reads and Bedtools to remove blacklisted regions using ENCODE mappability regions. Same pipelines used for input DNA and ChIP – Seq profiles.

2.11.5 ChIP-Seq Peaks Calling and Overlapping

Peaks were called using Model-based Analysis of ChIP-Seq 2 (MACS2). ESR1 peaks were called using input control with a q-value of 0.01, MDA-MB-231 peaks were called using input control and q-value of 0.05. FOXM1 peaks were called stringently with a q-value of 0.01 and identified in reference to Sanders et al. Default parameters and settings were used. Discrete viable peaks were obtained from typical chromosome

regions 1 – 22, X and Y and subsequent common peak lists were produced from replicates. 6431 FOXM1, 37473 ESR1, and 2915 MDA-MB-231 peaks were left for downstream analysis. ChIP-Seq peak files were overlapped with Bedtools intersect. MCF7 FOXM1 and ESR1 peaks were overlapped with 10% and MCF7 peaks were overlapped with MDA-MB-231 peaks at 1 bp.

2.11.6 Genomic Distribution of ChIP-Seq peaks

Distribution of ChIP-Seq peak locations throughout the genome was identified using UCSC Ensembl gene annotation files. Fold-enrichment was calculated as a ratio of the fraction of respective classes in the genome over the fraction of this class in the genomic composition using a log2 scale.

2.11.7 Motif Enrichment Analysis

Motif Enrichment Analysis was conducted using AME from the MEME suite. Fasta files were generated using getfasta from Bedtools for ESR1-positive and ESR1-negative FOXM1 peaks. The corresponding nucleotide sequences were analysed in AME to discover the significantly enriched motifs in each sequence utilising the JASPAR CORE 2016 vertebrate motif database. FOXM1 peaks were shifted by 50 kb as a control. Discriminative motif analysis was also performed in AME discovering the enriched motifs relative to the control sequence, where FOXM1 ESR1-negative sequence was used for FOXM1 ESR1-positive analysis, vice versa.

2.11.8 Gene Ontology Analysis

Gene Ontology (GO) Analysis was performed on FOXM1 ESR1-positive and ESR1-negative peaks, using all FOXM1 and MDA-MB-231 peaks as two separate controls. R package biomaRt, the Ensembl archive release version 75 and R package GenomicRanges were used to map FOXM1 ESR1-positive and ESR1-negative ChIP-Seq peaks to within gene body and 25 kb upstream of the transcription start site (TSS), generating two sets of gene list. R package GOstats was used to perform hypergeometric tests to identify the top 100 significantly overrepresented GO terms for each list, using the GO database from the GO consortium. These lists were

collapsed removing redundant terms leaving the top 20 over-represented GO terms. Hypergeometric tests were conducted for over- and under-representation for FOXM1 ESR1-positive and ESR1-negative, MCF7 and MDA–MB–231 peaks. Subsequent FDR corrected p-values were used to calculate, log10 (underrepresentation) – log10 (overrepresentation) for each of the top 20 terms across all 4 categories mentioned. Heatmap plotted with R package GGplots. complete 20 terms, top 5 from each class or web-based tool Morpheus (Broad Institute).

2.11.9 Comparative Binding Analysis

MCF7 TF ChIP-Seq profiles were pre-processed using the pipeline as described. Deeptools was used to perform binding analysis, utilising BigWig files generated from the BAM ChIP-Seq signal with the bamCoverage tool. Heatmaps and binding profiles were produced using FOXM1 ESR1-positive and ESR1-negative 1 pb peak summit files and FOXM1 summits shifted by 50 kb as a control. The summits were extended by 5000 bp either side and binding profiles plotted.

2.11.10 ATAC-Seq Data Analysis Pipeline

The data analysis is separated as three important steps: (i) pre-analysis (quality control and alignment), (ii) core analysis (peak calling), and (iii) further analysis at peaks, motifs, nucleosomes, and TF footprints.

2.11.11 ATAC-Seq Pre-Analysis: Quality Control and Alignment

The first step of ATAC-Seq analysis involves pre-alignment quality control, read alignment to reference genome, and post-alignment quality control and processing.

2.11.11.1 Pre-alignment quality control

The pre-alignment quality control and read alignment steps are standard as ChIP-Seq. Overrepresentation of Nextera sequencing adapters is often observed, hence this overrepresentation and low quality bases were removed for accurate read alignment using Trimmomatic with inputs of known adapter sequences.

2.11.11.2 Alignment

After read trimming, FastQC can be ran to validate the successful removal of adapter and low quality bases. Trimmed reads are then mapped to UCSC hg19 human reference genome using Bowtie2 for short paired end reads. Mapped reads of 50 million were set for open chromatin detection and differential analysis using Samtools.

2.11.11.3 Post-alignment processing and quality control

After sequence alignment, basic metrics of the aligned BAM files were collected using Picard. Improperly paired and low mapping quality reads were removed. Mitochondrial genome, which are more accessible due to lack of chromatin packaging, and ENCODE blacklisted regions, were discarded with Bedtools using ENCODE mappability regions. Duplicated reads were also removed to improve biological reproducibility using Picard MarkDuplicates.

2.11.12 ATAC-Seq Core Analysis: Peak Calling

The second step of ATAC-Seq data analysis is to identify accessible regions (known as peaks). MACS2 was used as the default peak caller for the ENCODE ATAC-Seq pipeline. Similar to DNase-Seq, the enzymatic cut by Transposase 5 will introduce bias due to binding preference and was adjusted when calling peaks. BAM file was converted to BED format and peaks called using MACS2.

2.11.13 ATAC-Seq Further Analysis

Further analysis of ATAC-Seq shows several aspects of transcriptional regulation at peaks, motifs, nucleosomes, and TF footprints. Binned genomes (or consensus peaks) were identified, normalized, and fragments counted and compared between siNSC and siFOXM1, as well as FOXM1/ESR1 positive and negative regions statistically. As stated, normalised read density profiles were produced using deeptools applying the following parameters: --binSize 10 –normalizeTo1x 2451960000. ATAC-Seq signal in respective regions was extracted and plotted using BOX Plots.

2.11.14 BOX Plots

The coverage over 1000 bp around the peak summit was calculated for ChIP-Seq and input samples, added a pseudo count of 0.01 and computed the enrichment ratio of ChIP-Seq over DNA input, which was transformed into log values.

2.11.15 Genomic Distributions

Blacklisted regions, RefSeq genes, and RepeatMasker data were downloaded from the UCSC table browser. STARR-Seq peak genomic distributions were done as previously described combining hg19 RefSeq genes, blacklisted and RepeatMasker annotations. All repeats from RepeatMasker and blacklisted regions were combined.

2.11.16 Core Promoter Motif Analysis

Position-weight matrices (PWM) for core promoter motifs conserved from fly to human were extracted from FitzGerald et al., 2006; Ohler et al., 2002. PWMs for TATA-box, Initiator (INR), downstream promoter element (DPE) and E-box were obtained from (Ohler et al., 2002) and PWM for TCT motif from (Parry et al., 2010). Each FOXM1 ChIP-Seq peak was assigned to the transcription start site (TSS) in 3 different ways: nearest TSS, within 5 kb distance or gene body. In the nearest TSS, each FOXM1 peak was uniquely assigned to the nearest TSS given RefSeq gene annotation. According to a 5 kb distance method, if a peak was within 5 kb distance of any TSS, it would be assigned to it. Notably, multiple peaks to TSS assignments are possible. In gene body assignment, if a peak was within the gene body, 5 kb upstream to the gene's TSS, the peak was assigned to this TSS – multiple assignments are possible. Core promoter element presence was counted using the MAST tool from MEME suite with $p < 10^{-3}$ and e-value < 200 cutoff in respective classes of FOXM1 ESR1-positive and ESR1-negative target TSS. Enrichments were calculated as a normalised abundance of motif occurrence in ESR1-positive over ESR1-negative target TSS. To both classes, a pseudo count of 1 was added.

2.11.17 Statistical Analysis

Unless otherwise stated, all data were presented as mean ± standard error of the mean (SEM) of three independent biological replicates. For qPCR experiments, each

independent experiment includes three technical replicates. Statistical analyses were performed using GraphPad Prism (GraphPad Software Inc). Statistical significance was labelled with an asterisk; *p<0.05, **p<0.01, ***p<0.005, ****p<0.001.

3 The functional differences between the FOXM1/ESR1 common transcriptional targets and FOXM1 specific targets in MCF-7 cell lines

3.1 Introduction

Despite the advances in breast cancer treatments, endocrine therapy resistance and the development of fatal metastatic tumours remain a common clinical concern (12, 116). A meta-analysis of gene expression signature from independent breast cancer patient data identified *FOXM1* as one of the 117 genes predictive of breast cancer survival (117). Additionally, overexpression of *FOXM1* has been associated with resistance in breast cancer therapy, therefore predictive of poor clinical outcome (40). Breast cancer covers a heterogeneous subset of diseases, all developed from breast tissues. Among different breast cancer subtypes, roughly 70% of breast cancers express ESR1 and fall under ESR1-positive subtype. In this subtype, ESR1 signalling cascade is the determining pathway contributing to tumour growth and disease progression in these tumours. In breast cancer cells, *FOXM1* expression is correlated with ESR1 activity, vice versa (67, 118, 119); additionally, this positive feedback loop contributes to endocrine therapy resistance (120). Literature evidence suggests that ESR1 and FOXM1 may act as co-regulators (67). Moreover, *FOXM1* has been previously stated as an ESR1 target gene and has been proposed as a prognostic marker of endocrine-positive cancers (121). Furthermore, anti-estrogen and chemotherapy treatment resistances have been associated with overexpressed *FOXM1* (56). In this chapter, I investigated two distinct FOXM1 transcriptional clusters, differentiated by the presence of ESR1 co-binding.

The raw data of the binding profiles investigated in this chapter were collected by ChIP analyses, with the detected RT-qPCR signals normalised to controls to deduce the significance of the bindings. ChIP is a molecular technique to study the relationship between a DNA-binding protein and a gene sequence. In most scenarios, the protein of interest is often a TF, histone molecule, co-activator, or co-inhibitor that catalyses or suppresses specific processes. ChIP is useful when the gene sequence of interest on a specific DNA region is known, as designated gene primer construct could tag on

these targeted regions, amplified by RT-qPCR, and later detected. In some circumstances, the research hypothesis is left open-ended – for example, a condition is altered, and we are interested to find out the resulting impact on the binding profile of a specific TF. The condition may be knock-out or knock-in of a gene, upregulation or mutation of a specific gene or protein, or changes in metabolomics, to name a few. During these situations, ChIP-seq comes in extremely fitting. ChIP-seq combines ChIP with massively parallel DNA sequencing to identify the unknown binding sites of DNA-associated protein of interest. Due to the limitations in resources, I utilised available ChIP-seq data uploaded on the public database, Cistrome Data Browser (CistromeDB) in place of conducting ChIP-seq experiments from scratch. GEO and CistromeDB ID of each ChIP-Seq data presented in this chapter are stated in Materials and Methods Section **2.11.2.1**. Visualisation of the transcriptional binding profiles and overlaps are illustrated with heatmaps and boxplots for the relevant factors to display the intensities and significances of these overlaps, and each of the pattern guides us to better understand the two FOXM1 clusters – FOXM1/ESR1 positive and ESR1 negative regions. Genomic distribution of ChIP-seq peak locations throughout the genome was identified using R/Bioconductor package "ChIPseeker" and ICSC Ensembl gene annotation files (109). To further understand the binding profiles of these FOXM1 binding clusters, motif enrichment analysis was performed to screen for consensus sequences that exist in these binding regions. Additionally, co-factors bindings were scanned by respective ChIP-seq data on public database. The top associated biological pathways of the binding profiles were investigated using GOanalysis.

ATAC-seq was performed in collaboration with **Dr. Raffaella Lucciola**, a then PhD student of Professor Jan Brosens from University of Warwick to study the chromatin accessibility in MCF-7 cell line upon FOXM1 depletion. ATAC-seq identifies loosely packed and accessible DNA regions by probing open chromatin with hyperactive mutant Tn5 Transposase that inserts sequencing adapters into open regions of the genome. Tn5 transposase cleaves and tags double-stranded DNA with sequencing adaptors. The tagged DNA fragments are then purified, PCR-amplified, and sequenced using next-generation sequencing (122, 123).

Genetic studies have shown that mutations in ESR1 are found in >20% of tumours that progress on endocrine therapies. ESR1 coding region mutations are detected in approximately 20% of aromatase inhibitor-resistant breast cancer. There is evidence

that find ESR1 mutations mostly happen after progression on endocrine therapies, shows that treatment-selective pressures are likely to regulate the acquisition of ESR1 mutations. Notably, the great majority of these mutations in the ESR1 localise to just a few amino acids within or near the critical H12 region of the ESR1 hormone ligand binding domain (LBD), where they are likely to be single allele mutations. 'Hotspots' in the LBD are at the consecutive amino acids L536, Y537, and D538, structural studies state that these residues control the agonist state of the LBD and that their mutation stabilises the receptor in the agonist state to promote the recruitment of co-activators, and consequently facilitate transcription of ESR1 target genes. The CRISPR-Cas9 technique for genome editing provides a highly targeted approach for cell deletion and insertion mutagenesis in mammalian cells, possibly permitting the more accurate evaluation of the functional role of gene mutations in model mechanisms (104). This technique was utilised to introduce mutations in the genomically encoded *ESR1* gene to facilitate direct functional comparison of wild-type and mutant *ESR1* breast cancer cells towards drug treatments. Particularly, this mutant would provide a model to study therapeutic approaches to target breast cancer cell lines harnessing *ESR1* mutation, which represents a main trait of breast cancer cell lines. In this chapter, I focused on Y537, the most frequently mutated ESR1 residue, in the estrogen-responsive and anti-estrogen sensitive MCF-7 cell line. In MCF-7 cell line with a genomically encoded ESR1-Y537S mutation, ligand-independent recruitment of ESR1 has been observed and regulation of gene expression in the absence of estrogen (18, 104). More notably, these cells grew in the absence of estrogen and show feature of resistance to anti-estrogens. Using MCF7-Y537S cell line kindly shared by **Dr. Alison Harrod**, I aimed to understand the relationship of FOXM1 and ESR1 in endocrine treatment response in breast cancer ESR1 positive cell line.

3.2 Aims and Objectives

ESR1-positive breast cancer has been widely studied, where prognostic factors and treatment options are abundant. However, the full picture of ESR1-modulated transcriptome, ESR1 mutations, their therapy response, and functional relationships are not precisely elucidated. Given the correlation between ESR1 (expression, cistrome, and mutations) and breast cancer development, understanding the exact mechanisms are essential to improve breast cancer survival. Additionally, endocrine therapies are often ineffective after a short initial responsive period in breast cancer treatments, resulting in therapy resistance and breast cancer recurrence. FOXM1 is undoubtedly a TF involved in the regulation of cell cycle. Aberrant expression and activity of FOXM1 leads to various cancer initiation and progression. The downstream regulation by these TFs are finetuned by co-factors, whether to activate or suppress gene expressions. In this chapter, I investigate the relationship between FOXM1 and ESR1, and how it affects downstream target genes and therapy responses, in breast cancer models MCF-7 cell line and MDA-MB-231 cell line. Furthermore, ESR1 mutations, particularly at hotspot Y537S, is often detected in endocrine-resistant metastatic breast cancer cells. Therefore, dissecting MCF7-Y537S mutation pathway may lower the chance of acquired endocrine resistance and prevent breast cancer relapse.

3.3 Results

3.3.1 ChIP and ATAC sequencing to assess transcriptional profile and chromatin accessibility of target regions

ChIP-seq and ATAC-seq were conducted in breast cancer MCF-7 cell line. The workflow and analysis sequences were illustrated in Figure 3.1. Cell lines were cultured and harvested for analysis to study transcriptional profile (ChIP-seq) and chromatin accessibility (ATAC-seq). For ChIP sequencing, cells were dissociated from culture dishes, isolated chromatin was fixed and fragmented with detailed protocols described in **Section 2.8 ChIP – qPCR**. Magnetic beads conjugated to antibody specific to epitopes of the target proteins were used to precipitate fragments bound to the protein of interest. ChIP-Sequencing data analysis followed guidelines in **Section 2.11.4 ChIP-Sequencing Data Processing**. Adapter primers were then used for sequencing and quality control assessment. The adapters were trimmed, and poor-quality reads were eliminated using Trimmomatic. The trimmed reads were mapped to the UCSC hg19 human genome using Bowtie2. Mapped reads were sorted using Samtools. Picard tools were used to remove duplicated reads and Bedtools to remove blacklisted regions using mappability regions. Same pipelines used for input DNA and ChIP – Seq profiles. Peaks were called using Model-based Analysis of ChIP-Seq 2 (MACS2). ESR1 peaks were called using input control with a q-value of 0.01, MDA-MB-231 peaks were called using input control and q-value of 0.05. FOXM1 peaks were called stringently with a q-value of 0.01 and identified in reference to Sanders et al (67). Default parameters and settings were used. Discrete viable peaks were obtained from typical chromosome regions 1 – 22, X and Y and subsequent common peak lists were produced from replicates. Total of 6,431 FOXM1, 37,473 ESR1, and 2,915 MDA-MB-231 peaks were left for downstream analysis. ChIP-Seq peak files were overlapped with Bedtools intersect. MCF7 FOXM1 and ESR1 peaks were overlapped with 10% and MCF7 peaks were overlapped with MDA-MB-231 peaks at 1 bp. On the other hand, ATAC-Sequencing was used to investigate chromatin accessibility of cell lines under certain experimental conditions. Detailed protocols are stated in **Section 2.10 Assay for transposase** accessible chromatin with high throughput sequencing (ATAC–Seq, using Illumina Hiseq2500 platform.

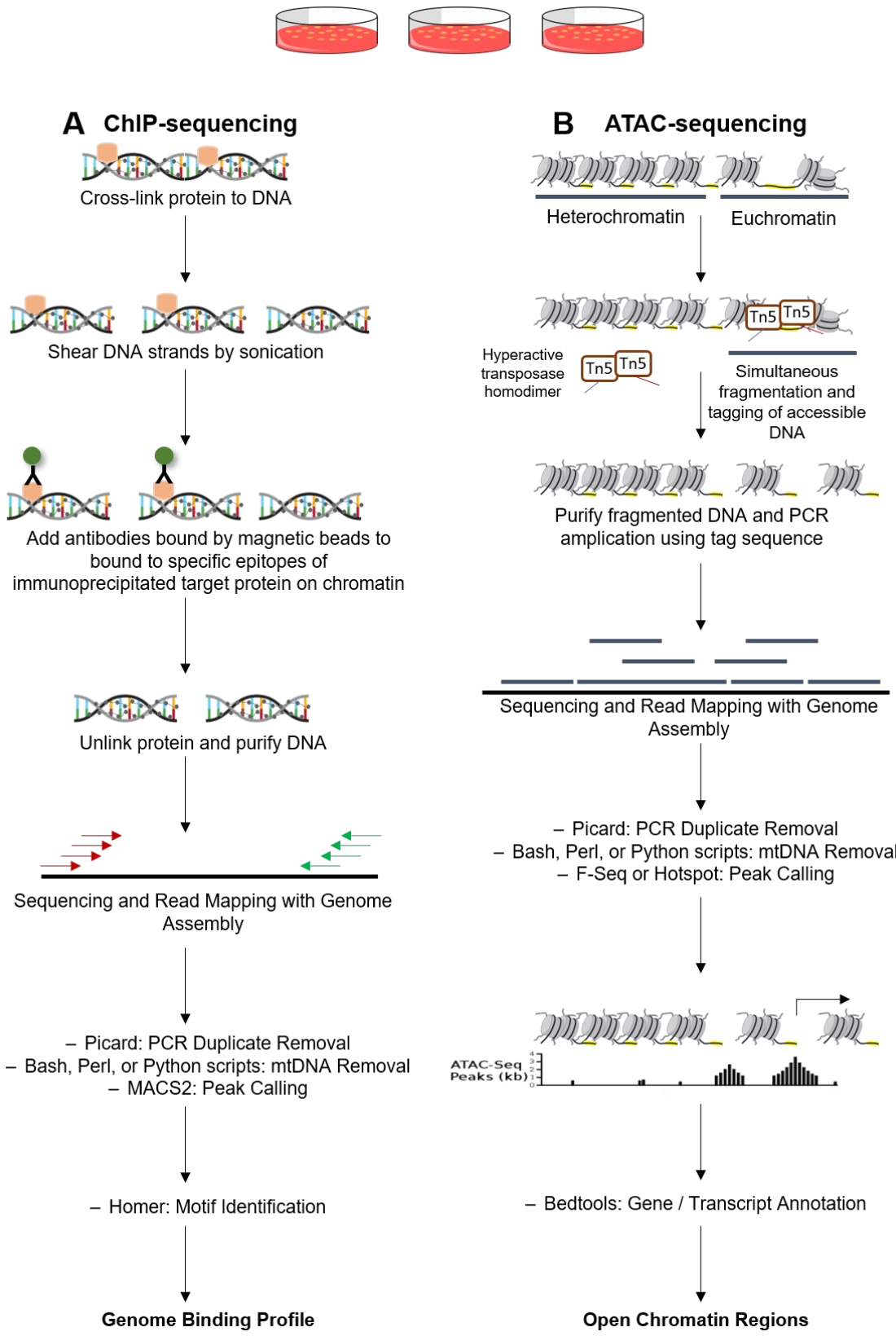


Figure 3.1 (A) ChIP-Sequencing (ChIP-Seq) and (B) ATAC-Sequencing (ATAC-Seq) workflow and analysis.

ChIP-Seq to decipher target TF binding on gene promoter regions and ATAC-Seq to investigate the chromatin accessibility under target conditions in harvested cell lines.

3.3.2 FOXM1 and ESR1 share similar transcriptional binding profile in MCF-7 cell line, and there was little overlap for FOXM1 binding profile between MCF-7 and MDA-MB-231

To identify and compare the subsets of FOXM1/ESR1 positive and FOXM1/ESR1 negative binding regions, breast carcinoma MCF-7 and MDA-MB-231 cell lines were investigated. Raw data from ChIP-seq was trimmed and mapped to UCSC human genome, hg19 using Bowtie2 then visualised on Integrative Genomic Viewer (IGV) (124-126). R package 'ChIPseeker' was then loaded on R Studio to annotate peaks to understand the genomic context of the bindings (109). The mapped data was then loaded on Python and sorted with 'bioconda' package with bedtools 'sort-bed' to sort peaks. In order to find the overlapped peaks (genomic bindings) of ESR1 and FOXM1, 'bedops' tool was used to intersect overlapped peaks (111, 112). FOXM1 peaks intersected with ESR1 were categorised as FOXM1/ESR1 positive, the rest labelled as FOXM1/ESR1 negative.

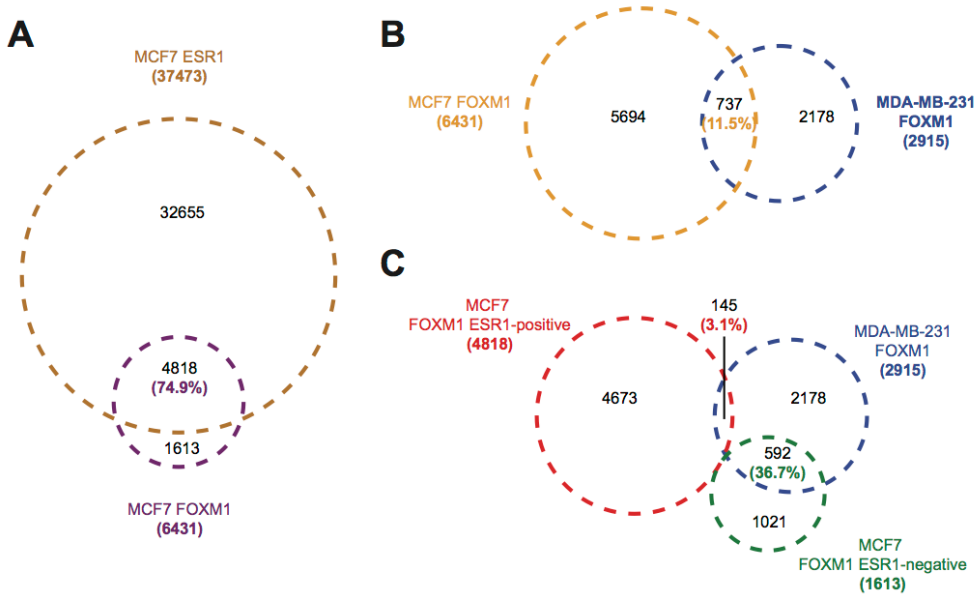


Figure 3.2 The overlap between ChIP-Seq peaks of FOXM1 and ESR1.

(A) Pairwise Venn diagram representing overlap of FOXM1 ChIP-seq peaks in MDA-MB-231 cell line and MCF-7 cell line. **(B)** Pairwise Venn diagram representing overlap between ChIP-seq peaks of FOXM1 and ESR1. **(C)** Venn diagram representing overlap between FOXM1/ESR1 peaks, FOXM1 only peaks in MCF-7 cell line and FOXM1 only peaks in MDA-MB-231 cell line.

As shown in Figure 3.2, ChIP-seq data of FOXM1 and ESR1 in MCF-7 cell line revealed a vast 74.9% overlap, indicating both the TFs bound on highly similar target gene regions in breast cancer MCF-7 cell line. I compared the similarity of FOXM1 transcriptional profile between ER-positive luminal A subtype MCF-7 and triple negative basal subtype MDA-MB-231 cell lines, ChIP-seq peaks showed a bare 11.5% overlap, displaying the high diversity of transcriptional binding across cell lines in different breast cancer molecular subtypes. FOXM1/ESR1 negative peaks in MCF-7 has a 36.7% overlap with that of MDA-MB-231, as opposed to 3.1% overlap FOXM1/ESR1 positive peaks between MCF-7 and MDA-MB-231 cell lines. The low percentage to almost none FOXM1/ESR1 positive regions in MDA-MB-231 cell line is not surprising as it is a known triple negative (ESR1, PR, and HER2 negative) breast cancer cell line. Furthermore, the low similarity (36.7%) of FOXM1/ESR1 negative transcriptional profile between MCF-7 and MDA-MB-231 cell line could be explained by the different cellular contexts and cell line binding specificity.

3.3.3 FOXM1/ESR1 co-binding regions are cell line specific, whereas FOXM1/ESR1 negative regions are conserved across various cancer cell lines

In order to characterise FOXM1/ESR1 positive and FOXM1/ESR1 negative cell lines, I proceeded to investigate these clusters across cancer cell lines to study its binding specificity. I hypothesised that FOXM1/ESR1 cell transcriptional profile are specific to ESR1 cell lines and endocrine regulated signalling pathways. ChIP-seq in five cancer cell lines were analysed: (i) MCF-7, (ii) MDA-MB-231, (iii) HeLa-S3, (iv) HepG2, (v) K562 (Figure 3.3). Each of these cell lines have distinct characteristics and respond to different treatments. MCF-7 is a breast carcinoma ESR1 positive cell line, also PR positive and HER2 negative, it is often the model for hormonal therapy research. Apart from MCF-7, all the other cell lines used are ESR1 negative. MDA-MB-231 is a basal, triple negative breast carcinoma cell line, indicated by the absence of ESR1, PR, and HER, it is frequently used for chemotherapy studies. HeLa-S3 is a human cervix carcinoma cell line. HepG2 is a human hepatocellular carcinoma cell line with epithelial-like morphology. Lastly, K562 is a chronic myelogenous leukemia (CML) cell line originated from human bone marrow.

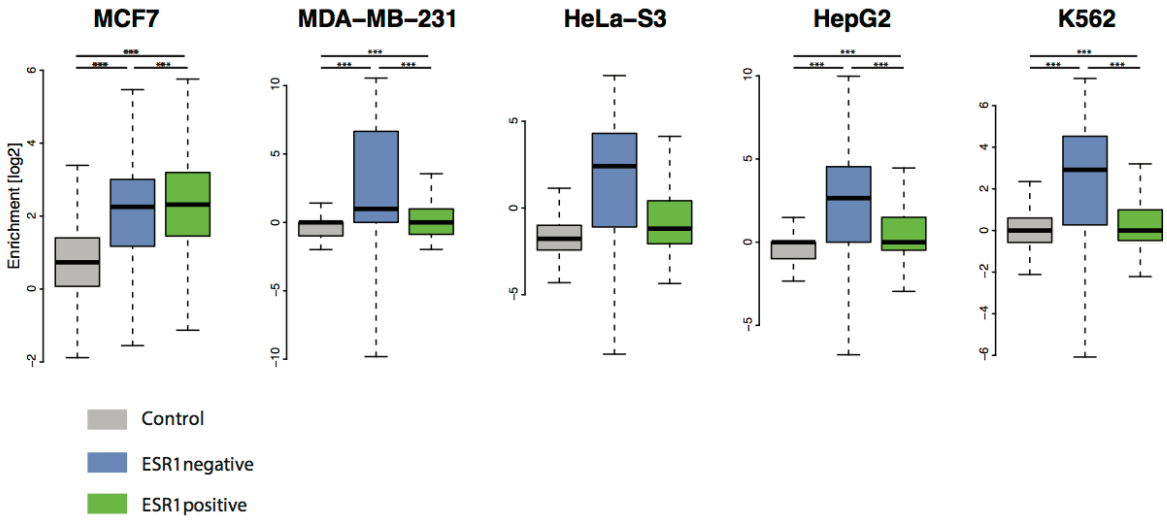


Figure 3.3 Conservation of activity of FOXM1/ESR1-negative regions in the absence of ESR1.

Enrichment of FOXM1/ESR1 common regions and FOXM1 only regions in MCF-7 cell line compared to other cancer cell lines, HeLa-S3, HepG2, K562, MDA-MB-231 using ChIP-Seq analysis. A negative control was included on the left side for each cancer cell line. Wilcoxon two-sided rank sum test was performed that was multiple test FDR corrected; *p<0.05, **p<0.01, ***p<0.001. Box and whiskers represent 5, 25, 50, 75 and 95% percentiles respectively.

In Figure 3.3, the results from these 2 clusters – ESR1 positive and ESR1 negative, were visualised with BOX plots. The coverage over 1,000 base pairs around the peak summit was calculated for ChIP-seq and input samples, added a pseudo count of 0.01 and computed the enrichment ratio of ChIP-seq over DNA input, which was transformed into log values.

In MCF-7, both FOXM1/ESR1 positive regions and FOXM1/ESR1 negative regions showed significant clusters compared to negative control. In the other cancer cell lines, FOXM1/ESR1 positive regions showed merely 3.1% conservation, whereas FOXM1/ESR1 negative binding peaks appeared to be conserved. It is known in literatures that ESR1 plays significant role in the growth of normal breast cells, as well as the development of breast cancer (127). Based on these observations, I hypothesised that the FOXM1/ESR1 positive binding cluster has a role in the development of ESR1 positive breast cancer and hormonal resistance. On the other hand, FOXM1/ESR1 negative bindings were conserved across cell lines, suggesting

these bindings may have preserved and fundamental roles in biological processes across cancer cell lines.

3.3.4 FOXM1/ESR1 positive and ESR1 negative binding regions display varied genome-wide distribution traits

Next, I looked at the genome-wide distributions of the ChIP-seq peaks (with FOXM1 pulldown) to understand the genomic characteristics of FOXM1/ESR1 positive and FOXM1/ESR1 negative regions. Distribution of ChIP-seq peak locations throughout the genome was identified using R/Bioconductor package “ChIPseeker” and ICSC Ensembl gene annotation files (109). FOXM1/ESR1 negative peaks were identified mostly in upstream and 5’ UTR regions (Figure 3.4). In other words, FOXM1/ESR1 negative peaks are in promoter regions of target genes. On the other hand, FOXM1/ESR1 positive peaks were detected mainly in the intergenic and intronic regions. Though FOXM1/ESR1 positive peaks also displayed bindings on the proximal core promoter regions, these findings support the notion that FOXM1/ESR1 positive peaks are mainly associated with enhancer regions that are looping different chromatin sites to regulate expression of distal genes whereas FOXM1/ESR1 negative genes are regulated proximally. Previous studies have shown that pioneer factors are dominated by the steroid hormone receptor family, for example FOXM1 and GATA3 are both involved in regulating the chromatin structure of ESR1 target regions to allow ESR1 to subsequently bind on these locations (128). FOXM1 may act as a pioneer factor in these FOXM1/ESR1 positive target regions to mediate ESR1 binding; additionally, these sites may be regulating mainly distal target gene regions.

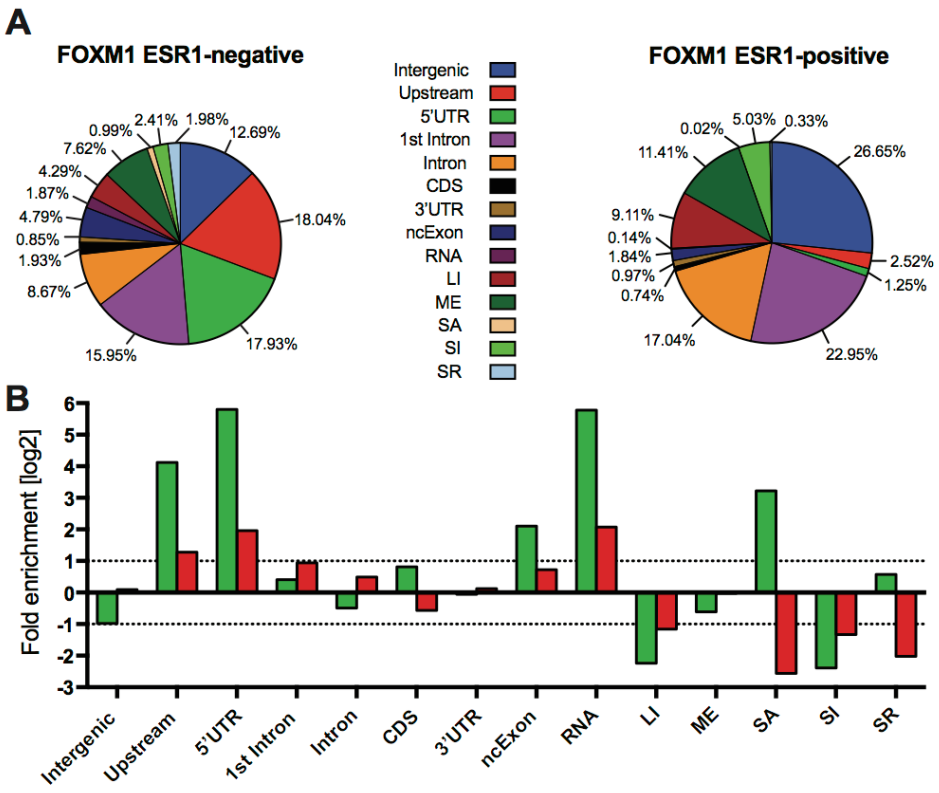


Figure 3.4 Genome-wide distribution of Chip-seq FOXM1 peaks.

(A) Genome-wide distribution of FOXM1/ESR1 peaks and FOXM1 only peaks in MCF-7. (B) Fold-enrichment of each class over the total genome composition. Fold-enrichment of the fraction of respective classes in the genome over the fraction of this class in the genomic composition using log2 scale. Intergenic = any regions not contained within the other classes. Upstream = regions within 2000 base pairs upstream of transcription start site (TSS). 5' UTR = 5' Untranslated Region. Intron = Every other intron but first intron. CDS = coding DNA sequence. 3' UTR = 3' Untranslated Region. ncExon = non-coding exon. RNA = DNA repeat classes, snRNA, srpRNA, rRNA, scRNA, RNA and tRNA. LI = Long interspersed nuclear element (LINE). ME = Mobile elements (long terminal repeat, DNA transposons). SA = satellite repeats. SI = short interspersed element (SINE). SR = simple repeats.

3.3.5 ChIP-seq profiles of epigenetic and enhancer markers in FOXM1/ESR1 positive and negative binding regions hint different genomic insights

To further demystify the pattern of these 2 clusters of FOXM1 transcriptional bindings, the transcriptional profiles of identified protein factors on these regions were investigated. ChIP-seq data of following protein factors in MCF-7 cell line were extracted: RAD21, STAG1, CTCF, EP300, and H3K27Ac (Figure 3.5).

The ChIP-seq peaks of all the stated factors were overlapped with FOXM1/ESR1 positive and ESR1 negative peaks, albeit to different extents. Similarly, MCF-7 TF

ChIP-seq profiles were pre-processed using the pipeline as stated in Figure 3.1: raw ChIP-seq profiles were converted into fastq format using Fastq-dump from the SRA toolkit. The adapters were trimmed, and poor-quality reads removed using Trimmomatic (129). The trimmed reads sorted using Samtools. Picard tools were used to remove multiple reads and Bedtools to remove blacklisted regions using ENCODE mappability regions. Peaks were called using Model-based Analysis of ChIP-seq 2 (MACS2) (130, 131). ChIP-seq peak files were overlapped with Bedtools intersect. Deeptools was used to perform binding analysis, utilising BigWig files generated from the BAM ChIP-Seq signal with the bamCoverage tool. Heatmaps and binding profiles were produced using FOXM1/ESR1 positive and negative 1 bp peak summit files, with FOXM1 summits shifted by 50 kb as a control. The summits were extended by 5,000 bp either side and binding profiles plotted. For the box plots, the coverage over 1,000 bp around the peak summit was calculated for ChIP-seq and input samples, added a pseudocount of 0.01 and computed the enrichment ratio of ChIP-seq over DNA input, which was transformed into log values.

The heatmaps and boxplots for the relevant factors displayed the intensities and significances of these overlaps, and each of the pattern guides us to better understand the two FOXM1 clusters. The weakened signal of *ESR1* summits in the FOXM1/ESR1 negative clusters acted as a control to ensure the cleanliness of the data before I proceeded to study the difference in patterns between the two clusters. These ChIP-seq data provided information of the epigenetic marks and enhancer properties-related bindings on the FOXM1/ESR1 positive and negative clusters.

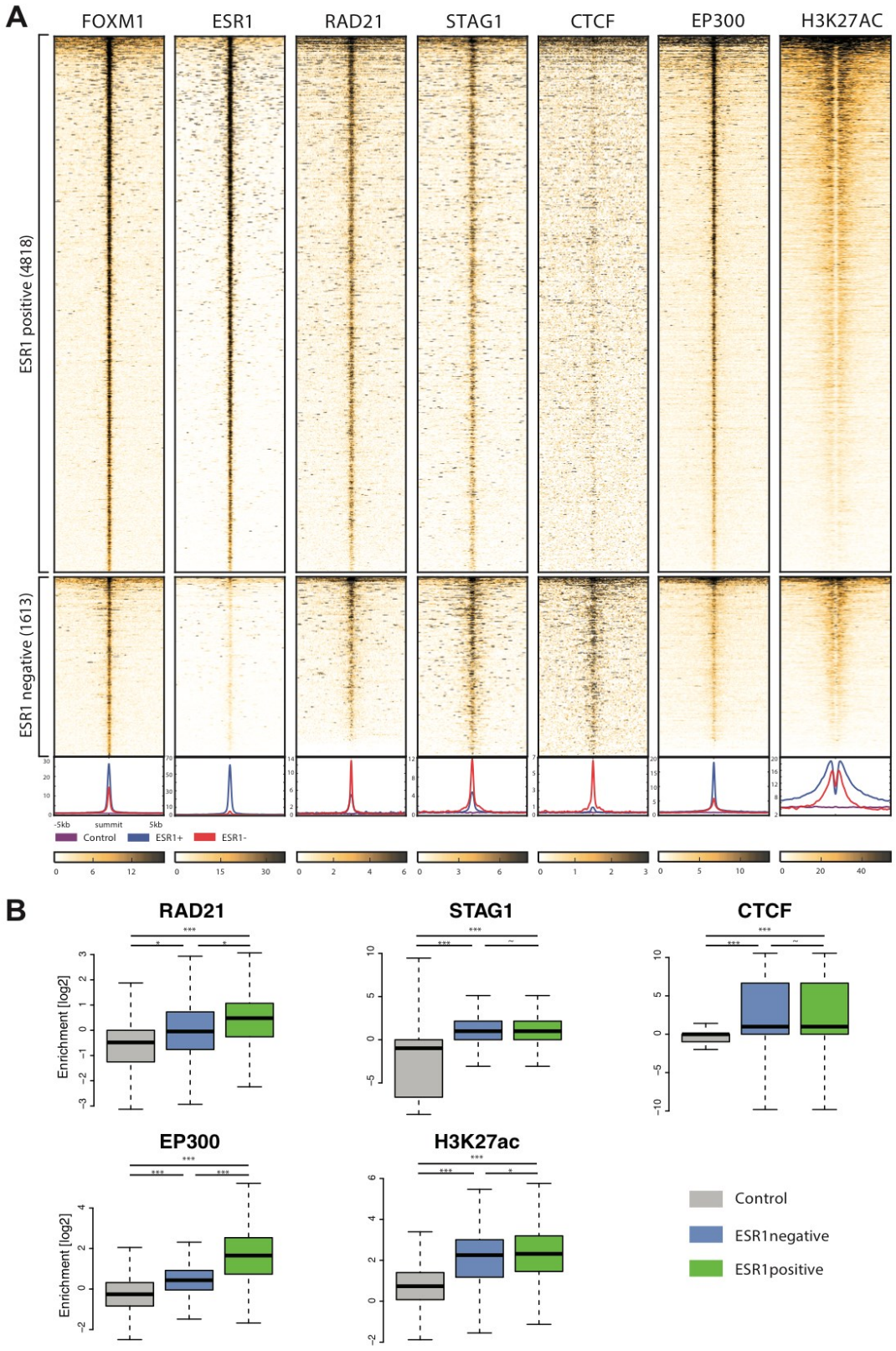


Figure 3.5 Chromatin landscape of FOXM1 bound regions.

(A) ChIP-seq signal of different chromatin markers across 5,000 base pairs both sides of FOXM1/ESR1 and FOXM1 summits in MCF-7. **(B)** Enrichment of respective chromatin mark from **(A)** over input DNA

control. Wilcoxon two-sided rank sum test was performed with multiple test FDR corrected; *P<0.05, **P<0.01, ***P<0.001. Box and whiskers represent 5, 25, median, 75 and 95% respectively.

Enhancer regions are known to amplify signals produced by TFs, thereby selectively activating the transcription of target genes (132). These enhancers are DNA regions that contain TF binding sites that trigger the transcription activities of specific promoter regions. Previous studies showed that two groups of enhancers exist in *Drosophila melanogaster* and ovarian somatic cells, which involved in the regulation of housekeeping and developmental genes (133). To further understand the activity of these FOXM1 transcription regulatory regions, I investigated a list of known enhancer markers on the identified FOXM1 target gene regions to investigate the enhancer-core promoter activity in MCF-7 cell line. Heatmaps showed ESR1 negative cluster to have strong enrichment of *RAD21*, *STAG1*, and CCCTC-binding factor (*CTCF)* compared to ESR1 positive cluster, which showed a higher *EP300* bindings (Figure 3.5). Moreover, other enhancer-associated chromatin features, such as *H3K27* acetylation (115), was observed in all FOXM1 peaks, regardless presence of ESR1. It is essential to understand the significance of these bindings. Firstly, looping is a suggested mechanism for enhancers to regulate and activate promoter regions. The CTCF is typically involved in chromatin looping (134). Additionally, STAG1 and RAD21 are factors that constitute part of the mitotic cohesion complexes, a complex that binds to enhancers and mediate looping independently of CTCF, though CTCF sites involved in chromatin looping are often co-occupied by cohesion complexes (135). In FOXM1/ESR1 positive regions, the occupancy of acetyltransferase coactivator, EP300, was detected (115). Differential enrichment of all these epigenetic markers suggests that the two FOXM1 clusters, regardless of the presence of ESR1, are present in active regulatory regions, with either enhancer or promoter related features.

3.3.6 ESR1 binding discriminates the biological pathways of FOXM1/ESR1 peaks and ESR1 negative peaks

Knowing that the two clusters are in active chromatin regulatory regions with varied enhancer-related properties (Figure 3.5), it would shed some light to find out co-factors that may be present in these regions to modulate the transcriptional state of these clusters. For that reason, I screened for consensus motif sequences present in these

FOXM1 binding regions for the predicted TFs that are likely to bind onto the same regions. Motif enrichment analysis of these FOXM1 clusters revealed various classes of TF binding (Figure 3.6). Motif Enrichment Analysis was conducted using AME from the MEME suite (136). Fasta files were generated using getfasta from Bedtools for ESR1 positive and negative peaks. The corresponding nucleotide sequences were analysed in AME to discover the significantly enriched motifs in each sequence utilising the JASPAR CORE 2016 vertebrate motif database. FOXM1 peaks were shifted by 50 kb as a control. Discriminative motif analysis was also performed in AME discovering the enriched motifs relative to the control sequence, where FOXM1 ESR1 negative sequence was used for FOXM1 ESR1 positive analysis, vice-versa.

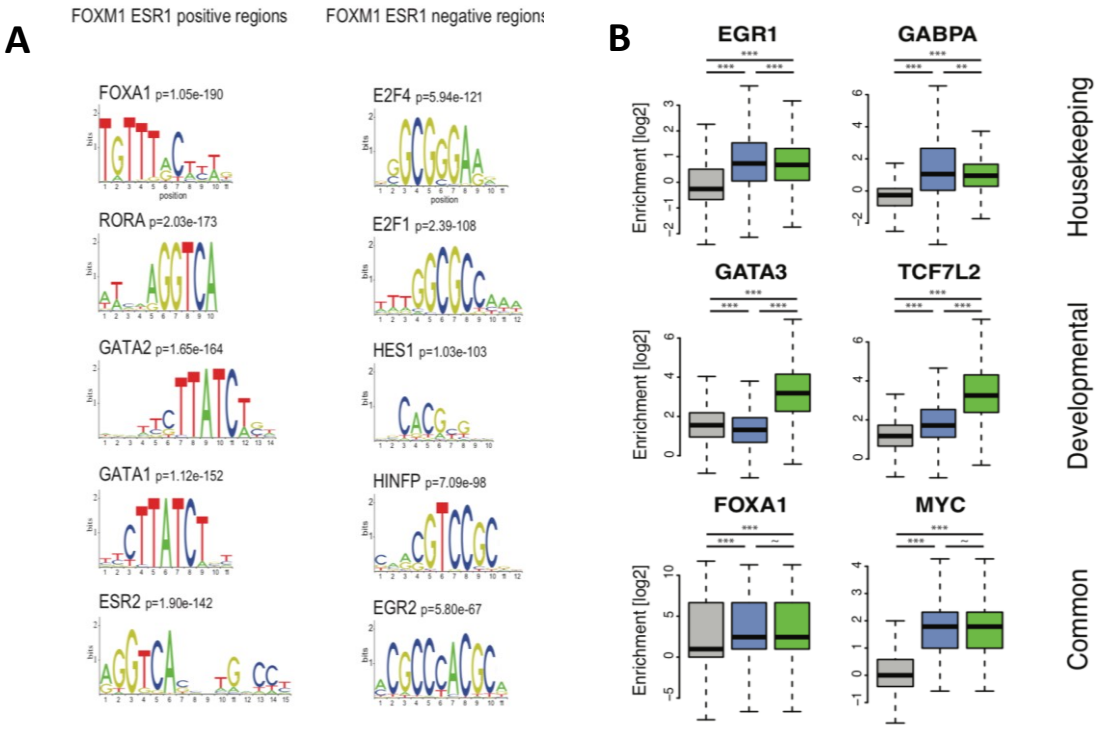


Figure 3.6 ESR1 binding discriminates between the biological pathways of FOXM1/ESR1 peaks and FOXM1 peaks.

(A) Discriminative motif enrichment analysis displaying significantly enriched motifs for the nucleotide sequences of both FOXM1/ESR1 and FOXM1 only peaks. FDR corrected p-values are presented, and symbols derived from JASPAR core 2016 vertebrate database **(B)** Enrichment of binding of selected TFs across FOXM1/ESR1 peaks and FOXM1 only peaks in MCF-7 cell line in comparison to input DNA controls where TF binding is. Increased in FOXM1 only regions (housekeeping), FOXM1/ESR1 regions (developmental) or the same (bottom).

Discriminative motif enrichment analysis is a powerful tool to identify enriched motifs in the identified DNA sequences in ChIP-seq profiles of different experimental conditions under study to truly understand the biological process involved (137). The annotated motifs were then matched with motif database JASPAR. In Figure 3.6, FOXA1 motif was discovered in FOXM1/ESR1 positive cluster, suggesting two likely scenarios for FOXA1, as a pioneer factor, to be involved in the transcriptional state of this cluster: (i) FOXA1 may bind and alter the chromatin state of these target regions to allow subsequent TFs, such as FOXM1 and ESR1, to bind and activate transcription of these regions (138), or (ii) FOXA1 and FOXM1 share highly similar motifs, therefore FOXM1 may bind to the recognised FOXA1 motif (48, 67). Similarly, ESR2 motif may be identical to ESR1 (139), therefore detected in FOXM1 ESR1 positive cluster in Motif Enrichment Analysis.

The other dimension to look at the outcome of the motifs screening would be to study the fundamental roles of the factors binding to these motifs. Discriminative motif enrichment analysis displayed different biological process enrichments for FOXM1/ESR1 positive and ESR1 negative regions. The FOXM1/ESR1 positive regions showed motifs for TFs associated with developmental pathways: FOXA1 in liver and subthalamic nucleus development and endocrine resistance (140, 141), GATA1 and GATA2 in blood development (142), and RORA in cerebellar development and bone formation (143). On the other hand, FOXM1/ESR1 negative regions displayed motifs for TFs associated with relatively fundamental biological processes: E2F1 and E2F4 in cell cycle and cell growth (144), EGR2 in cell growth and bone homeostasis (145). Furthermore, HES1 prevents differentiation of embryonic stem cells into myogenic or neural cells (146), and HINFP is involved in cell cycle regulatory mechanism (147). In a nutshell, this suggests a plethora of functional roles between FOXM1/ESR1 positive and negative regions that falls into a broad range of developmental or fundamental processes.

Further, downstream analysis can be used to validate predicted motifs. One methodology is to confirm the binding of TFs under study through individual ChIP experiments, another method is through the overlap of ChIP-seq peak summits to the enriched motifs. Accordingly, enrichment of various TF bindings revealed insights: In FOXM1/ESR1 positive cluster, GATA3 is associated with luminal cell differentiation (148) (Figure 3.6). Additionally, GATA3 regulates FOXA1 in modulating ESR1 binding.

TFs significantly enriched in FOXM1 ESR1 negative cluster showed different transcriptional roles; for example, GABPA binds to promoter of housekeeping genes (149). Commonly, FOXA1 was enriched in both classes, consistent with its role as a pioneer factor; whereas MAX-MYC heterodimer is involved in cell cycle regulation (150).

3.3.7 GO enrichment analyses show fundamental biological process differences between FOXM1/ESR1 positive and ESR1 negative binding regions

Furthermore, GO analysis was conducted to investigate the biological pathways involved given the predicted motifs binding (151). GO Analysis was performed on FOXM1/ESR1 positive and negative peaks, using all MCF-7 peaks and MDA-MB-231 peaks as two separate controls. R package “biomaRt”, the Ensembl archive release version 75 and R package “GenomicRanges”, were used to map ChIP-seq FOXM1 ESR1 positive and negative peaks to within gene body and 25 kb upstream of the transcription start site (TSS), generating two sets of gene list (152, 153). R package “GOstats” was used to perform hypergeometric tests to identify the top 100 significantly overrepresented GO terms for each list, using the GO database from the GO consortium. These lists were collapsed removing redundant terms leaving the top 20 over-represented GO terms. Hypergeometric tests were conducted for over- and under- representation for FOXM1 ESR1 positive and negative, MCF-7 and MDA-MB-231 peaks. For each of the top 20 terms across all 4 categories mentioned, subsequent False Discovery Rate (FDR) corrected p-values were transformed into log scale and used to evaluate whether they are overrepresented or underrepresented. Heatmap for the top GO terms categorised under either housekeeping or developmental roles was then plotted with R package “GGplots” or web-based tool Morpheus, developed by Broad Institute, with complete 20 terms, top 5 from each class (Figure 3.7).

The Gene Ontology Consortium defines GO as an international standard to annotate genes (154). Ontologies are standardised vocabularies to communicate and annotate knowledge for molecular functions, biological processes and cellular components. The GO database keeps track of known genes and their associated GO terms. R package “GOstats” finds GO terms that are statistically significant within an experimentally-

derived gene list based on a control set of genes. GOstat will result in a list of p-values that state how specific certain GO terms are for the selected gene list. Over-representation analysis is a widely used statistical method to assess whether genes identified to belong to a specific GO term (or known biological functions or signalling pathways) are present more than expected (over-represented) in the experimentally-derived gene list, i.e. the 4 categories of interest (154). The more overrepresented the gene list, more genes in the set overlaps with the identified genes associated with specific GO term. The p-values of underrepresented GO terms are displayed without colours in the visualised heatmap, whereas the overrepresented terms lean towards dark maroon (Figure 3.7).

The top significantly overrepresented GO terms showed variable enrichments. FOXM1/ESR1 positive peaks showed mainly pathways essential for cell survival such as metabolism and chromatin organisation, whereas FOXM1/ESR1 negative peaks showed genes crucial for sustaining cell type-specific processes, such as tissue development and regulation of signalling. Interestingly, there is no overlap in GO terms between FOXM1/ESR1 positive and negative peaks. To support the fundamental housekeeping roles regulated by FOXM1/ESR1 negative peaks, the peaks seem to be conserved in MCF-7 and MDA-MB-231 cell lines. This finding aligns with the conservation of FOXM1/ESR1 negative binding profile across various cell lines (Figure 3.3). Additionally, these findings are in line with the enhancer-related footprints, which demonstrated the differences in gene regulation and the downstream biological processes between FOXM1/ESR1 positive and negative clusters (Figure 3.5, Figure 3.6). Indeed, study had shown that enhancers core promoter specificity separates housekeeping and developmental gene regulation (133).

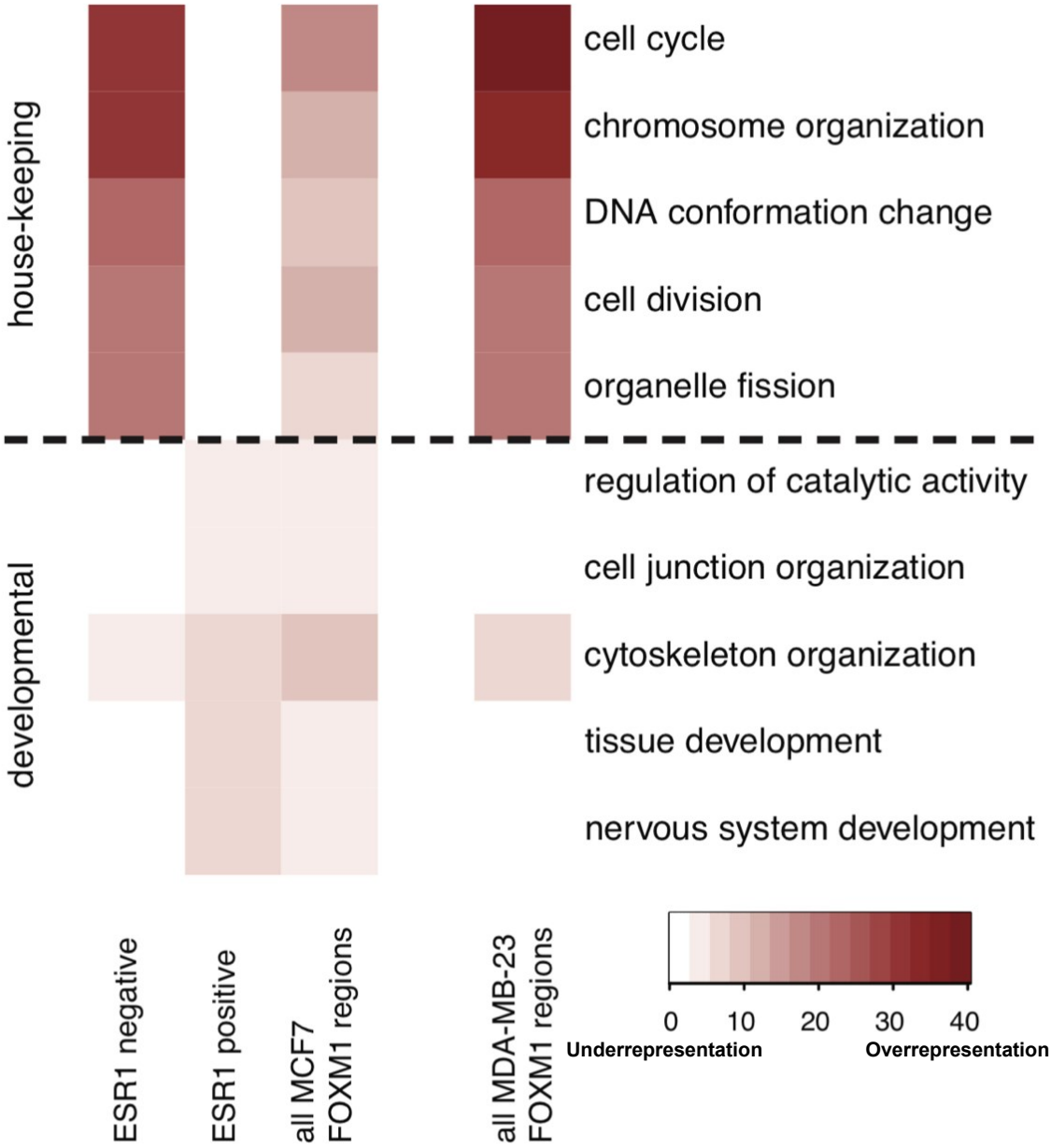


Figure 3.7 GO enrichment analysis shows housekeeping roles for FOXM1 specific target regions whereas FOXM1/ESR1 regions displays developmental roles.

GO enrichment analysis illustrating top 5 overrepresented non-redundant GO terms derived from GO consortium database, FDR multiple test corrected p-values for FOXM1/ESR1 peaks and FOXM1 only peaks in MCF-7 cell line and all MDA-MB-231 FOXM1 peaks. Log10 (underrepresentation) indicates no enrichment of target genes identified in the respective pathways, whereas Log10(overrepresentation) represents enrichment of target genes associated in respective pathways.

3.3.8 FOXM1 knockdown impacts chromatin accessibility in MCF-7 cell line

I proceeded to study the relationship between FOXM1 and ESR1 on the regulation of these two FOXM1 clusters. Transient knockdown assay is useful to study the conditional impact of the absence of a molecule on cell lines. ATAC-seq is a technique to measure the genome-wide chromatin accessibility of cells (Figure 3.1). Downregulation of *FOXM1* was achieved with short-interfering (si)RNA targeting the mammalian *FOXM1* DNA sequence. *FOXM1* knockdown impacted chromatin accessibility in MCF-7 cell line (Figure 3.8). Specifically, FOXM1/ESR1 negative regions was affected more than that of FOXM1/ESR1 positive regions upon *FOXM1* depletion, with a median difference of 2.17 and 0.85 respectively.

Notably, when *FOXM1* was transiently knocked down, ATAC-seq signal in all ESR1 bound regions was significantly downregulated, indicating that chromatin accessibility decreased in these regions. This could imply that the transcriptional activity in these regions is impacted in the absence of *FOXM1* and suggests that FOXM1 is essential in altering chromatin structure and regulating *ESR1* transcriptional activity upstream.

Furthermore, ChIP targetting FOXM1 (ChIP: FOXM1) was performed in fulvestrant-treated MCF-7 cell line (Figure 3.8B). Fulvestrant inhibits ESR1 downstream mechanism by binding to ESR1 and leading to its degradation (Figure 1.3). Five promoter regions were selected each for FOXM1/ESR1 positive (9 – 13) and negative (1 – 5) regions, ChIP primers designated to tag each of these regions were utilised to amplify these regions during ChIP-qPCR if they were pulled down by FOXM1 antibody, which indicates that FOXM1 is binding to these regions. Primer pair binding to no known mammalian promoter regions were included as negative control. The primer sequences are detailed in **Section 2.4.2 RT-qPCR Primer Design**. ChIP: FOXM1 showed positive fold enrichment over negative control in both FOXM1/ESR1 positive and negative cluster. Notably, increase in FOXM1 enrichment was observed in FOXM1/ESR1 positive cluster when MCF-7 cell line was treated with fulvestrant, suggesting a compensatory role of FOXM1 potentially happening to replace ESR1 when it was degraded, on their co-regulated promoter regions. Alternatively, FOXM1 could be binding more on the target regions to increase transcriptional activity as a rescue mechanism for MCF-7 to resist fulvestrant treatment.

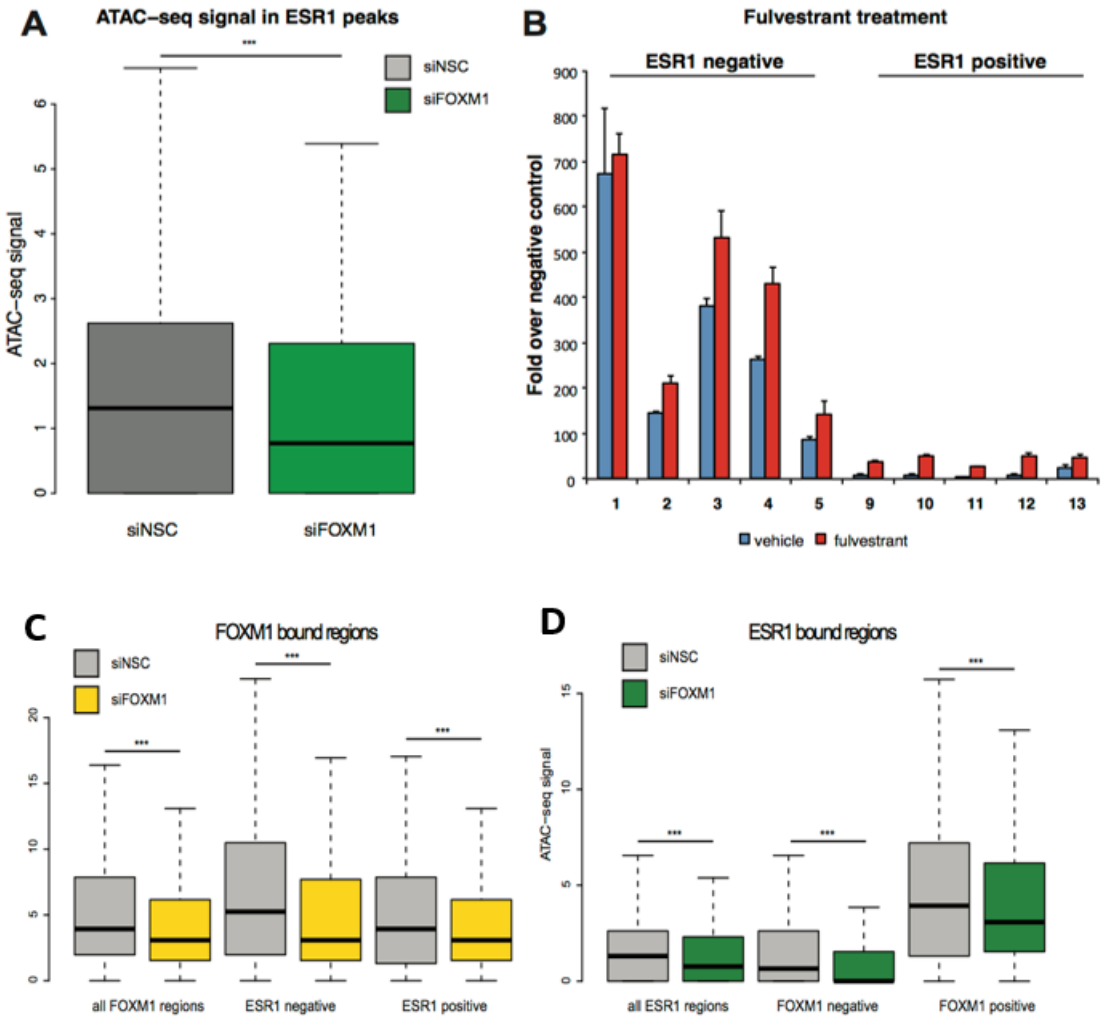


Figure 3.8 The chromatin accessibility impact on FOXM1 and ESR1 bound regions upon FOXM1 silencing in MCF-7 cell line.

(**A**) ATAC-seq signal detection and quantification in ESR1 peaks upon FOXM1 silencing and in the control cells; ***p<0.001 statistical difference; box and whiskers represent 5, 25, median, 75 and 95% quartiles respectively. **(B)** ChIP-qPCR enrichment of FOXM1 binding in FOXM1/ESR1 regions and FOXM1 only regions upon fulvestrant treatment. Results shown are representative of three replicates and standard deviation respectively. **(C)** Normalised ATAC-seq signal in MCF-7 cell line in FOXM1 regions upon FOXM1 silencing. **(D)** Normalised ATAC-seq signal in MCF-7 cell line in ESR1 regions upon FOXM1 knockdown.

3.3.9 Recruitment of FOXM1 to estrogen-responsive targets are dependent on ESR1 binding

As mentioned, ChIP is a technique to study the binding of protein molecules onto target DNA regions. ChIP was conducted to study the binding of TFs, FOXM1, ESR1, and FOXA1, on the promoter regions of FOXM1 ESR1 positive and negative regions (Figure 3.9). *CCND1* and *ACTB* promoter regions, known as non-targets of FOXM1 or ESR1, are included as negative controls. A heatmap analysis was conducted to compare the relative binding intensity of FOXM1, ESR1, and FOXA1 in selected FOXM1 ESR1 negative and ESR1 positive regions (Figure 3.9). All 3 TFs did not display positive signal on *CCND1* and *ACTB* promoter regions, indicating no transcriptional binding in these cell lines at the harvest timepoint and treatment conditions. Firstly, FOXM1 binds to all regions of FOXM1 ESR1 negative regions irrespective of the cell lines (MCF-7 or FulR), or treatment (untreated or fulvestrant treated). On FOXM1 ESR1 positive regions, ESR1 displays stronger binding relative to FOXM1 in MCF-7 cell line. In fulvestrant resistant cells, both FOXM1 and ESR1 did not show binding. Finally, FOXA1 showed binding in both FOXM1 ESR1 positive and negative regions, again supporting its role as a pioneer factor in most promoter regions with active transcriptional activity.

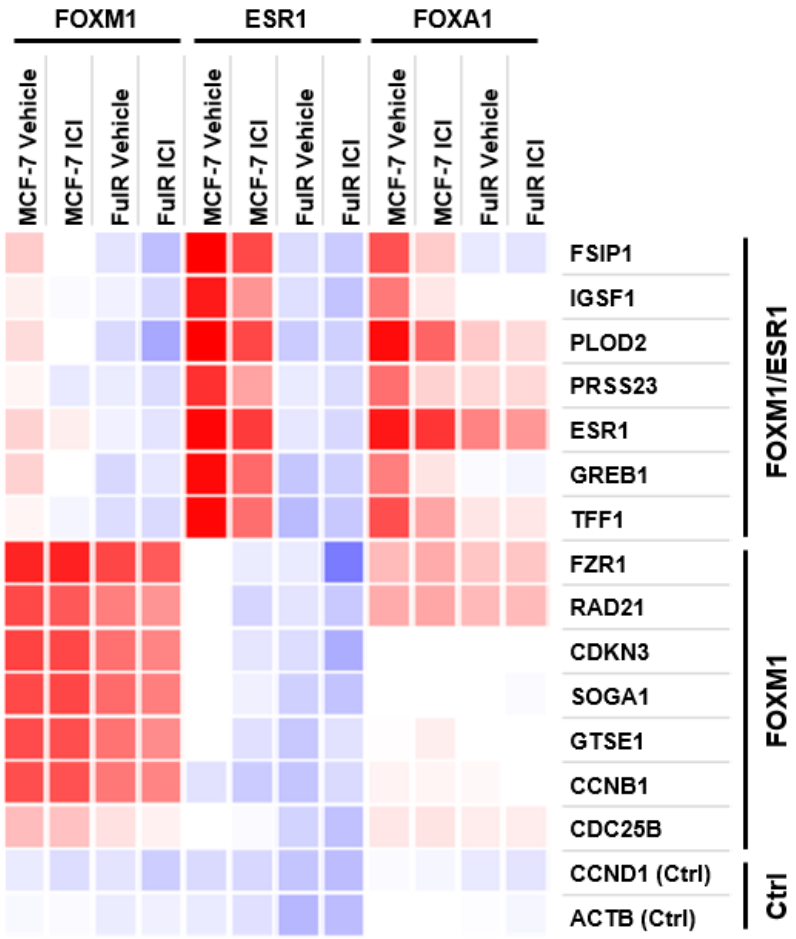


Figure 3.9 FOXM1 binding to the regions that are absent of ESR1 is not affected.

Heatmap analysis was conducted to show the effect of ICI on FOXM1/ESR1 co-bound promoter regions and FOXM1 specific promoter regions in MCF-7 and MCF-7 FulR cell lines. **(B)** ChIP coupled with RT-qPCR analyses showed that ICI did not impact on FOXM1 and FOXA1 binding on FOXM1 specific target promoter regions. Red is a positive signal, indicating a binding; blue is a negative signal, indicating a non-binding.

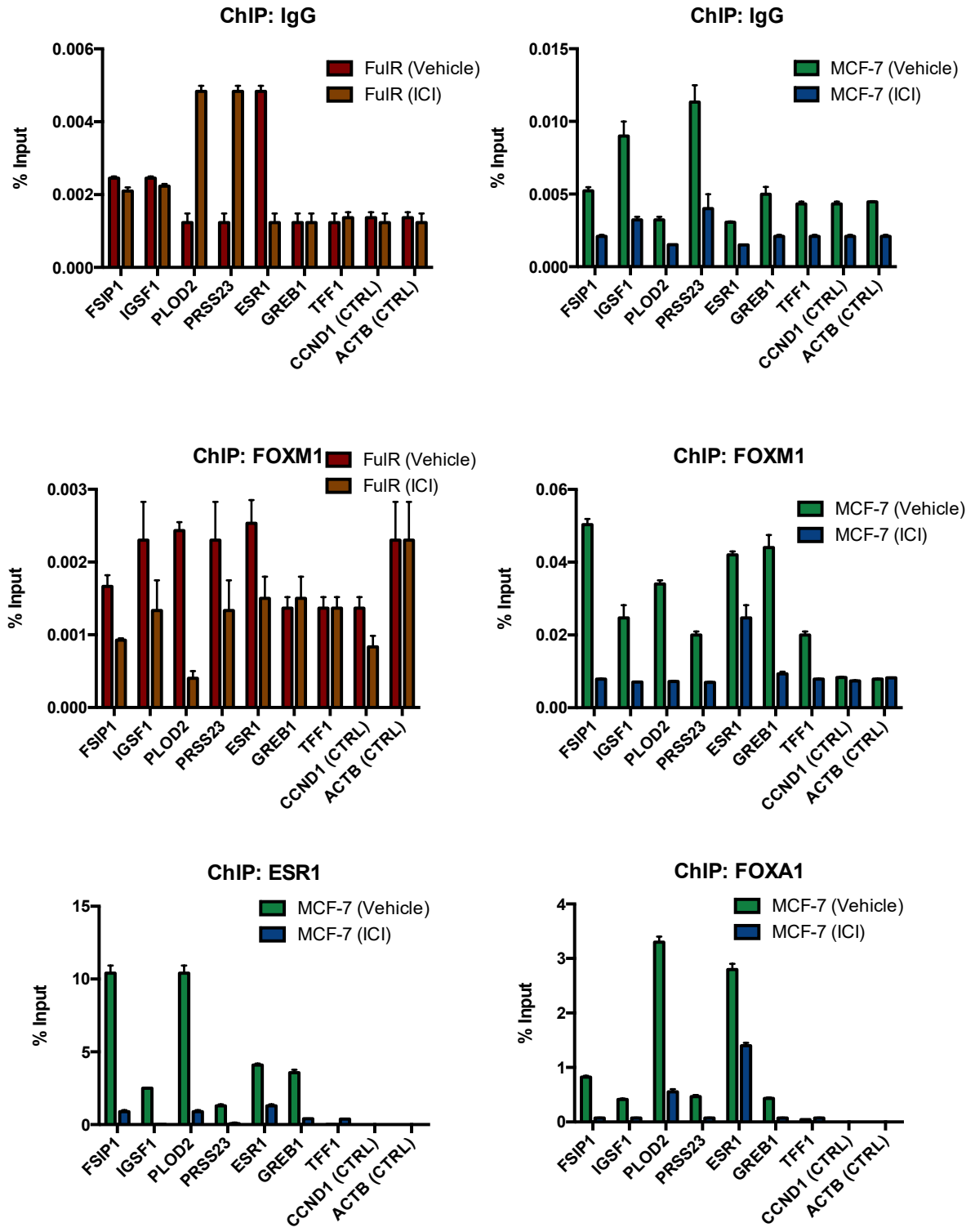


Figure 3.10 Recruitment of FOXM1 to estrogen-responsive targets are dependent on ESR1 binding.

(A) Wild type MCF-7 cell line was treated with 100 nM ICI 182,780 (also known as ICI or fulvestrant) for 3 h, 6 h, 12 h, 24 h or left untreated. Western blotting was performed using cell lysates extracted after drug treatment. **(B)** ChIP and RT-qPCR analysis were performed to assess binding of FOXM1, ESR1, or FOXA1 protein on FOXM1/ESR1 co-bound regions. Results and error bars are representative

of three technical replicates of ChIP experiment. As the data shown represent mean ± SEM of three technical repeats, no statistical analysis was performed.

3.3.10 Thiostrepton has greater effects in inhibiting ESR1 and ESR1-target gene expression compared to FDI-6

To further study the relationship between FOXM1 and ESR1, I investigated the inhibitory effect of FOXM1 inhibitors, thiostrepton and FDI-6, on the expression level of FOXM1, ESR1, and the identified target molecules (Figure 3.11). When MCF-7 cell line were treated with thiostrepton, both *FOXM1* and *ESR1* mRNA levels were significantly downregulated compared to untreated control; this effect was similarly translated into the mRNA level of FOXM1/ESR1 target molecules, *GREB1* and *TFF1*. Expectably, *ESR1* mRNA was significantly downregulated by tamoxifen (4-OHT) and fulvestrant (ICI). *FOXM1* mRNA was significantly downregulated by FDI-6, albeit at a lower level than thiostrepton. This led to the hypothesis that thiostrepton has greater inhibiting effect on *FOXM1* than FDI-6. Additionally, thiostrepton may suppress *ESR1* mRNA.

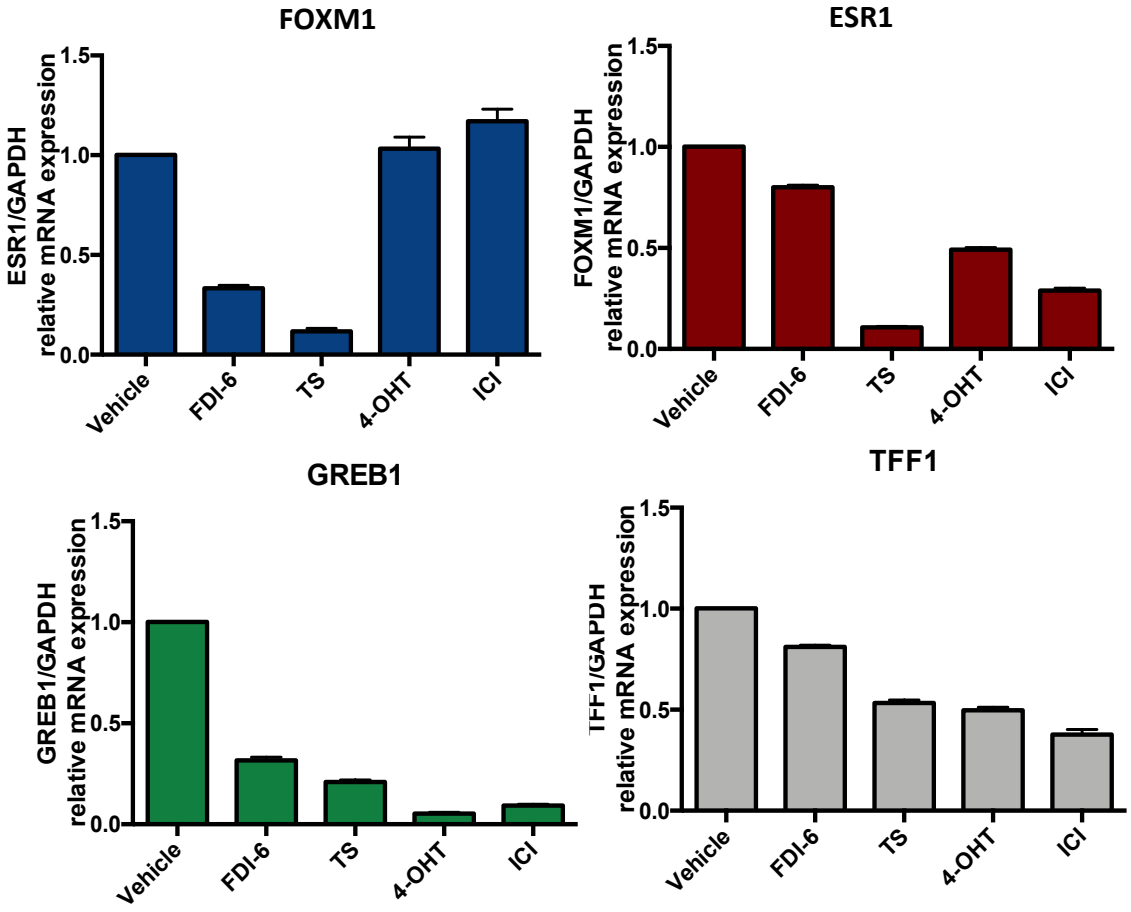
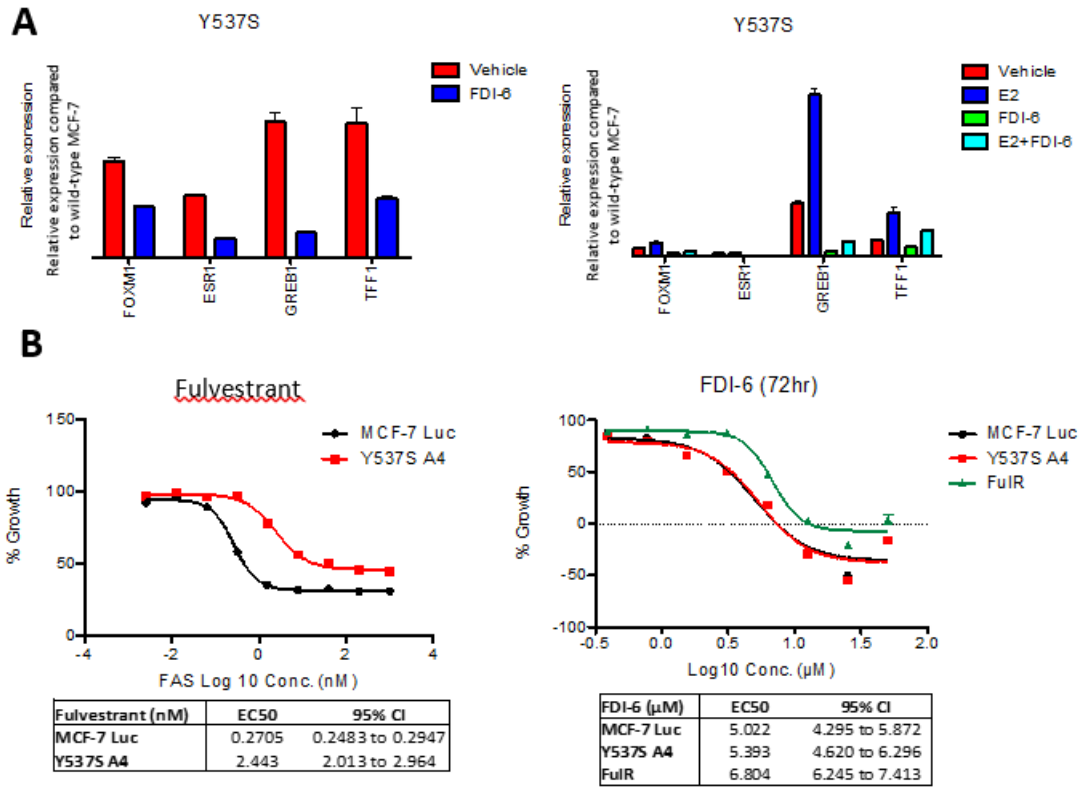


Figure 3.11 Thiostrepton has greater effects in inhibiting ESR1 and ESR1-target gene expression compared to FDI-6.

Gene expressions of FOXM1, ESR1 and estrogen-responsive target genes, GREB1 and TFF1 were measured by RT-qPCR analysis using RNA prepared 16 h following addition of ligands, FDI-6 (20 μM), thiostrepton (TS) (10 μM), 4OHT (10 nM), or ICI (100 nM), relative to the vehicle-treated MCF-7 cell line (*n* = 3 technical replicates). As the data shown represent mean ± SEM of three technical repeats, no statistical analysis was performed. TS, thiostrepton, 4OHT, tamoxifen. ICI, fulvestrant.

3.3.11 FDI-6 inhibits the expressions of ESR1 and ESR1-target genes in ESR1 mutant MCF-7 cell line

Furthermore, I observed the effect of FOXM1 inhibition by FDI-6 in wild type MCF-7, fulvestrant resistant MCF-7, and ER-mutant MCF-7-Y537S cells.I investigated cell growth rate in wild-type MCF-7, mutant MCF-7-Y537S with constitutively active ESR1, and fulvestrant resistant MCF-7, FulR. The mutant MCF-7 cell line was created with CRISPR-Cas9 technique.



Fulvestrant (nM)	EC50	95% CI
MCF-7 Luc	0.2705	0.2483 to 0.2947
Y537S A4	2.443	2.013 to 2.964

FDI-6 (μM)	EC50	95% CI
MCF-7 Luc	5.022	4.295 to 5.872
Y537S A4	5.393	4.620 to 6.296
FulR	6.804	6.245 to 7.413

Figure 3.12 FDI-6 inhibits the expressions of ESR1 and ESR1-target genes in ESR1 mutant MCF-7 cell line.

(A) Gene expressions of estrogen-responsive genes were measured in mutant Y537S MCF-7 cell line. For conditional treatment, estrogen, E2 was 10 nM; FDI-6 was 20 μM. **(B)** Cells were treated with serial concentration of fulvestrant for 48 h. Mean growth is shown relative to that for untreated cells. EC50 = the dose of drug causing 50% of maximum effect for any measured biological effect of interest, which in this case represents the point which cell growth is inhibited by 50%. As the data shown represent mean ± SEM of three technical repeats, no statistical analysis was performed.

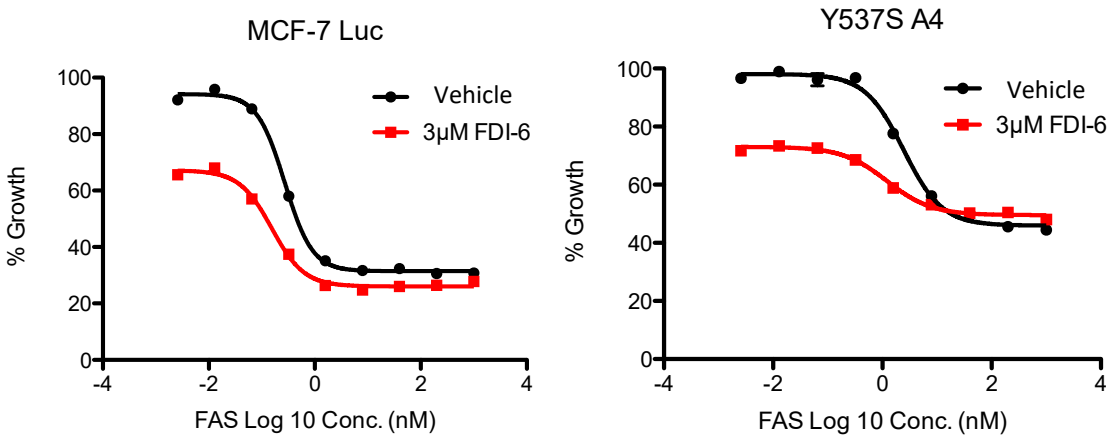
A study by Harrod et al. has showed that fulvestrant was a more effective inhibitor of MCF-7-Y537S growth compared to tamoxifen, in the absence of estrogen(104). Additionally, ESR1 target gene expression is ligand-independent and is often upregulated in MCF-7-Y537S cells. In this chapter, the mRNA expression level of *FOXM1*, *ESR1*, *GREB1*, and *TFF1* were detected by RT-qPCR in wild-type MCF-7 and MCF-7-Y537S mutant (Figure 3.12). The relative mRNA expression of target genes in MCF-7-Y537S were normalised against wild-type MCF-7 cell line. As observed, *FOXM1* level remained the same in untreated MCF-7-Y537S compared to MCF-7 cell line, whereas *ESR1* mRNA expression in MCF-7-Y537S is half of MCF-7. Interestingly, target genes *GREB1* and *TFF1* have approximately 1.5x higher mRNA

level than MCF-7. All *FOXM1*, *ESR1*, *GREB1* and *TFF1* mRNAs were impacted more in MCF-7-Y537S than in MCF-7 when cells were treated with *FOXM1* inhibitor, FDI-6.

Furthermore, FDI-6 was more effective in the inhibition of *FOXM1* in MCF-7-Y537S than MCF-7 cell line, evidently an additional 50% decrease in *FOXM1* mRNA was seen in mutant MCF-7-Y537S cells. In MCF-7-Y537S cells, the additional reduction of *FOXM1* mRNA resulted in the drastic decrease of the mRNA of *GREB1*, *TFF1*, and the already lower *ESR1*.

3.3.12 Combinatorial treatment of FDI-6 and fulvestrant on MCF-7 and MCF7-Y537S cell lines show promising results

Next, I studied the effect of fulvestrant or FDI-6 on the cell growth in different variations of MCF-7 cell lines (Figure 3.13). MCF-7-Y537S cell lines were roughly 10-fold more resistant to fulvestrant than MCF-7 cell line, with EC50 achieved at 2.443 nM fulvestrant and 0.2705 nM fulvestrant respectively. EC50 indicates the dose of drug causing 50% of maximum effect for any measured biological effect of interest, which in this case represents the point which cell growth is inhibited by 50%. MCF-7 cell lines reacted more positively to FDI-6, as the EC50 for FDI-6 in wild-type MCF-7, MCF-7-Y537S, and MCF-7-FulR cells were similar at 5.022 μM, 5.393 μM, and 6.804 μM respectively. An optimum concentration of 20 μM FDI-6 has always been the standard guideline to treat all MCF-7 cell lines for all conditional experiments. This suggests that the concentration of FDI-6 required to kill 50% of native or resistant MCF-7 cell lines will unlikely be exceeding the toxicity level for these cell lines, and even for patients' use. Moreover, FDI-6 presents itself as the potential alternative to treat endocrine resistant MCF-7 cell lines that no longer respond to ESR1-targeted drugs, such as fulvestrant.



	ICI (nM)		ICI (nM) + 3μM FDI-6	
	EC50	95% CI	EC50	95% CI
MCF-7 Luc	0.2705	0.2483 to 0.2947	0.1542	0.1325 to 0.1795
Y537S A4	2.443	2.013 to 2.964	1.211	0.9149 to 1.604

Figure 3.13 Effects of FDI-6 on growth in ESR1 mutant cells in combination with fulvestrant.

Wild type MCF-7 cell line or mutant Y537S MCF-7 cell line were treated with increasing concentration of fulvestrant or combined fulvestrant and FDI-6 (3 μM) treatment for 48 h. Mean growth is shown relative to that for untreated cells. EC50 = the dose of drug causing 50% of maximum effect for any measured biological effect of interest, which in this case represents the point which cell growth is inhibited by 50%.

FDI-6 seemed to be a potential alternative to endocrine treatment when MCF-7 cell line has gained resistance (Figure 3.12). Combinatorial drug regimens are frequently used in breast cancer treatments when cells are either no longer responding to monotherapies, or when relapse has occured. The main challenge in finding the optimum drug treatment lies in identifying a balanced treatment dose weighing between the effectiveness and toxicity. In Figure 3.13, I looked into the possibility of a combinatorial treatment with fulvestrant and FDI-6. Wild-type MCF-7 and MCF-7-Y537S cell lines were treated with constant 3 μM of FDI-6 and increasing concentration of fulvestrant. The cell growth of wild-type MCF-7 were reduced 50% when treated with 0.1542 nM of fulvestrant and 3 μM of FDI-6, whereas in MCF-7-Y537S 1.211 nM fulvestrant and 3 μM of FDI-6 were required for 50% cell reduction. The concentration of fulvestrant and FDI-6 in combinatorial treatment were both lower than that used in monotherapies. This experiment acted as a prelimary outcome to

investigate the effect of this combinatorial treatment in MCF-7 with fulvestrant resistance, potentially providing an alternative treatment option.

3.4 Discussion

Transcriptional networks are tightly regulated by positive (co-activators) and negative (co-repressors) regulators. Presence of co-regulators on genomic regions are crucial for the recruitment of co-factors to ESR1 transcriptional complex, such as CBP/P300, SRC-2, and GRIP1 (67), or these co-regulators can be a co-factor of *ESR1* transcription itself. FOXM1 is known to modulate its transcriptome by the recruitment of co-activators (69, 155). Unsurprisingly, FOXM1 cistrome and interactome have shown to be better predictors of breast cancer metastatic outcome than *FOXM1* expression itself (155). In this chapter, I studied the relationship between FOXM1 and ESR1, and how it affects downstream target genes and therapy responses, in breast cancer models MCF-7 and MDA-MB-231 cell lines.

To study the DNA profiles in these cancer cell lines, ChIP-seq and ATAC-seq were two techniques adopted. ATAC-seq studies the chromatin accessibility of DNA region based on the accessibility of transposon, whereas ChIP-seq studies the binding profile of a specific TF on the open chromatin region. These two methods are usually complementary. Through ChIP-seq, I showed two subsets of FOXM1 transcriptional binding profile, differentiated by the presence of ESR1 binding regions (Figure 3.2). A zoom in view of these target genes showed that FOXM1 and ESR1 transcription profiles shared high similarities, whereas FOXM1 profiles differ widely between breast cancer MCF-7 and MDA-MB-231 cell lines. This stark difference differentiated two subsets of FOXM1 transcriptome and is likely dictated by presence of ESR1.

FOXM1/ESR1-positive target genes are specific to ESR1-positive luminal A subtype MCF-7 cell line, whereas FOXM1/ESR1-negative target genes are conserved across cancer cell lines (Figure 3.3). In MCF-7 cell line, FOXM1/ESR1 positive target sites represent another layer to the intricate transcriptional network of FOXM1 and ESR1 in breast cancer. FOXM1 may show cell-line dependent patterns of DNA binding due to recruitment by different transcriptional complexes responsible for driving proliferation depending on cellular context. These observations led to my hypothesis that there are two broad FOXM1 transcriptional programmes in MCF-7 cell line.

Over a million of enhancers regulate cell- and tissue- specific expression of human genes, including circadian genes, core clock functions and ultimately behaviour (156). Deregulation of enhancer sequence and activity leads to cancer development (156). Enhancers can function from a distance from promoter region irrespective of the

orientation relative to target genes. As core promoters usually support low-level basal transcription, enhancers finetune and amplify the signals produced by TFs on cell context dependent basis, leading to various scenarios of gene expression. This also explains why the expression level of a TF is not fully representative of its activity. Genome-wide distribution analysis of FOXM1/ESR1-positive binding regions indicated that most of the bindings were found in intronic and intergenic regions, suggesting a comparable pattern to nuclear hormone receptors (127) (Figure 3.4). This characteristic illustrates a typical enhancer-proximal core promoter regulation on target genes. FOXM1/ESR1-negative bindings, however, are mostly located in 5' UTR and upstream regions, representing a distal promoter regulation.

In 2015, Zabidi et al. has supported the genome-wide presence of enhancer-core promoter specificity of thousands of enhancers to one of two types of core promoters, this differential regulation separates housekeeping and developmental genes in *Drosophila* and mammalian ovarian somatic cells (133). According to self-transcribing active regulatory region sequencing (STARR-seq) in embryos S2 cells (S2), larval brain (BG3) and ovarian somatic cells (OSCs) in *Drosophila melanogangster*, there are two broad classes of housekeeping and developmental (or regulated) promoters. The housekeeping genes are ubiquitously expressed, share enhancers that overlapped or proximal with a transcription start site and conserved across all tested cell types. GO analysis showed that these genes are involved in metabolism, RNA processing and cell cycle. On the other hand, the developmental genes specific enhancers differed strikingly between cell types and from housekeeping enhancers. They regulate cell type-specific and developmental genes, also located mostly in introns and intergenic regions, may feature more complex regulatory machinery, and activate target genes from distal core promoter regions of various distances. Any overlapped between housekeeping and developmental enhancers indicate common core promoters that can be activated to different extents depending on cell contexts (such as housekeeping genes that are further activated in specific cell types). These enhance-core promoter specificities are hypothesised to be modulated by defined biochemical compatibilities between different *trans*-acting factors of enhancers and different paralogues for components of the general transcription apparatus at core promoters. Such paralogues may have promoter-selective, cell type- and developmental stage- specific functions, therefore these enhancer-core promoter

specificities could be adopted widely to characterise and separate different transcriptional programmes.

Features that are related to enhancer activity, such as co-factors bindings (e.g. CBP/EP300), regulatory factors, or histone modifications can predict the location of enhancers (157). Apart from factor bindings on enhancer regions, notion supports that transcriptional control are regulated through physical contract between enhancers and promoters (132). To understand the enhancer-core promoter specificity in FOXM1 regions in mammalian MCF-7 cell line, enhancers related markers are probed in the identified two subsets of FOXM1 transcriptional regions (Figure 3.5). Based on the ChIP-seq signals of these known enhancer- and promoter- regulatory factors, the potential mechanism in these FOXM1 regions in MCF-7 cell line are proposed (Figure 3.14).

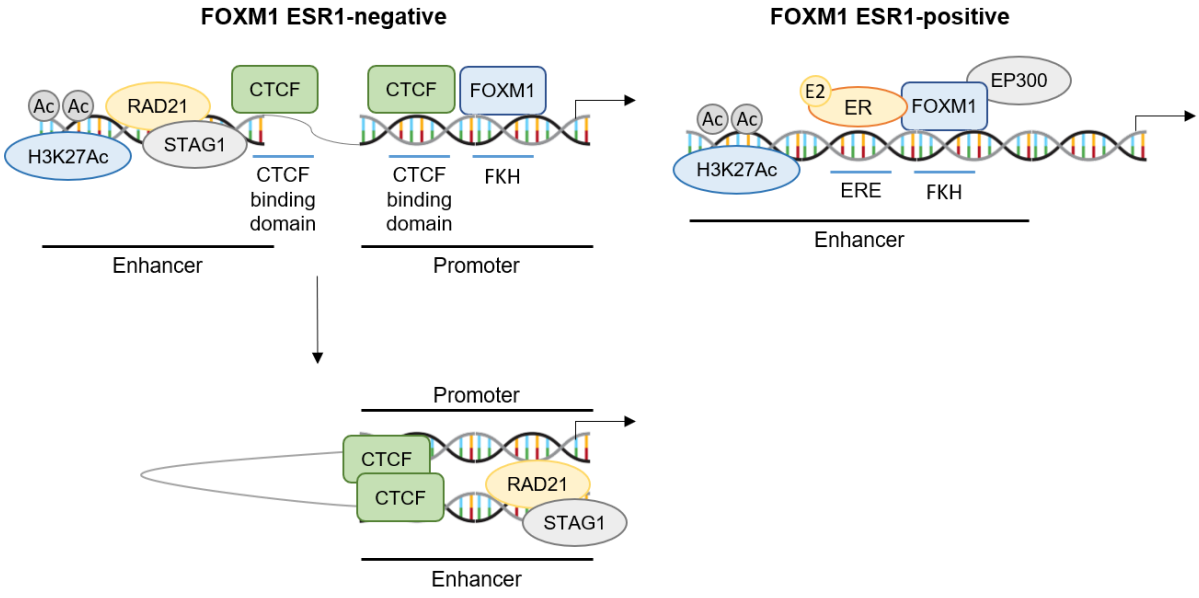


Figure 3.14 Proposed simplified models for enhancer- and promoter- regulatory properties for (left) FOXM1/ESR1 negative and (right) ESR1 positive target regions in MCF-7 cell line.
Enhancer-core promoter interactions and the binding of TFs, co-factors, and histone modifications regulate cell-type and developmental stage-specific gene expression.

Active enhancers have been shown to correlate with acetylation of lysine 27 of the adjacent histone H3 (H3K27Ac) and methylation of lysine 4 (H3K4me1). CBP/EP300 complex acetylate H3K27 in enhancer regions to drive promoter activation and gene

transcription. In addition, they can be bound and transcribed by RNA polymerase II (RNAPII), which produces enhancer RNAs and are often hypomethylated in mammalian cells. CTCF dimerization is stated to stabilise the formation of chromatin loops and recruitment of cohesion complexes formed by RAD21 and STAG1 (132). Going forward, the interaction between these regulatory factors and DNA regions should be further explored with co-immunoprecipitation and ChIP in specific gene regions. The promoter gene activity, frequency, and intensity regulated by specific enhancers can be studied by reporter assays using various reporter genes and can be either visualized (by LacZ staining, fluorescence or *in situ* hybridization) or quantified using bioluminescence (e.g., in luciferase assays). Additionally, differential binding of co-factors, epigenetic markers and RNAPII can be investigated by co-immunoprecipitation, ChIP, and DNase I hypersensitive assay on these FOXM1 target gene regions. STARR-seq in combination with co-factor disruption by CRISPRi or inhibition with small molecule inhibitors can be used to investigate genome-wide enhancer and co-factors identification, validation, and measurement of regulatory capacity and activity under different conditions in breast cancer cell lines.

Additionally, response to steroid hormone stimulus was one of the top processes. Motif enrichment analysis showed the top significantly enriched motifs in FOXM1 binding targets were TF family members of Forkhead, leucine zipper, GATA, and nuclear hormone receptor families (Figure 3.6). Expectably, FOXM1 was previously reported to regulate its own expression and activity by a positive auto-regulatory loop(71). Additionally, the recruitment of pioneer factor FOXA1 to FOXM1/ESR1 positive and FOXM1/ESR1 negative gene regions are identical and non-biased, as shown by enrichment of FOXA1 binding across FOXM1/ESR1 positive and negative peaks (Figure 3.6B). The motif enrichment of FOXA1, GATA, ESR2, E2F, HES, HINFP, EGR2 in the nucleotide sequences of FOXM1 peaks suggests potential interaction of FOXM1 with these co-activators to regulate gene expression (Figure 3.6A). Indeed, FOXM1 operates via diverse DNA binding mechanisms, either through direct DNA binding; or binding at certain genomic sites by recruitment and interacting with co-factors or transcriptional complexes (47). The comparison between discriminative motif enrichment analysis and ChIP-seq binding profiles showed predicted transcriptional factors that are key in the biological processes under study for identified binding regions of the two FOXM1 transcriptional programme (Figure 3.6).

Housekeeping genes that are fundamental regardless of developmental stages hence common across cell types were identified in FOXM1/ESR1 negative regions, whereas developmental genes that are specific to breast cells and developmental stages were shown in FOXM1/ESR1-positive regions. Unsurprisingly, common targets, such as *FOXA1* and *MYC*, were identified in both FOXM1 transcriptional programmes. FOXA1 is a known pioneer factor that can bind on compact chromatin, subsequently allowing other TFs to access the gene regions; *MYC* is a known oncogene that is involved in global chromatic structure regulation via histone acetylation (158, 159). It is understandable that they serve housekeeping and developmental roles, where the fundamental expression of both factors is crucial across all cell types, and a further enrichment may happen in specific cell types and even during cancer development. To confirm these direct TF bindings, a clear DNase I digital footprinting signature and individual ChIP techniques should follow. Similar approaches can be conducted systematically to identify delicate specificity for members of transcriptional factor superfamily that share highly identical motifs, such as the Forkhead members in this case. Furthermore, different breast cancer cell lines and varied developmental stages can be investigated to study the characteristic of the two FOXM1 transcriptional programmes.

GO enrichment analysis of FOXM1 bound genes in MCF-7 cell line showed top enriched biological processes in both housekeeping and developmental roles (Figure 3.7). In FOXM1/ESR1 negative regions, there is clear indication of housekeeping roles in cell division and cell cycle phases, as supported by similar heatmap intensities for both MCF-7 and MDA-MB-231. In FOXM1/ESR1 positive regions, developmental roles were enriched, such as tissue development and nervous system development. This supports the cell line and developmental stage specific roles of the FOXM1/ESR1 positive transcriptional programme in breast cancer MCF-7 cell line. The gene expression profiles of other cell lines with ESR1 and in different developmental stages should be explored.

In a study by Millour et al., ESR1 ligands (oestradiol (E2), tamoxifen and fulvestrant) regulate the expression level of FOXM1 protein and mRNA in ER-positive cell lines ZR-75-1 and MCF-7 cell lines (118). Genome-wide analysis indicated that FOXM1 uniquely regulate the transcription of subset of genes independent of ESR1 even in ESR1-positive cell lines, or that ESR1 potentially function along with FOXM1 at

different sites to loop together different chromatin locations (120). When ESR1 is inhibited by tamoxifen, FOXM1 replaces ESR1 and turns into main TF to induce *ERK2* to drive enhanced activation of MAPK signalling pathway. I showed that a similar alternative growth factor signalling pathway may happen upon fulvestrant treatment. The occupancy of TFs and co-factors can maintain the euchromatin state of selected DNA regions, which is loosely packed form of DNA to allow complex binding which may consist of other TFs and RNA Pol II on target promoters to drive gene expression. *FOXM1* knockdown significantly decreases the chromatin accessibility of ESR1 peaks (Figure 3.8). This role of FOXM1 is potentially associated with fulvestrant resistance in MCF-7 cell line as fulvestrant increases FOXM1 binding on both sets of FOXM1 binding regions.

The FOXM1 target gene regions in both FOXM1 transcriptional programmes identified from ChIP-seq in MCF-7 cell line were selected as readouts for further functional experiments, particularly *GREB1* and *TFF1* in FOXM1/ESR1 binding regions (Figure 3.9). The binding of FOXM1 on these FOXM1/ESR1 target regions were inhibited by ICI (fulvestrant) in MCF-7 cell line, whereas it was unaffected in fulvestrant-resistant MCF-7 (Figure 3.10). Two FOXM1 inhibitors, FDI-6 and thiostrepton, achieved the same inhibition effect on gene expression of *FOXM1*, *ESR1*, *GREB1*, and *TFF1*, although thiostrepton accomplished a larger suppressive effect (Figure 3.11). The co-regulatory relationship of FOXM1 and ESR1 seems to be a positive feedback loop.

The inhibition of FOXM1 activity in MCF-7 cell line by FDI-6 and thiostrepton validated the importance of FOXM1 in the modulation of ESR1 signalling cascades (Figure 3.11). Other than that, a differential effect between FDI-6 and thiostrepton in MCF7 cell line was observed. Thiostrepton has been the widely accepted FOXM1 inhibitor to suppress the transcriptional regulation of downstream target genes by FOXM1 (160, 161). It is a natural thiazole antibiotic that inhibits FOXM1 action by forming a tight complex with FOXM1 target DNA regions and FOXM1 protein itself via its binding domain, resulting in a downregulation of mRNA expression of *cyclin B1*, a downstream target of FOXM1 (161). As a result, thiostrepton selectively induces cell cycle arrest and cell death in breast cancer cells. Unfortunately, the usage of thiostrepton is confounded by its inhibition of the 20S proteasome. It triggers significant toxicity well below the dosage required to induce FOXM1 inhibition in cancer cells. These side-effects of thiostrepton result in uncertainties about its dominant mode of action and

therefore cannot be allowed to unambiguously target FOXM1. Later, another novel inhibitor FDI-6 was developed. In contrast to thiostrepton, FDI-6 is inactive against the proteasome, and the dosage required to target *FOXM1* is within the range of tolerable toxicity in breast cancer cells, implicating that this novel agent suffers less of the off-target effects carried by thiostrepton (162). Moreover, the smaller molecular size, greater aqueous solubility, and a scaffold open to future optimisation are strong advantages of FDI-6. Importantly, FDI-6 regulated downregulation of known FOXM1 downstream target genes, *CENPA*, *KIF20A*, *NEK2*, and *CDKN3*, which all have strong FOXM1 binding motifs in their promoter regions, aligned with the effect of a previous siFOXM1 study (162, 163). ChIP-qPCR confirmed that FDI-6 specifically targets FOXM1-regulated genes with FOXM1 binding, no effect on genes regulated by other homologous FOX TFs, indicating its specificity (Data not shown). This might explain the stronger inhibition of thiostrepton on ESR1 and ESR1 target genes, potentially occurred due to the off-target effect and less selectivity of thiostrepton. Considering the diverse implications of all FOX TFs, the specificity of FDI-6 might be more ideal to lessen the fatal side-effects from knocking down other FOX TFs. Having said that, other FOXM1 inhibitors should be tested to find the right fit for combinatorial therapy in ESR1-positive breast cancer; additionally, the same treatment should be tested in MDA-MB-231 cells and other breast cancer models to investigate the functionality of these novel inhibitors.

Furthermore, mutations in the *ESR1* gene have been highly detected in metastatic breast cancers and are usually associated with endocrine resistance and disease relapse (12, 15). Therefore, understanding the impact of these mutations on ESR1 signalling in breast cancer cells is essential to develop more effective endocrine- or chemo- therapies when ESR1 mutations are identified in advanced breast cancers. CRISPR-Cas9 technique to replace genes provides an influential new approach to gain a more accurate investigation of the effect of gene mutation and avoid problems inherent in overexpression experiments. MCF7-Y537S, a CRISPR-Cas9 generated MCF-7 mutant model, is constitutively active without dependence on E2. The tyrosine 537 in the ESR1 gene is replaced by serine 537 using CRISPR-Cas9 approach (104). The model was used to understand the mutated mechanism and resulting endocrine resistance in breast cancer cells since tyrosine Y537 in the LBD of ESR1 is the most highly mutated region in endocrine-resistant, advanced, and metastatic ESR1-positive

breast cancers. In MCF7-Y537S model, FDI-6 significantly decreased the expression of ESR1 target genes, *GREB1* and *TFF1*; and *ESR1* (Figure 3.12).

It is shown here a combinatorial treatment of FOXM1 inhibitor with endocrine therapy may exert an effect in FOXM1/ESR1 positive target sites, this may explain the mechanism behind the development of endocrine resistance characterised by MCF7-Y537S model. The combinatorial treatment of small molecule FOXM1 inhibitor FDI-6 with fulvestrant inhibited the expression of FOXM1/ESR1 target genes (*GREB1* and *TFF1*), FOXM1, and ESR1, in MCF7-Y537S cell line; and this resulted in effective growth inhibition with lower dosage of drugs (Figure 3.13). The combinatorial therapy was more effective than monotherapy alone in MCF7-Y537S cell line. These findings suggest the potential mechanism behind the development of endocrine resistance in MCF7-Y537S model. The importance of alternate anti-estrogen SERDs besides fulvestrant, or the combined treatment of SERMs and SERDs with novel targeted therapies, have been stated in the treatment of metastatic breast cancer with *ESR1* mutations (16). A study showed that a combination of oral estrogen inhibitor SERD, AZD9496, with targeted therapy, PI3K pathway inhibitor or CDK4/6 inhibitor, showed increased suppressive effects when compared to monotherapy, hinting the potential utility of combinatorial therapy in patients with mutant ESR1 (16). Additionally, there is an increasing rationale for effective combinations of endocrine therapy with other drugs that target deregulated signal transduction cascades in ESR1-positive breast cancer. Extended duration of endocrine therapy can be correlated with an acquired upregulation of estrogen-independent ESR1 signalling, often through EGFR/HER2, alongside with crosstalk activation of ESR1-dependent gene transcription that ultimately causes endocrine resistance. To overcome this resistance, combinatorial treatments usually include fulvestrant and various signal transduction inhibitors (STIs), such as tyrosine kinase inhibitors and mTOR antagonists (164). Pre-clinical data suggests that combinatorial treatment of endocrine and STIs result in larger growth inhibition than either of the monotherapy and thus postpone the development of resistance in resistant, advanced, metastatic breast cancer cells.

In the future, the optimum dosage and toxicity profile of the combinatorial treatment should be tested. The priority is to understand optimal dosage at a toxicity level tolerable by patients. Whole genome sequencing and other omics study have allowed the development of personalised combinatorial targeted therapy over the general one-

size-fits-all cancer therapies (165). Combinatorial treatments might also include agents that target epigenetic processes, such as HDAC inhibitors, and other chromatin-regulating compounds, such as demethylases, that could battle off resistance that is involved in chromatin-regulated, reversible drug resistance in cancer stem cell populations (166).

4 The relationship between FOXM1 and DDIT3 in breast cancer cell lines

4.1 Introduction

4.1.1 Breast cancer and UPR

Breast cancer covers a heterogeneous set of diseases with different prognoses, physiological and histological features – leading to distinct treatment options available for each subtype. Different breast cancer subtypes are typically diagnosed based on the presence of three receptor proteins: ESR1, PR, and HER2. A different method of subtyping is by the differential expression of fifty select genes (PAM50), which infer the "intrinsic" subtype that can assist clinicians on the right breast cancer treatments and increase survival rates (167). In PAM50 analyses, tumours with a gene expression profile representative of luminal epithelial cells belong to the luminal subtype and are usually hormone receptor-positive (ESR1+ PR+). Over 70% of breast tumours of all breast cancers are luminal and about half are often responsive to ESR1 targeted treatments, with the remaining half being de novo resistant. Even so many initially responsive tumours develop resistance to these agents later.

Among the many reasons for endocrine therapy resistance, the activation of the UPR pathway is one of them (74). UPR effector, GRP78, has an important role as a regulator of endocrine resistance. As a matter of fact, estrogen triggers the expression of GRP78 and XBP1, both being UPR molecules overexpressed in, 60 – 70% and 80 – 90% of human breast cancers, respectively (74). In this chapter, I investigated the binding of FOXM1 and DDIT3 protein on the promoter region of *GRP78* and *GRP94* in MCF-7 cell line using ChIP. Additionally, I studied the impact of FOXM1 expression on the expression levels of *GRP78* and *GRP94* in the presence of epirubicin. The correlation between expression level of TFs, FOXM1 and DDIT3, and chaperone proteins, GRP78 and GRP94 were observed using publicly available breast cancer patient samples.

4.1.2 Abnormal UPR Signalling in Breast Cancer

It has been extensively proposed that abnormal and/or chronic activation of UPR, as well as the upregulation of UPR molecules, are associated with breast cancer progression and with resistance to apoptosis and drug therapy in breast cancer cells. In breast cancer, prolonged UPR signalling may contribute to most, if not all, hallmarks of cancer and therapeutic resistance (168). UPR is by default inactive in normal cells but active in tumour cells, this distinct difference poses a strong advantage for the development of novel breast cancer treatments targeted at ER stress. Additionally, UPR features different therapeutic possibilities for different breast cancer subtypes, including TNBC that is notorious for lack of available treatment options.

The UPR is demonstrated to grant breast cancer cells with resistance to radiation therapy, tamoxifen, paclitaxel, vinca alkaloids, cisplatin, doxorubicin, HDAC inhibitors, and microtubule interfering agents (6, 168). In contrast, the UPR can promote breast cancer cell death upon other treatment options such as bortezomib, lapatinib / obatoclax joint treatment, and pan-peptidylarginase deiminase. These opposite findings emphasise the context-specific nature of UPR output and the importance of understanding the contribution of each UPR arm independently. Since the level of ER stress is also a crucial factor to determine cell fate, it is possible that the therapeutic dosage achieved in tumour cells is also crucial to decide whether UPR drives cell survival or cell death rather than the drugs' mode of action.

In breast cancer cells, stress occurs from hypoxia and nutrient deprivation arisen from cytotoxic and endocrine therapies and other evidence propose that prolonged and severe activation of the UPR is linked to therapy resistance and disease recurrence. Additionally, evidence suggested a connection between ESR1 signalling and the regulation of the UPR. GRP78 and XBP1 are upregulated in endocrine- and chemotherapy- resistant breast cancers. Therefore, a clear molecular understanding of the chronic UPR signalling pathway may provide novel targets to overcome drug resistance.

4.1.3 DDIT3

DDIT3 is the main initiating molecule for the pro-apoptotic arm of UPR (Figure 1.8). The elimination of DDIT3 leads to cells escaping death. It plays as a downstream

effector of ER stress signalling pathways at the convergence of IRE1, PERK, and ATF6 axes, and overexpression of the 27-kDa DDIT3 protein promotes cytotoxicity and consequent apoptosis. It has been proposed that DDIT3 may act as a valuable biomarker of the aggressiveness of tumours in mesothelioma and is predictive of drug sensitivity in thyroid cancer (169, 170). However, the predictive value of DDIT3 has not been fully understood in breast cancer. In 2014, Zheng et al. have studied the predictive values of GRP78 and DDIT3 in 250 tumour samples from female patients diagnosed with invasive ductal breast carcinoma (87). It was shown that DDIT3 correlated with prolonged DFS, whereas GRP78 indicated the opposite effect. In the GRP78-positive population, DDIT3 overexpression correlated with a lower risk of recurrence. This suggested that a high level of expression of DDIT3 is predictive of a switch to a pro-apoptotic role. The integration of GRP78 and DDIT3 into the standard prediction model significantly improved its accuracy.

4.2 Aims and Objectives

Recent limelight has been placed on the importance of ER stress and stress-induced UPR pathways in breast cancers with increasing clarity on their molecular mechanisms. Particularly, it has been reported that abnormal activation of UPR signalling, as well as the upregulation of UPR components, are involved in breast cancer development, anti-apoptotic pathways and chemotherapy resistance in breast cancer cells (74, 86, 90, 168, 171). Therefore, UPR molecules have risen as potential targets for the development of new breast cancer predictive markers or even therapeutic targets. Nonetheless, further investigation of these novel or potential ER stress markers is essential to justify their potential predictive value in breast cancers. For instance, GRP78 and DDIT3 are highly expressed in breast cancer, they might serve as potential therapeutic targets for specific subtypes in this heterogeneous set of diseases to help clinicians make better-informed decisions on prescribed adjuvant chemotherapy options (87). Previous research from our laboratory has discussed a relationship between FOX TF and ER stress (172-174). Additionally, another study proposed a link between ER stress downstream effector DDIT3 and FOXO3 in apoptosis in neuronal cells (80). In this chapter, I researched into the role of FOXM1 in the context of the UPR signalling pathways, focusing on breast cancer cell lines as study models. Particularly, I deep dived into the relationship between FOXM1 and DDIT3 in different breast cancer subtypes.

4.3 Results

4.3.1 FOXM1 and DDIT3 bind to the promoter regions of *GRP78* and *GRP94* in MCF-7 cell line

As a rule of thumb, cells mainly favour survival when they encounter stressful conditions. For instance, cells would preferably initiate the pro-survival arm of the UPR if protein homeostasis can be restored (86). Specifically, when cells suffer from ER stress, GRP78 and GRP94 would strive to increase protein folding capacity and restore protein homeostasis in ER lumen.

Preliminary study in our laboratory has suggested a relationship between FOXM1 and DDIT3. Since DDIT3 has a well-known role in the regulation of ER stress, I investigated the possibility of a relationship between FOXM1 and DDIT3 in this mechanism. Using ChIP-seq data of FOXM1 and DDIT3, a list of overlapped UPR regulators were identified through the presence of binding peaks on their promoter regions. Out of all targets, I detected positive binding peaks of both FOXM1 and DDIT3 on the gene promoters of *GRP78* and *GRP94.* To confirm the bindings, ChIP analyses were performed using anti-FOXM1 and anti-DDIT3 antibodies with designed primers targeting *GRP78* and *GRP94.*

The RT-qPCR experiments validated the binding of FOXM1 and DDIT3 on the promoters of *GRP78* and *GRP94* in MCF-7 cell line. Additionally, the same experiments were performed in the presence of epirubicin (1 μM) or epirubicin-resistant cell line, MCF-7Epi^{R}. Epirubicin treatment of 24 hours significantly disrupted FOXM1 and DDIT3 bindings on the promoters of *GRP78* and *GRP94* in MCF-7 cell line. In MCF-7Epi^{R} cell line, FOXM1 ChIP showed an enhanced binding on the promoter regions of *GRP78* and *GRP94* when compared to negative controls. Surprisingly the 24-hours epirubicin treatment was also effective in epirubicin resistant cell line, with diminishing bindings observed upon treatment. On the other hand, DDIT3 ChIP on the promoter regions of *GRP78* and *GRP94* did not show visible bindings in MCF-7Epi^{R} cell line. However, when epirubicin resistant cell line was treated for 24 hours, DDIT3 demonstrated increased occupancy on *GRP78* and *GRP94* promoters (Figure 4.1).

The initial hypothesis suggested a co-regulatory relationship between FOXM1 and DDIT3 on common transcriptional targets. It was thought that these two known TFs

would piggy-back each other and act as either co-activator or co-suppressor. Preliminary data in our laboratory did not show a protein-protein interaction between FOXM1 and DDIT3 when immunoprecipitation experiment was performed in MCF-7 cell line (Data not shown). Here, the ChIP results showed that FOXM1 and DDIT3 are directly bound on the promoter regions of *GRP78* and *GRP94* (Figure 4.1*)*. In the next section, the functional implications of these protein-DNA interactions are further explored.

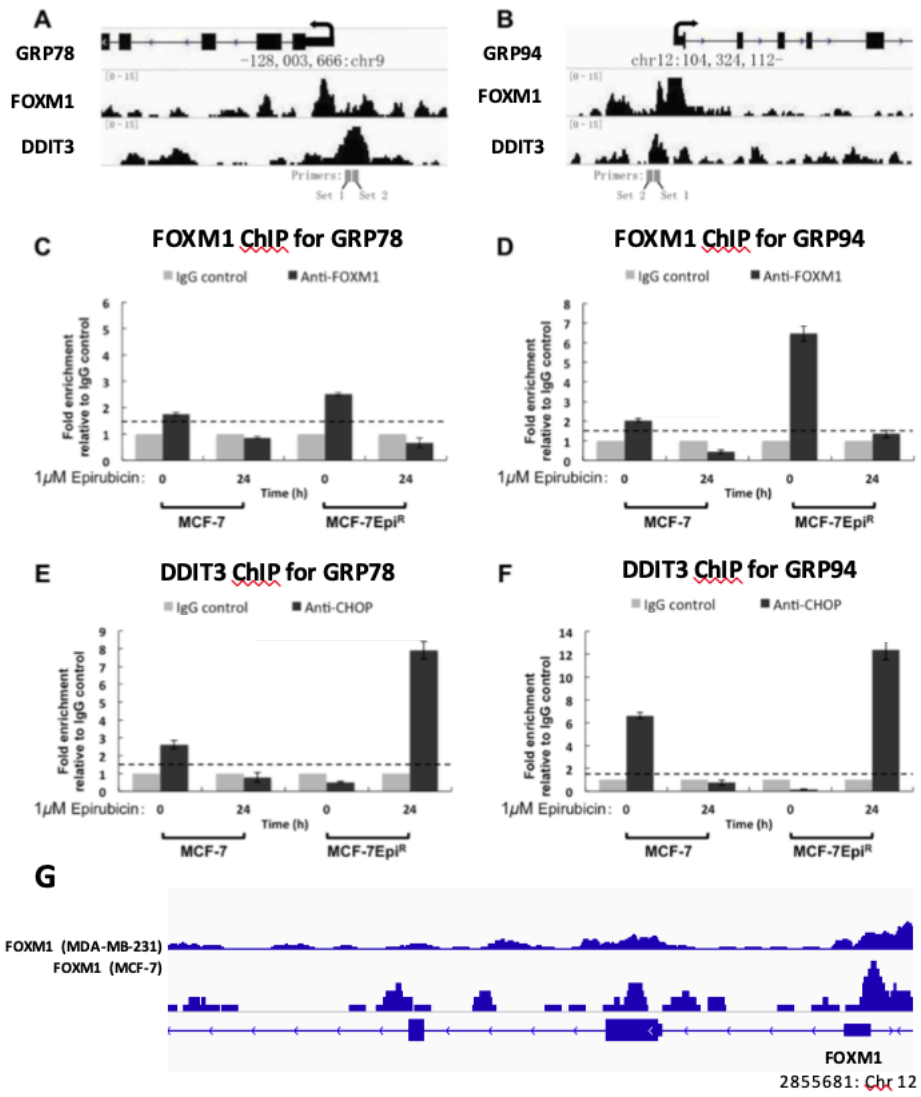


Figure 4.1 Differential binding effects seen for FOXM1 and DDIT3 binding on the promoter regions of GRP78 and GRP94.

(A) ChIP-Seq findings suggest binding of FOXM1 and DDIT3 on the promoter regions of *GRP78* and *GRP94.* Top panels show gene structure provided by RefSeq database, arrows indicate transcription start sites. **(B - F)** ChIP on *GRP78* and *GRP94* promoter regions with anti-FOXM1 and anti-DDIT3 antibodies in MCF-7 and MCF-7EpiR cells with 24h epirubicin treatment (1 μM) or untreated. Results

are representative of 3 technical replicates of ChIP experiments. As the data shown represent mean ± SEM of three technical repeats, no statistical analysis was performed. **(G)** Reference peaks of ChIP: FOXM1 in MCF-7 and MDA-MB-231 cell line on *foxm1* promoter region.

4.3.2 Epirubicin treatment downregulates the expression of FOXM1 and DDIT3

Time-course treatment with 1 μM of epirubicin decreased the protein expression level of FOXM1 and DDIT3 (Figure 4.2). This decrease was not reflected in the protein expression level of GRP78 and GRP94. Circling back to our hypothesis, the decrease in FOXM1 and DDIT3 did not affect the protein level of GRP78 and GRP94. If there was indeed a regulatory mechanism, it would mean that there could be other signalling pathways that had compensated for the absence of FOXM1 and DDIT3. Additionally, the protein expression of FOXM1 and DDIT3 are not the best measure of their transcriptional activity. Furthermore, the protein expression and activity of UPR molecules are temporal-dependent on ER stress, therefore only happen on a contextual basis.

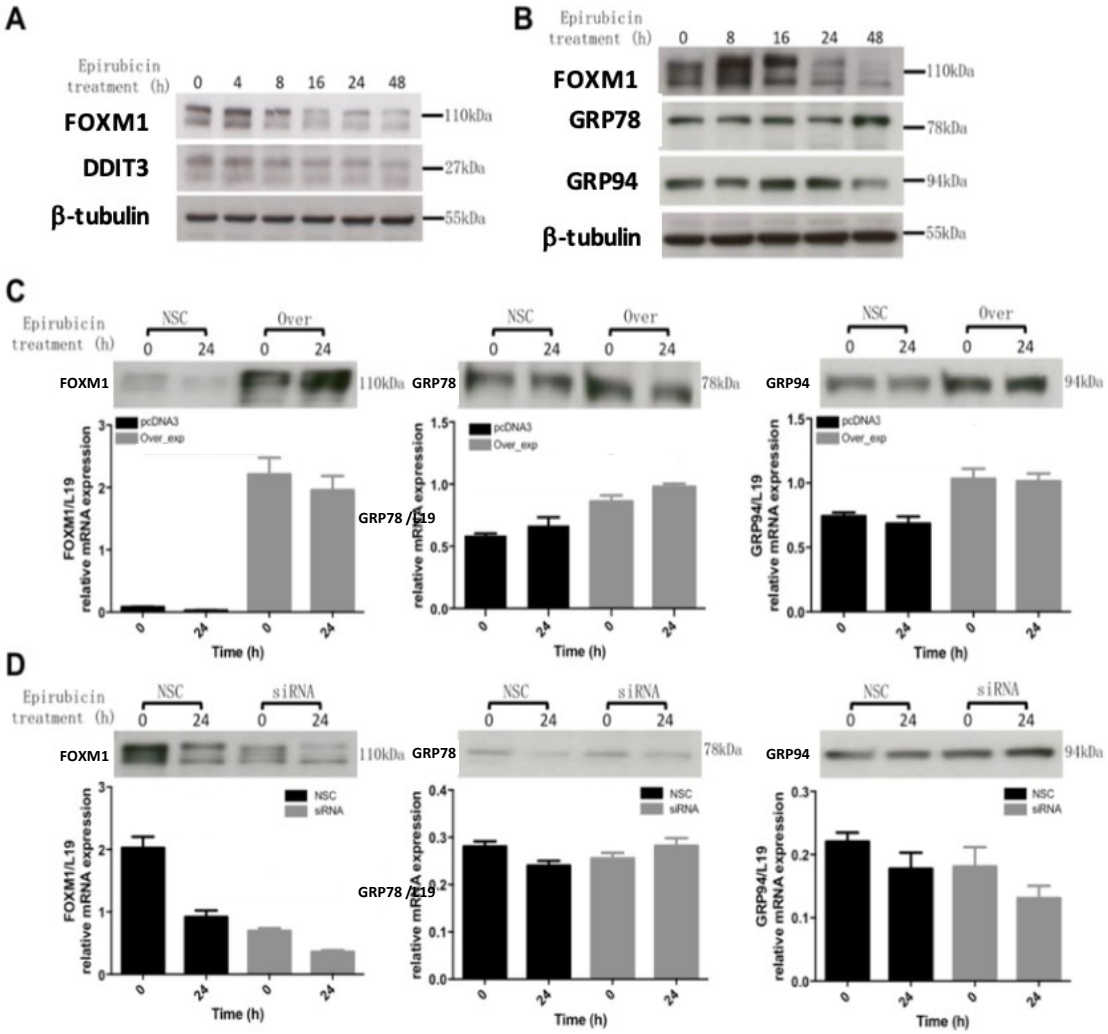


Figure 4.2 The overexpression or downregulation of FOXM1 impacts the expression levels of GRP78 and GRP94 in the absence and presence of epirubicin.

(A) – (D) Cell lysates were harvested from MCF-7 cell line after epirubicin treatment (1 μM) at the indicated timepoints for western blotting and RT-qPCR analyses. RT-qPCR experiments measuring mRNA level were carried out in technical triplicates, data were normalised to housekeeping genes, L19 and presented as mean ± SEM. **(C)** and **(D)** MCF-7 cell line were cultured for both experiments. Overexpression of FOXM1 was transient via transfection with either FOXM1 plasmid construct, pcDNA-FOXM1 or empty vector control, pcDNA3. Downregulation of FOXM1 was achieved with siRNA targeting FOXM1 and a control siRNA targeting no known mammalian gene (NSC) was included. Untreated cells and cells treated with epirubicin for 24h were harvested at 48h after transfection. As the data shown represent mean ± SEM of three technical repeats, no statistical analysis was performed.

4.3.3 FOXO3 activation represses FOXM1 and DDIT3 expression

The expression level of FOXM1 was significantly reduced upon epirubicin treatment in MCF-7 cells (Figure 4.2). This inhibition may happen via the transcriptional repression of FOXO3 as previously reported (175). FOXO3 may act as an important upstream modulator of *FOXM1* expression under stress conditions. The activation of FOXO3 is tightly regulated by various post-translational modifications, such as phosphorylation by the serine/threonine kinase AKT that causes the inactivation of FOXO3 and guides it to degradation (176). The DL23 system is featured by colorectal adenocarcinoma DLD1 cells with a fusion protein of a constitutively active FOXO3 mutant lacking AKT phosphorylation sites and transcriptionally inactive ESR1 with a modified ligand-responsive domain that induces FOXO3 activity in response to 4-hydroxytamoxifen (4OHT). Additionally, the DL23 cells contain similar levels of inducible active FOXO3 mutants compared to endogenous FOXO3 in DLD1 cells, thus providing a suitable system to isolate and visualise the impact of FOXO3 activation.

The microarray data was adopted to investigate the global changes in gene expression in DL23 cells following FOXO3 activation (Figure 4.3). A substantial decrease in the expression level of FOXM1 and DDIT3, with downregulation of 3.16-fold and 2.49-fold respectively, was observed after 4OHT treatment. In contrast, no significant difference in *GRP78* and *GRP94* expression was detected. To further relate the suppression of FOXM1 and DDIT3 to their transcriptional roles, different ChIP-seq data were analysed to study the effects of FOXO3 activation. Only a small overlap of target gene promoter regions was observed of FOXM1 and FOXO3, either endogenously induced by AKT inhibition in DLD1 cells or exogenously activated by 4OHT treatment in DL23 cells (Figure 4.3). The majority of FOXO3 binding sites are located at the distal intergenic regions and only a minority of FOXO3 binding sites were detected within 1 kb upstream of the TSS. Moreover, none of the promoters of *FOXM1*, *DDIT3*, *GRP78* or *GRP94* were targeted by active FOXO3, as no peak region was identified (data not shown). This proposes that FOXO3 may regulate the expression of *FOXM1* and *DDIT3* via other mechanisms. Furthermore, ChIP-seq analyses of RNA Pol II and the active histone mark H3K27Ac were performed to study the role of FOXO3 in regulating transcription initiation and enhancer activity (Figure 4.3). As a reference to the DL23 system with inducible active FOXO3 responsive to 4OHT, identical peak locations of RNA Pol II occupancy were observed in DLD1 cells before and after 4OHT treatment.

However, the recruitment of RNA Pol II and H3K27Ac levels near the TSS remained largely unimpacted upon FOXO3 activation in DL23 cells.

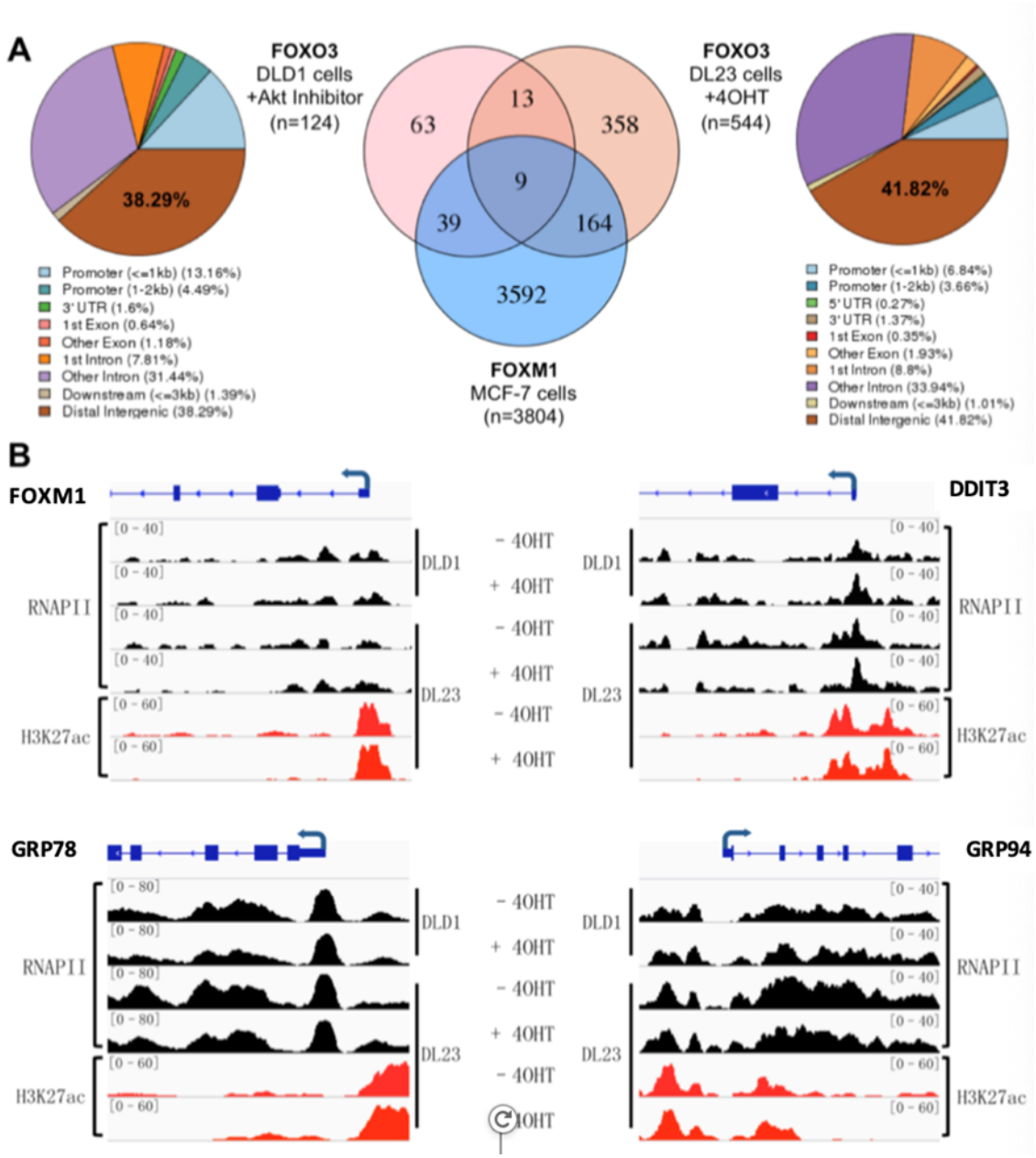


Figure 4.3 The relationship between FOX TFs and DDIT3, GRP78 and GRP94.

(A) Global distribution of binding regions of activated FOXO3 in DLD1 and DL23 cells. Venn diagram shows the overlap between binding profiles on promoter regions of activated FOXO3 in DLD1 cells (pink), FOXM1 in MCF-7 cells (blue) and activated FOXO3 in DL23 cells (brown). n = the total number of genes with enriched peaks identified within promoter regions from ChIP-Seq analyses. **(B)** Binding peaks of RNA Polymerase II and histone modification H3K27Ac around transcription start site, in the absence and presence of 4-OHT treatment.

4.3.4 The correlation of FOXM1 and DDIT3 with GRP78 and GRP94 were observed in breast cancer patient samples

To further understand the correlation between the molecules, differential gene expression profiles were plotted by consolidating 1,215 invasive breast carcinoma patient samples extracted from The Cancer Genome Atlas (TCGA) database (Figure 4.4). Raw gene expression data was analysed with Illumina HiSeq 2000 RNA Sequencing (RNA-Seq) system. The expression profiles were categorised into two separate subgroups, either with 'high' or 'low' expression. The gene expression of *GRP78* and *GRP94* were significantly upregulated in breast cancer samples that have high expression of FOXM1 or DDIT3. Additionally, the overexpression of FOXM1 positively correlates with high tumour grade, whereas DDIT3 has a negative correlation with malignancy.

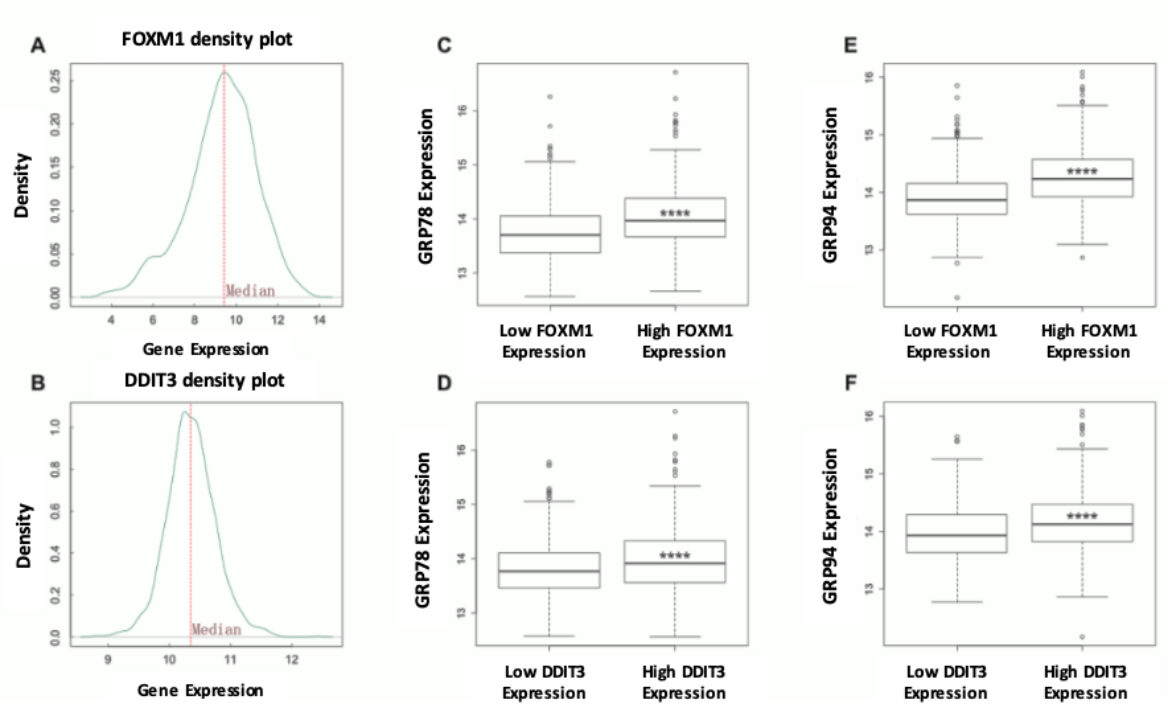


Figure 4.4 The correlation between expression level of TFs, FOXM1 and DDIT3, and chaperone proteins, GRP78 and GRP94.

A) - (B) Distribution of gene expression data. Median was used to divide 'high' and 'low' expression level. **(C) – (F)** Differential gene expression of GRP78 and GRP94 induced upon differential gene expression of FOXM1 and DDIT3. ****p < 0.0001. I am thankful to Mr. Yuliang He for this figure.

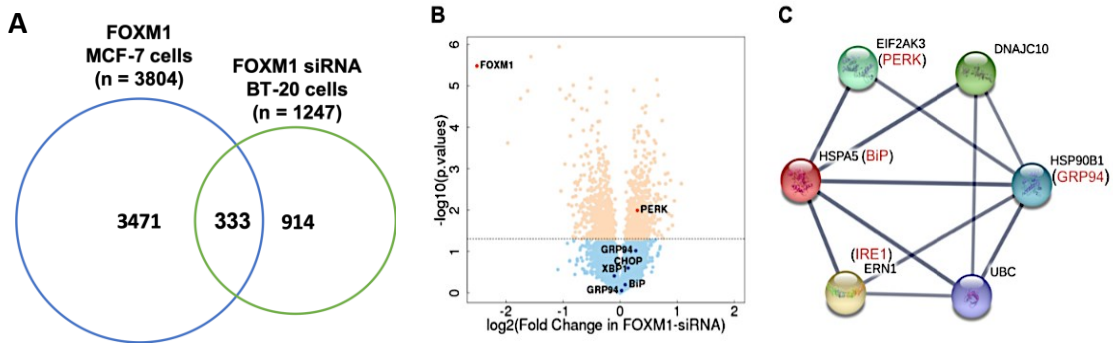


Figure 4.5 Microarray data of FOXM1 knockdown.

(A) Venn diagram shows the overlap between FOXM1 binding profiles upon FOXM1 knockdown in BT-20 cells as compared to wildtype MCF-7 cells. **(B)** Volcano plot shows statistical significance of gene expressions in the absence and presence of FOXM1 knockdown. Unpaired t-tests were performed to test the statistical significance and p-values < 0.05 was set as the cut-off threshold, as shown by the dotted line. Y-axis represents the -log10 of p-values, whereas the log2(Fold Changes) in FOXM1 knockdown versus control is shown on the x-axis. Each dot represents one gene. **(C)** String plot shows known or predicted interactions between protein molecules of interest. I am grateful to Mr Yuliang He for this figure.

Among the profile of genes that were differentially expressed upon FOXM1 knockdown in TNBC BT-20 cell line, the expression of *PERK* was mildly increased (Figure 4.5). Some of the known or predicted protein-protein interactions of GRP78 and GRP94 were presented in the STRING plot, which sheds some light on the currently known regulatory network of the UPR. GRP78 and GRP94 may work with DnaJ heat shock protein family member C10 (DNAJC10) and ubiquitin C (UBC), which are involved in the degradation of misfolded and unfolded proteins, to maintain ER homeostatic balance.

4.3.5 The transcriptional binding profile of FOXM1 and DDIT3 in cancer cell lines

Similarly, all ChIP-seq data in different cell lines were downloaded from the Cistrome database (177). ChIP-seq data of FOXM1 in breast cancer MCF-7 and MDA-MB-231 cell lines were selected to compare against ChIP-seq data of DDIT3. Due to a lack of public ChIP-seq data of DDIT3 in breast cancer cell lines, DDIT3 transcriptional profile in myxoid liposarcoma MLS402 cell lines was compared and analysed (Figure 4.6). The deviation in cell lines involved (from the focus in breast cancer models) added

ambiguity to the experiments and is the main incompleteness of this study, therefore these results may only serve as preliminary work to serve as pointers to unravel future studies (Figure 4.6).

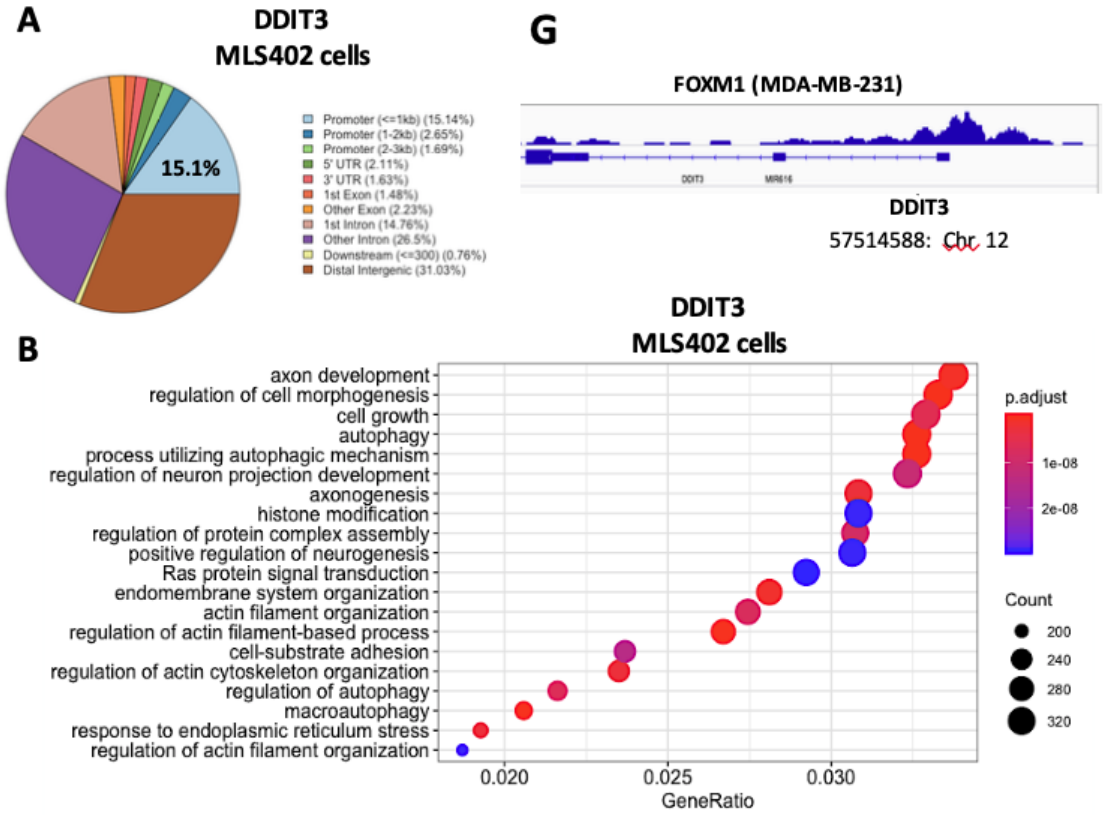
Studies have suggested a subtype-specific role of ER stress and UPR in breast cancer, specifically in luminal A and basal subtypes (74). On the other hand, myxoid liposarcoma is a malignant tumour that originates from deep soft tissues in fatty tissues, which is characterised by FUS-DDIT3 fusion oncoprotein that functions as a mutated TF (178). Like FOXM1, FUS is another diverse protein associated with myriad transcriptional regulations, DNA damage repair mechanisms, and splicing processes. As mentioned, FOXM1 is a bona fide oncogenic TF that is known for its tumorigenic role in breast cancer, whereas DDIT3 is recognised for its apoptotic role in cells when ER stress goes out of control.

Due to the limitations in resources, ChIP-seq data uploaded on the public database, CistromeDB, were used as main resource in place of conducting ChIP-seq experiments from scratch. GEO and CistromeDB ID of each ChIP-Seq data presented in this chapter are stated in **Section 2.11.2.1 Cistrome Data Browser (Cistrome DB) Data Mining**. Raw data from ChIP-seq was trimmed and mapped to UCSC human genome, hg19 using Bowtie2 then visualised on IGV. R Bioconductor package 'ChIPseeker' was then loaded on R Studio to annotate peaks to understand the genomic context of the bindings (109). The mapped data was then loaded on Python and sorted with 'bioconda' package with bedtools 'sort-bed' to sort peaks. After the identified genes are associated with the ChIP-seq peaks, pathway enrichment analysis was performed using 'clusterProfiler' package developed by Guangchuang Yu (110).In order to find the overlapped peaks (genomic bindings) of FOXM1 and DDIT3, 'bedops' tool 'intersect' command was used to intersect overlapped peaks (111, 112).

GO analysis from 'clusterProfiler' indicates that FOXM1 and DDIT3 are associated with widely diverse biological processes. The role of FOXM1 lies in RNA splicing and processing, and cell cycle regulation in breast cancer subtype MCF-7 and MDA-MB-231 cell lines. On the other hand, DDIT3 is involved in ER stress, cell growth, cytoskeleton and membrane organisation, signal transduction of MAPK/ERK/Ras pathway, and autophagy. The genome-wide binding of DDIT3 in myxoid liposarcoma MLS402 cell line showed approximately 31% of genome binding were in the intergenic

regions. Roughly 19% of DDIT3 bindings were identified within 3 kb of Transcription Start Site (TSS), among which 15% was within 1 kb of TSS.

Next, I proceeded to investigate the overlap of ChIP-seq peaks of FOXM1 in MCF-7 cell line and DDIT3 in MLS402 cell line. Roughly 57.4% of FOXM1 transcriptional targets in MCF-7 are shared with that of DDIT3 in MLS402 cell line, 30% of the common target bindings happened within 1 kb of TSS. GO analyses showed that the functional pathways of these overlapped genes were associated with the regulation of protein serine/threonine kinase activity, RNA metabolic processes such as RNA splicing, protein catabolic processes such as translation, translocation and organisation, and p53 signal transduction. On the other hand, 58% of FOXM1 bindings in MDA-MB-231 cell line overlapped with the transcriptional profile of DDIT3 in MLS403 cell line. Comparably, these transcriptional targets were involved in RNA metabolic processes, RNA splicing, nuclear translocation, and ER stress.


A
DDIT3
MLS402 cells
15.1%
Promoter (<=1kb) (15.14%)
Promoter (1-2kb) (2.65%)
Promoter (2-3kb) (1.69%)
5' UTR (2.11%)
3' UTR (1.63%)
1st Exon (1.48%)
Other Exon (2.23%)
1st Intron (14.76%)
Other Intron (26.5%)
Downstream (<=300) (0.76%)
Distal Intergenic (31.03%)
G
FOXM1 (MDA-MB-231)
DDIT3
57514588: Chr 12
B
DDIT3
MLS402 cells
axon development
regulation of cell morphogenesis
cell growth
autophagy
process utilizing autophagic mechanism
regulation of neuron projection development
axonogenesis
histone modification
regulation of protein complex assembly
positive regulation of neurogenesis
Ras protein signal transduction
endomembrane system organization
actin filament organization
regulation of actin filament-based process
cell-substrate adhesion
regulation of actin cytoskeleton organization
regulation of autophagy
macroautophagy
response to endoplasmic reticulum stress
regulation of actin filament organization
0.020
0.025
0.030
GeneRatio
p.adjust
1e-08
2e-08
Count
200
240
280
320
C
Common binding targets
FOXM1 MCF-7 cells and DDIT3 MLS402 cells
30%
Promoter (<=1kb) (29.97%)
Promoter (1-2kb) (2.26%)
Promoter (2-3kb) (1.5%)
5' UTR (3.33%)
3' UTR (0.95%)
1st Exon (1.65%)
Other Exon (1.63%)
1st Intron (12.29%)
Other Intron (19.2%)
Downstream (<=300) (0.53%)
Distal Intergenic (26.69%)
FOXM1
MCF-7 cells
(n = 12404)
5280
(29.4%)
7124
(39.7%)
5561
(31%)
DDIT3
MLS402 cells
(n = 12685)
D
Common binding targets
FOXM1 MDA-MB-231 cells and DDIT3 MLS402 cells
51.8%
Promoter (<=1kb) (51.82%)
Promoter (1-2kb) (2.21%)
Promoter (2-3kb) (1.12%)
5' UTR (5.32%)
3' UTR (0.42%)
1st Exon (1.9%)
Other Exon (1.18%)
1st Intron (7.88%)
Other Intron (11.96%)
Downstream (<=300) (0.27%)
Distal Intergenic (15.92%)
FOXM1
MDA-MB-231
cells
(n = 8948)
3752
(22.8%)
5196
(31.6%)
7489
(45.6%)
DDIT3
MLS402 cells
(n = 12685)

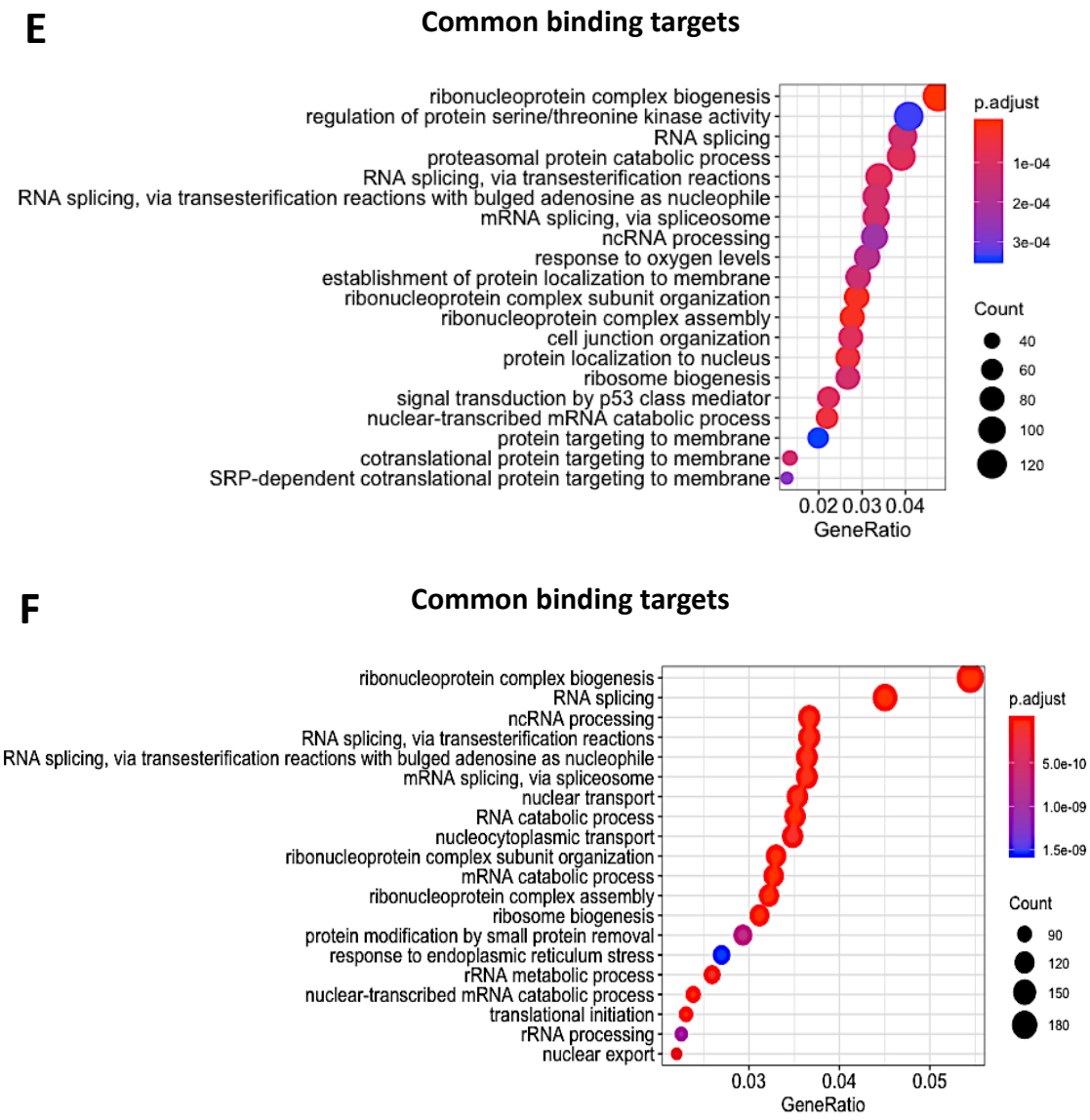


Figure 4.6 The potential transcriptional co-regulatory role of FOXM1 and DDIT3 in breast cancer subtypes.

(A) Global distribution of genomic binding regions and **(B)** GO analyses of DDIT3 in myxoid liposarcoma MLS402 cells. Common targets between DDIT3 in MLS402 cells and **(C)** FOXM1 in MCF-7 cell line, **(D)** FOXM1 in MDA-MB-231 cell line. GO analyses of common targets between DDIT3 in MLS402 cells and **(E)** FOXM1 in MCF-7 cell line, **(F)** FOXM1 in MDA-MB-231 cell line, **(G)** FOXM1 ChIP-Seq in MDA-MB-231 cell line detected a binding peak within 1kb of TSS of DDIT3 gene region.

4.4 Discussion

UPR pathways are activated upon ER stress in cells, it plays an essential part in tumour progression and the development of chemotherapy resistance (73). The role of UPR in breast cancer is still clouded with plenty of unanswered questions. Fundamental research has reported promising evidence to target UPR molecules, such as IRE1/XBP1, PERK, and GRP78, to better breast cancer treatment outcomes in patients. To date, UPR consists of three known, different, but interdependent signalling cascades, each with a plethora of known functions, hence, decoding the exact mechanisms on how UPR is involved in each cancer type and tumour growth on a contextual basis remains to be a challenging task.

In this chapter, I demonstrated the transcriptional regulation of FOXM1 and DDIT3 on the UPR signalling molecule GRP78, as well as another ER chaperone GRP94 (Figure 4.1). These molecular chaperones are critical for the restoration of protein folding for cellular adaptation and survival under stressful conditions. However, the abnormal upregulation of GRP78 may also indicate chemotherapy resistance (179). In the ChIP experiment, the binding of FOXM1 and DDIT3 on *GRP78* and *GRP94* promoter regions in MCF-7 cell line diminished upon epirubicin treatment, which could be attributed to the suppression of FOXM1 and expression of DDIT3 via FOXO3 activation, as reinforced by microarray analysis (180).

In untreated MCF7-Epi^R cell line, FOXM1 molecule bound on the promoter region of *GRP78* and *GRP94* and this may result in an increase of protein-folding capacity in ER and restoration of ER homeostasis (Figure 4.1). On the contrary, when MCF7-Epi^R cell line was treated with epirubicin for 24 hours, FOXM1 dissociated from GRP78 and GRP94 whereas DDIT3 bound on the promoter regions. DDIT3 may act as a rheostat in the regulation of epirubicin resistance and potentiates either a pro-apoptotic and pro-survival effect in MCF7-Epi^R cell line.

However, epirubicin treatment, *FOXM1* knockdown, or isolated FOXO3 did not impact the expression level of *GRP78* or *GRP94* (Figure 4.2). This proposes that other arms of the UPR signalling pathways are compensating to ensure expression of these protein chaperones, despite the downregulation of FOXM1 and DDIT3. Despite the removal of FOXM1 binding on *GRP78* and *GRP94* promoters upon epirubicin treatment, the binding of DDIT3 to the promoter regions was drastically amplified in

MCF7-Epi^{R} cell line (Figure 4.1). If the binding of FOXM1 and DDIT3 has a transcriptional effect on the expression of *GRP78* and *GRP94,* the effect would be reflected in promoter activity, thus measurable with luciferase reporter assay.

The activation of DDIT3 via the PERK/ATF4 pathway is important to the transcriptional activation of autophagy-related genes necessary for the formation and functional role of the autophagy organelle, namely the autophagosome (77). This was supported by the GO term analysis of DDIT3 in MLS402 cell line (Figure 4.6B). Under stress conditions, autophagy acts as a protective mechanism to promote cell survival by degrading unfolded or misfolded proteins and maintaining ER homeostatic balance. Furthermore, DDIT3 binding profiles in MLS402 cells were found mainly in distal intergenic regions, hinting a potential for its enhancer binding property (Figure 4.6A). Due to the limitation in resources, the binding profile of DDIT3 in breast cancer cell lines were not investigated and must be done in the future. Interestingly, the common binding targets of DDIT3 in MLS402 cells and FOXM1 in MCF-7 cells showed similar phenomenon, again indicating that the cell line specific co-regulatory transcriptional relationship of these two factors could lie in the enhancer regions (Figure 4.6A).

It is proposed that the cytotoxic effects of anthracyclines are modulated by the activation of FOXO3, which is associated with the regulation of FOXM1. Anthracyclines can trigger the activation of FOXO3 mainly though the suppression of AKT activity, the endogenous activation of FOXO3 via AKT inhibition in DLD1 cells and the initiation of constitutively active FOXO3 mutant with the absence of AKT phosphorylation sites in DL23 cells, hence hopefully mimic the effects of epirubicin treatment (Figure 4.3). Microarray data analysis suggests a substantial reduction in FOXM1 and DDIT3 expression levels following FOXO3 activation (180), which aligns with a decrease in both of their protein expression levels upon epirubicin treatment (Figure 4.2). Having said that, the expression at the gene level could be affected by mRNA stability, which may not be a reliable readout of the direct impact of TF activity on downstream target genes. Otherwise, the changes in RNA Pol II occupancy could be more of a direct readout of the transcription rate and a better measure of TF activity (61). Therefore, here the genome-wide RNA Pol II binding profiles are investigated to visualise the transcriptional role of FOXO3 (Figure 4.3). Neither a change in RNA Pol II binding profiles nor a shift in binding was seen upon 4OHT treatments, suggesting that FOXO3 activation does not suppress the expression of FOXM1 or DDIT3 via the

recruitment of RNA Pol II or the pause release mechanism. The comparisons between the changes in FOXM1 binding on all coding regions from ChIP-Seq data and gene expression profiles from microarray analysis have shown only a low degree, less than 20%, of overlap in downregulated genes (Figure 4.5). This suggests the intrinsic difference in the nature of the two technologies – microarray screening and ChIP-seq analysis, in characterising the transcriptional binding profiles upon FOXO3 activation.

Nevertheless, ChIP-Seq data from histone modification H3K27Ac, a critical enhancer mark associated with active promoters, was analysed to investigate the effects of FOXO3 activation on enhancer activity (181). Expectably, the level of H3K27Ac modification was mostly unaffected by the activation of FOXO3 in DLD1 and DL23 cells. Additionally, the activation of FOXO3, either triggered endogenously or exogenously, does not enhance binding on the promoter regions of *FOXM1* and *DDIT3*, or their downstream targets, *GRP78* and *GRP94*. This is a contrast to a recent study demonstrating potential mechanisms of how FOXO3 negatively regulates FOXM1 in breast cancer cells: 1) by repressing FOXM1 expression via binding to *FOXM1* promoter; 2) by competing for FOXM1 target gene promoter regions and their antagonistic functions in regulating target gene activity; 3) by recruiting HDACs, that controls active histone marks and promotes chromatin remodelling, on the promoter regions of FOXM1 downstream targets (182, 183). However, this ambiguity could be attributed to the genetic alteration profiles and the differential regulation of gene expression by FOXO3 across cell types (62). Other than that, though DL23 cells provide a system that prevents unwanted effects of AKT suppression and isolate the impact of FOXO3 activation, exogenously or endogenously activated FOXO3 only shared a small sample of mutual target promoter regions. At the same time, there is only a 34% overlap in all identified regions (61). In conclusion, ChIP-Seq analysis in the DL23 cells may not entirely represent the binding profiles of endogenous FOXO3 on the promoter regions, upon activation by epirubicin treatment. A different study has also reported an inhibition of FOXM1 by the activation of FOXO3, as a downstream target of the proto-oncogene Myc (184). FOXO3 activation may induce the expression of a transcriptional inhibitor to displace the binding of Myc on *FOXM1* promoters, hence suggesting an alternative pathway where *FOXM1* expression could be targeted by FOXO3 under stress conditions (185). Interestingly, the binding patterns of FOXO3 on the intergenic regions have been connected to the activation of enhancers in the

regulation of gene expression. The enriched localisation of active histone modifications could indicate FOXO3 binding to enhancers, which consequently further amplifies the levels of enhancer activity marks to upregulate specific target genes (176).

In a nutshell, FOXM1 can work with an important UPR signalling effector, DDIT3, to upregulate protein chaperones, GRP78, and GRP94 that restores protein folding and maintains ER homeostatic balance. This mechanism may act as a switch between the pro-survival and pro-apoptosis arm of the UPR, as indicated by the regulation of FOXM1 on oncogenes and DDIT3 on tumour suppressor / apoptotic target genes. This draws some insights into the involvement of FOXM1 in the UPR signalling pathway to drive cell survival upon ER stress and therefore advancing our knowledge of how the upregulated expression of FOXM1 can confer chemotherapy resistance in breast cancer cells. This regulation could be cancer- and cell line-specific and therefore may serve as a promising therapeutic target for breast cancer treatment and chemotherapy resistance.

5 The relationship between FOXM1 and DNMTs in breast cancer cell lines

5.1 Introduction

5.1.1 Epigenetics and Breast Cancers

Literature has observed differential expression of the genes associated with genetics and epigenetic modifications of chromatin, such as methylation and acetylation status of DNA and histone/protein, in breast cancer tissues relative to normal breast tissues (3, 95, 96). Sequential treatment with demethylating agents, namely 5-Aza-deoxycytidine and HDAC inhibitor Trichostatin A, that target aberrant DNA methylation and histone modifications; then doxorubicin showed resensitisation of breast cancer resistant cells to doxorubicin (186). “Druggable” epigenetic enzymes are particularly rapidly evolving and are important targets of adjuvant therapies in personalised medicine for the reversible process of epigenetic modulation in breast carcinogenesis, including proliferation, energetics, immortalisation, angiogenesis, and invasion (187).

5.1.2 FOXM1 and epigenetic regulation in breast cancer

Global gene expression analysis has showed that *FOXM1* is an oncogene that is highly overexpressed in various cancers, making it an attractive biomarker with its role as master regulator in tumorigenesis (56, 188). FOXM1 has been known as a transcriptional activator involved in driving expression of cell cycle genes, as elaborated in **Introduction Section 1.6.3 FOXM1**. Not surprisingly, upregulated expression of *FOXM1* in human breast cancer is associated with an undifferentiated tumour phenotype and a negative clinical prognosis. On the other hand, zinc-finger transcription factor GATA-3 is highly correlated with ESR1 expression and is a marker for positive clinical outcome of breast cancer. It is critical as a key regulator of luminal differentiation for proper mammary gland development. Hyperplasia is often an initial stage in the development of cancer, which is observed after an acute loss of *GATA-3* in the mammary gland, acknowledging it as a tumour suppressor. As tumour grade increases, the expression of *GATA-3* is silenced by several mechanisms including DNA methylation.

The expression of *FOXM1* and *GATA-3* in human tumour arrays extracted from ONCOMINE™ breast cancer datasets reveals an inverse expression pattern of *FOXM1* and *GATA-3* (189). Interestingly, Carr et al. in 2012 has identified a novel function of FOXM1 as a transcriptional suppressor in the differentiation of luminal epithelial progenitors, which the pathway was modulated via the interaction between FOXM1, Rb and DNA methyltransferase 3B (DNMT3B) (128). Through interaction with DNMT3B, FOXM1 promotes methylation of the *GATA-3* promoter in an Rb-dependent manner. This resulted in transcriptionally silencing of *GATA-3* expression, resulting in an inhibition of GATA-3 downstream network of genes and preventing the transition from progenitor to differentiated mammary luminal cells. Other than that, FOXM1/Rb complex can suppress *GATA-3* in breast cancer MDA-MB-453 and MCF-7 cells. In MDA-MB-468 cells with Rb mutations, FOXM1 was unable to suppress *GATA-3* expression. This indicates that FOXM1 is likely a co-regulator in transcriptional suppression of target genes, rather than a standalone regulator.

5.2 Aims and Objectives

Breast cancer epigenetics is a field of ongoing research that continues to expand our knowledge of the heterogeneity in breast cancer signalling. Epigenome studies showed stratification of breast cancer subtypes, hinting essential roles in breast cancer diagnosis and treatment (190). Further investigations are required to reveal the deregulation of epigenetic programmes and the crosstalk with other relevant pathways in breast tumorigenesis. In this chapter, I explored the relationship between master regulator, FOXM1, and DNMT 1, 3A, and 3B, in breast cancer cell lines basal subtype MDA-MB-231 and luminal A subtype MCF-7 cells. Briefly, I investigated protein-protein interactions using co-immunoprecipitation. Next, I investigated the role of FOXM1 on the regulation of DNMT1, 3A, and 3B expression via knockdown experiments. Lastly, I studied the differences between genomic binding profiles of FOXM1 and DNMTs in cancer cell lines, and the associated biological pathways using ChIP-seq analysis.

5.3 Results

5.3.1 FOXM1 and DNMT3 form protein complex in MCF-7 and MCF7-Epi^R breast cancer cell lines

Co-immunoprecipitation (co-IP) experiments were conducted to investigate the interactions between FOXM1 and DNMT3A/B at protein level in native MCF-7 cell line and epirubicin-resistant counterpart, MCF7-Epi^R cell line (Figure 5.1). Both cell lines were treated with 1 μM of epirubicin for either 6 hours, 18 hours, or left untreated before cells were harvested for co-IP. The amount of protein in the input controls were prepared as one-tenth fraction of the IP samples to avoid oversaturation of the protein samples. FOXM1 pulldown was performed with an anti-FOXM1 antibody in both MCF-7 and MCF7-Epi^R cell lines. A DNMT3A/B antibody was then used to probe for FOXM1 interaction with the DNMT3A/B on Western blot. A non-specific immunoglobulin (IgG) antibody was included in all experiments to serve as negative control. Figure 5.1 showed the interaction between FOXM1 and DNMT3A/B protein molecules. Additionally, FOXM1 and DNMT3A/B protein bands were absent in the IgG-immunoprecipitated samples, thereby minimised the possibility of a false positive.

On a separate note, as supported by Karunarathna *et al* in 2016 (191), the protein expression of FOXM1 decreased upon epirubicin treatment in MCF-7 cell line, whereas in the resistant counterpart it remained constant. This suggests that FOXM1 may be suppressed by epirubicin in native MCF-7 cell line but not in MCF7-Epi^R. In contrast, the endogenous level of DNMT3A/B was unaffected by epirubicin treatment.

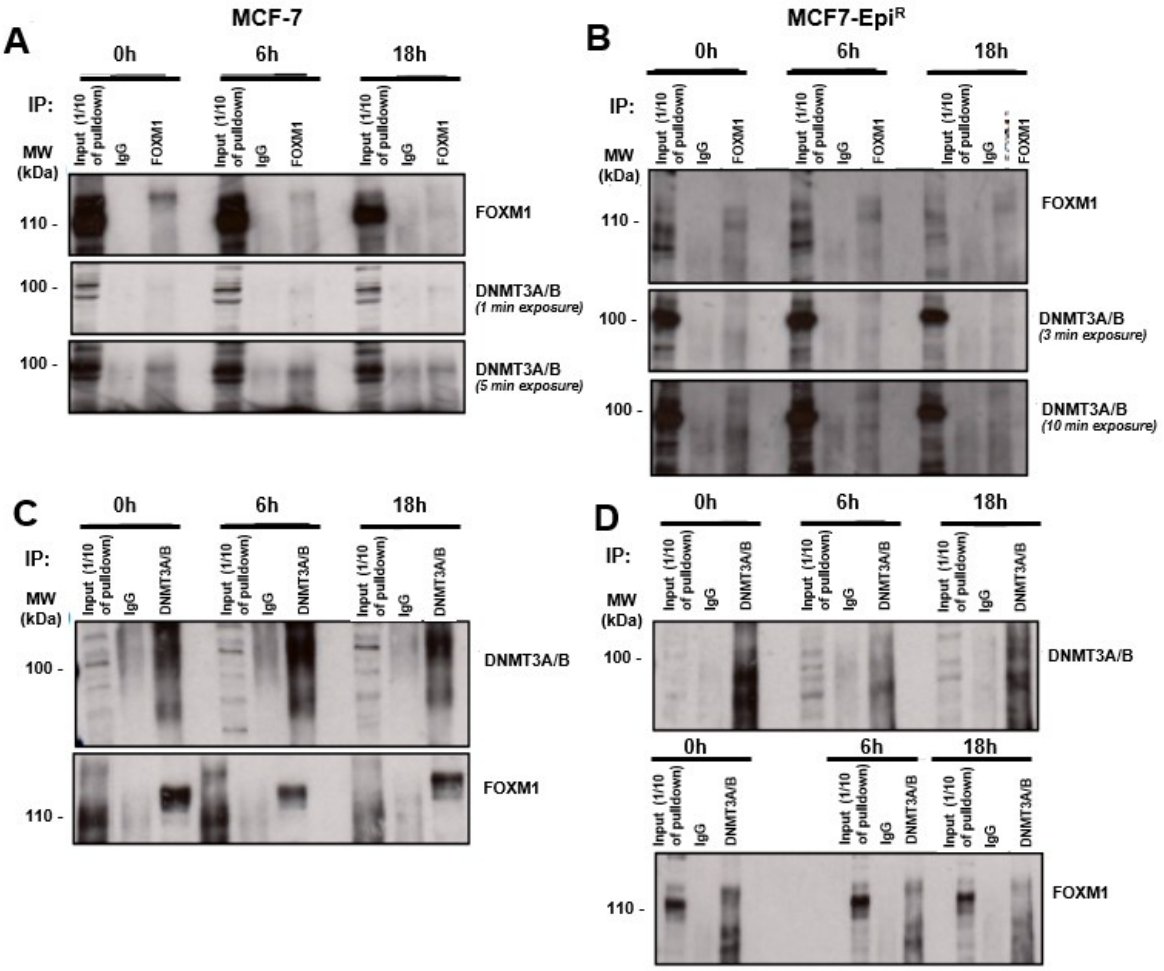


Figure 5.1 FOXM1 interacts with DNMT3A/B at the protein-protein level shown by co-immunoprecipitation (co-IP) in native MCF-7 and epirubicin-resistant counterpart MCF7-EpiR cell lines.

Both cell lines were either treated with 1 μM epirubicin for 6 hours, 18 hours, or left untreated as control. Co-IPs of FOXM1, DNMT3A/B and IgG were carried out using protein lysates extracted from MCF7 and MCF7-Epi^R with antibodies of anti-FOXM1, anti-DNMT3A/B or anti-IgG (negative control). The inputs represented the endogenous levels of proteins. The molecular weight (MW) is labelled in kDa **(a)** co-IP of FOXM1 and IgG in MCF7 and **(b)** in MCF7-Epi^R; Protein concentration of inputs were 1/10 fraction that of IP. **(c)** co-IP of DNMT3A/B and IgG in MCF-7 and **(d)** in MCF7-Epi^R; Inputs were 1/10 of IP **(d)** in MCF7-Epi^R. Figure shown is representative of one independent experiment.

5.3.2 FOXM1 does not interact with DNMT1 in both breast cancer MCF-7 and MCF7-Epi^R cell lines

Next, an identical co-IP experiment was conducted for FOXM1 and DNMT1 (Figure 5.2). FOXM1 was not detected visibly on DNMT1 pulldown IP, vice versa. This finding

suggests that FOXM1 and DNMT1 protein molecules do not form complex in both breast cancer cell lines.

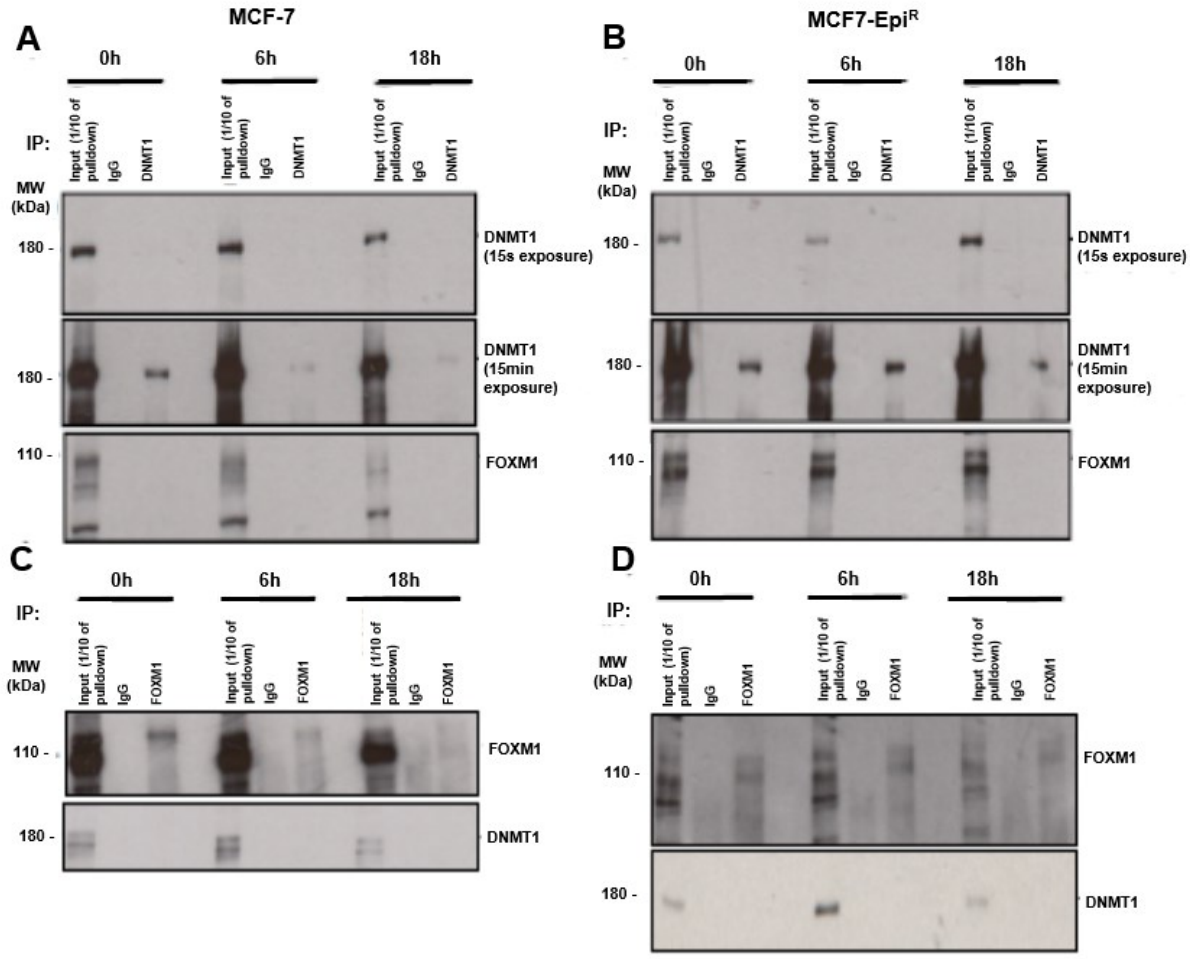


Figure 5.2 FOXM1 and DNMT1 does not interact at the protein level from co–IP experiments in MCF7 and MCF7-Epi^R cell lines.

Cells were harvested either upon 1 μM epirubicin treatment for 6 hours, 18 hours, or untreated as a negative control. Co – IPs using antibodies targeted at DNMT1 or FOXM1 were carried out using protein lysates extracted from MCF7 and MCF7-EpiR cell lines. Antibody targeted at IgG was also used as a negative control. The molecular weight (MW) of the protein was labelled in kDa. **(a)** co – IP of DNMT1 and IgG in MCF7 and **(b)** in MCF7-EpiR cell lines. **(c)** co – IP of FOXM1 and IgG in MCF7 and **(d)** in MCF7-EpiR cell lines. All inputs used were loaded at protein concentration of 1/10 fraction of IP protein lysates. Figure shown above is a representative result of one independent experiment.

5.3.3 Downregulation of FOXM1 does not affect the protein expression level of DNMTs in MCF-7 and MCF7-EpiR cell lines

The interaction between FOXM1 and DNMT3A/B is hypothesised to methylate target genes on the promoter regions, resulting in the silencing of target gene expressions. To investigate that, *FOXM1* was knockdown by siRNA and expressions of target molecules were observed on RT-qPCR (Figure 5.3). A non-specific short interfering RNA (siNSC) known to target no mammalian genes was used as a negative control. Cells were either treated for 18 hours or left untreated to observe the effect of epirubicin treatment. *FOXM1* was successfully depleted when compared against negative control in both MCF-7 and MCF7-EpiR cell lines. The protein expression of DNMT3A/B and DNMT1 were unaffected by *FOXM1* silencing. The protein expression of DNMT3A/B were not significantly different between MCF-7 and MCF7-EpiR cell lines. However, the expression of DNMT1 and DNMT3A/B were upregulated in MCF-7 cell line upon epirubicin treatment. These findings indicate that FOXM1 may interact with DNMT3A/B and regulate its methylation activity on target regions but does not regulate DNMT3A/B protein expression (Figure 5.1, Figure 5.3).

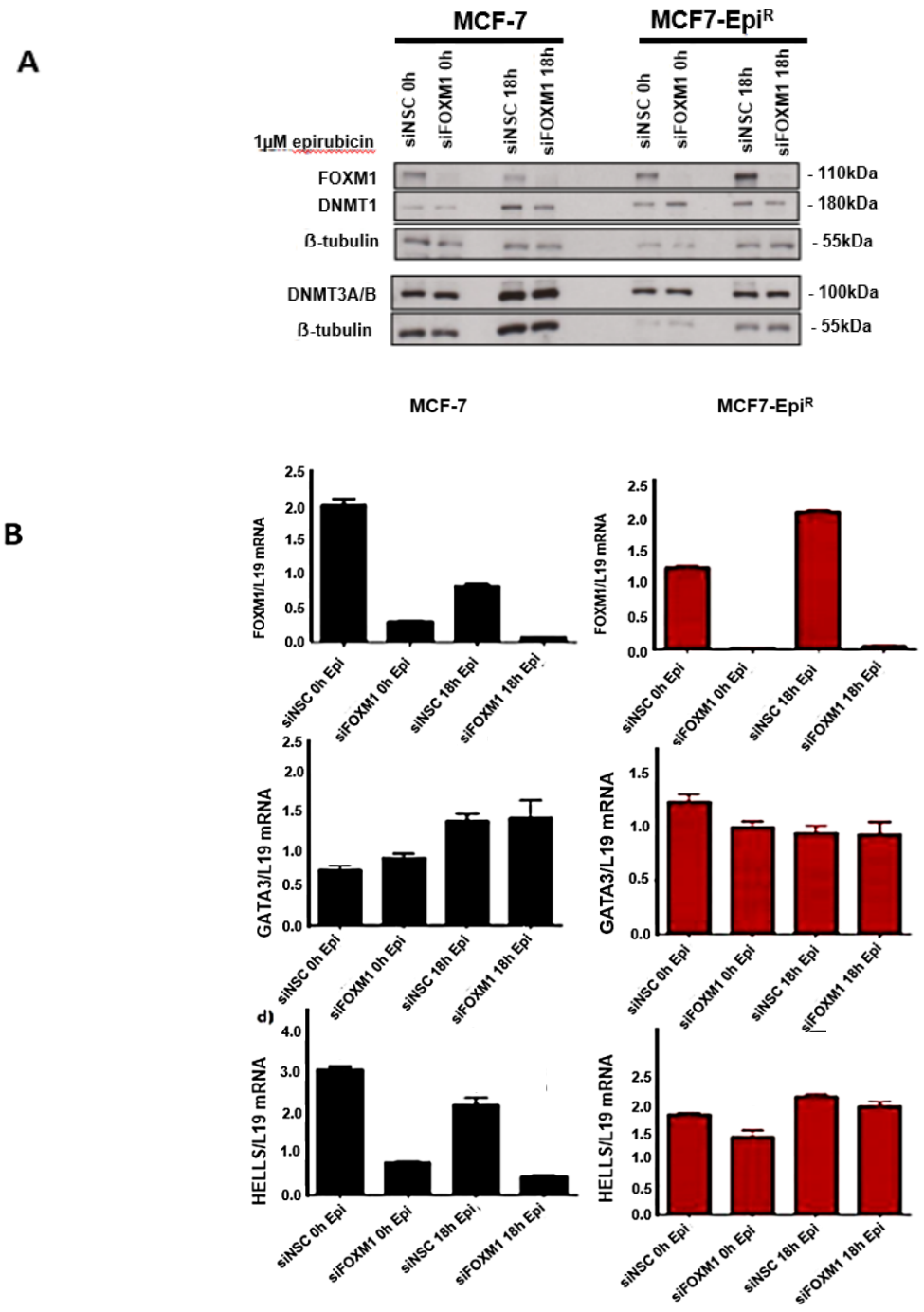


FIGURE 5.3 Expression analyses after downregulation of FOXM1 in MCF-7 and MCF7-Epi^R cell lines.

Cell lines were transiently transfected with siFOXM1 and siNSC (negative control). Cells were either treated with 1 µM epirubicin or left untreated as negative control. **(A) Protein expression analysis of FOXM1 knockdown in MCF-7 and MCF7-Epi^R cell lines.** Molecular weights are labelled in kDa. Western blot–SDS PAGE was carried out using extracted protein lysates from the mentioned cell lines. Protein expressions were observed for FOXM1, DNMT1, DNMT3A/B, and β-tubulin. β-tubulin protein expression was included as a loading control. The results represent two technical experiments. **(B) Gene expression analyses were performed by qRT-PCR after FOXM1 downregulation in MCF-7**

(left panel) and MCF7-Epi^R (right panel) cell lines. RNA was extracted and converted to cDNA from extracted cell lysates either treated for 18 hours of 1 μM epirubicin or left untreated as a control. L19 expression was measured for each sample and the respective average mean was used to normalise gene expressions of interest compared in MCF7 and MCF7-EpiR cell lines: **(a)** FOXM1 **(b)** GATA3 **(c)** HELLS. Three technical repeats were performed. As the data shown represent mean ± SEM of three technical repeats, no statistical analysis was performed.

5.3.4 The mRNA expression of *GATA3* and *HELLS* were impacted by the aberrant expression of FOXM1 in MCF-7 and MCF7-Epi^R cell lines

GATA3 and *HELLS* were recognised as the target genes of FOXM1 and DNMT3B from previous literatures (128, 192). As reported, expression of *GATA-3*, a tumour suppressor involved in differentiation of normal mammary luminal cells, was suppressed by promoter methylation via FOXM1-DNMT3B in a Rb dependent manner. FOXM1-DNMT3B interaction in MDA-MB-231 cell line led to *GATA3* methylation and subsequent inhibition of *GATA3* gene expression. *HELLS*, a chromatin remodelling/DNA helicase involved in DNA methylation, was identified as a downstream target of FOXM1 in head and neck squamous cell carcinomas (HNSCCs) (192). In primary human oral keratinocytes, *HELLS* serves as a downstream target of FOXM1 and interacts with DNMT1 and DNMT3B to inhibit p16^{INK4A}. Collectively, it was shown that FOXM1 acts though HELLS, DNMT1, and DNMT3B to inhibit *p16^{INK4A}* gene expression (193, 194).

The role of FOXM1 in transcriptional suppression via DNMTs was studied by investigating expression of these target genes in breast cancer cells, MCF-7 and MCF7-Epi^R cell lines via *FOXM1* knockdown experiments (Figure 5.3). RNAs were extracted from whole cell lysates and converted to cDNA, where gene expressions were analysed by qRT-PCR. *L19* housekeeping gene was included as a standard control to which all mRNA values were normalised against. Upon downregulation of *FOXM1* mRNA and protein levels in MCF-7 and MCF7-Epi^R cell lines, the mRNA expression of *GATA3* increased slightly upon *FOXM1* downregulation, however, this effect was not significant. On the other hand, mRNA expression level of *HELLS* was downregulated upon *FOXM1* depletion in MCF-7 cell line and decreased upon epirubicin treatment.

5.3.5 FOXM1 and DNMTs share similarities in transcriptional binding profiles

To assess the relationship between FOXM1 and DNMTs (1/3A/3B), I studied their genomic binding profiles from publicly available ChIP-seq profile of these factors. I focused on investigating their bindings on proximal promoter regions of target genes (Figure 5.4). Briefly, ChIP-seq profiles were downloaded from CistromeDB public database (161): FOXM1 in MCF-7 cell line and MDA-MB-231 cell line, DNMT1 in K562 cell line, and DNMT3A and DNMT3B in epidermal stem cell lines. ChIP-seq profiles of FOXM1 illustrated the genomic binding profiles on the proximal promoter regions of target genes (Figure 5.4). The genome-wide FOXM1 binding profiles in MCF-7 and MDA-MB-231 cell lines widely varied. Approximately 50% of FOXM1 binding in MCF-7 cell line (n = 12,404) overlapped with MDA-MB-231 cell line (n = 8,948). On the other hand, ChIP seq for DNMT1 in K562 cell line revealed roughly 33.3% of bindings within 3kb of TSS of target genes, among which 28% of the bindings were detected within 1kb of promoter regions. The other 26.4% of DNMT1 bindings were in the distal intergenic regions. Furthermore, the binding profile of DNMT3A in epidermal stem cell line was analysed. Roughly 27.3% of DNMT3A bindings seated within 3kb of TSS, and 23.4% of them were covered within 1kb of the promoter regions. Another 36% of bindings were in the intergenic regions. Lastly, DNMT3B has a small 19.3% of binding detected in the intergenic regions and another 43.6% of bindings were within 3kb of TSS, of which 39.3% were located within 1kb of the promoter regions.

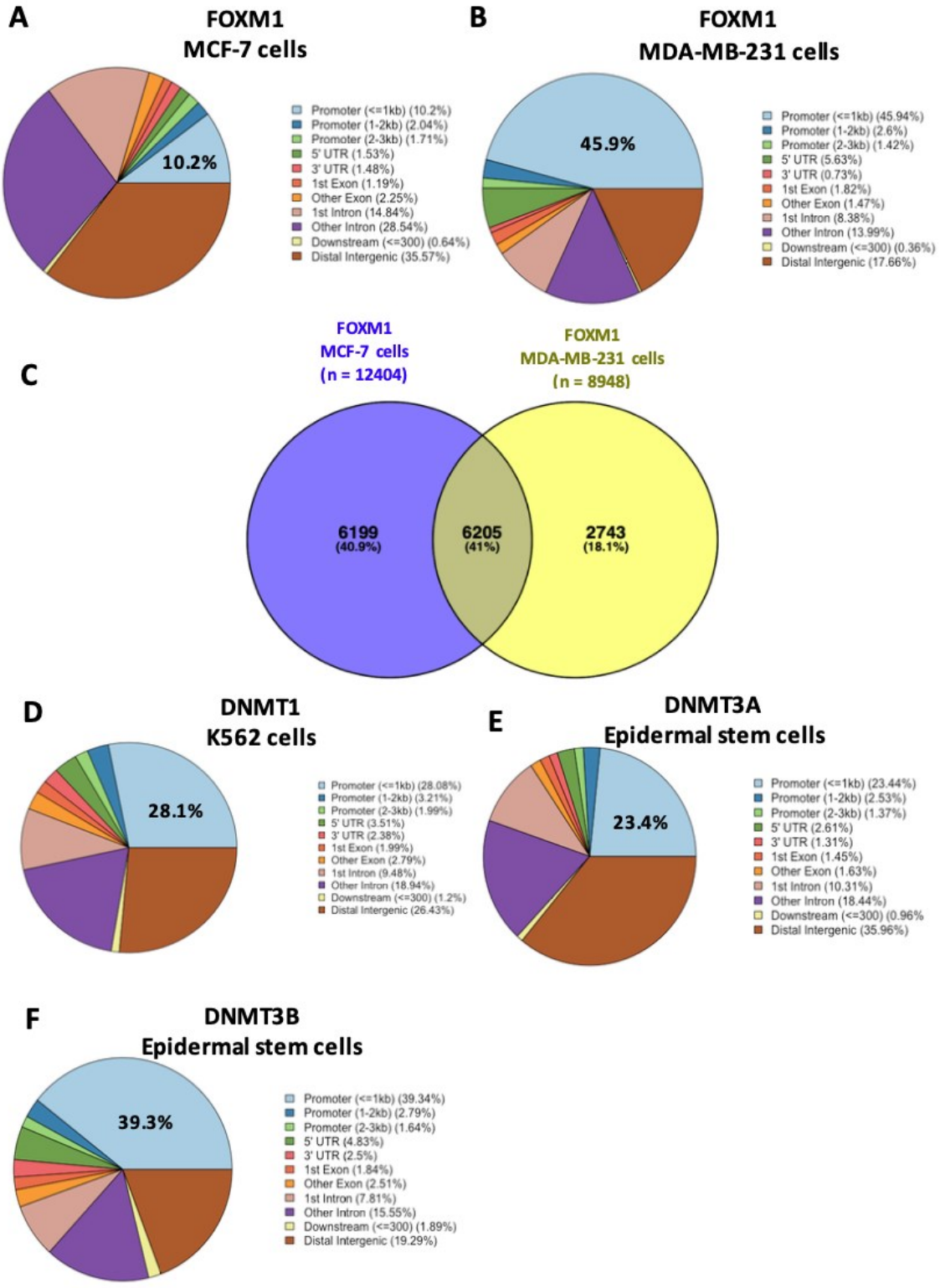


FIGURE 5.4 Genome-wide binding regions of FOXM1, DNMT1, DNMT3A, DNMT3B in cancer cell lines.

FOXM1 in **(A)** MCF-7 cells and **(B)** MDA-MB-231 cells. **(C)** Venn diagram shows the intersection between FOXM1 promoter binding peaks in MCF-7 cells (left) and MDA-MB-231 cells (right). n = the total number of genes with enriched peaks identified at promoter regions of known genes from ChIP-

seq analyses. Global distribution of genome binding regions of **(D)** DNMT1 in K562 cells, **(E)** DNMT3A and **(F)** DNMT3B, both in epidermal stem cells.

5.3.6 The known biological function of FOXM1 and DNMTs

Overall, it is known that aberrant methylations of the promoter region of tumour suppressor genes lead to gene silencing and de-methylation of oncogenes drive activation of these genes, thereby promoting tumorigenesis (96). It is known that DNMT1, DNMT3A, and DNMT3B have specific downstream target genes and regulates independent signalling pathways that abnormal promotions or suppressions could lead to tumorigenesis (93). However, cell-line specific and context-dependent differences in the regulation of DNMTs are not thoroughly understood. To characterise the biological functions of the candidate gene list, GO term analysis was conducted with R package 'ClusterProfiler'. The top 20 most significant biological categories were selected and compared (Figure 5.5). All three DNMTs were enriched with a role in ribonucleoprotein complex biogenesis. DNMT1 has a role in RNA splicing, chromatin modification, histone modification, and protein processing. Notably, DNMT1 has a role in the regulation of tumour suppressor p53 signalling pathway and cell cycle checkpoint. On the other hand, DNMT3A has less enriched GO terms, possibly due to lesser binding peaks detected in ChIP-seq analysis. DNMT3A has a role in the Wnt signalling pathway in ribonucleoprotein processing. Lastly, DNMT3B has additional roles in RNA splicing, mRNA, and protein catabolic processes, protein translocation and translation, and ER stress response.

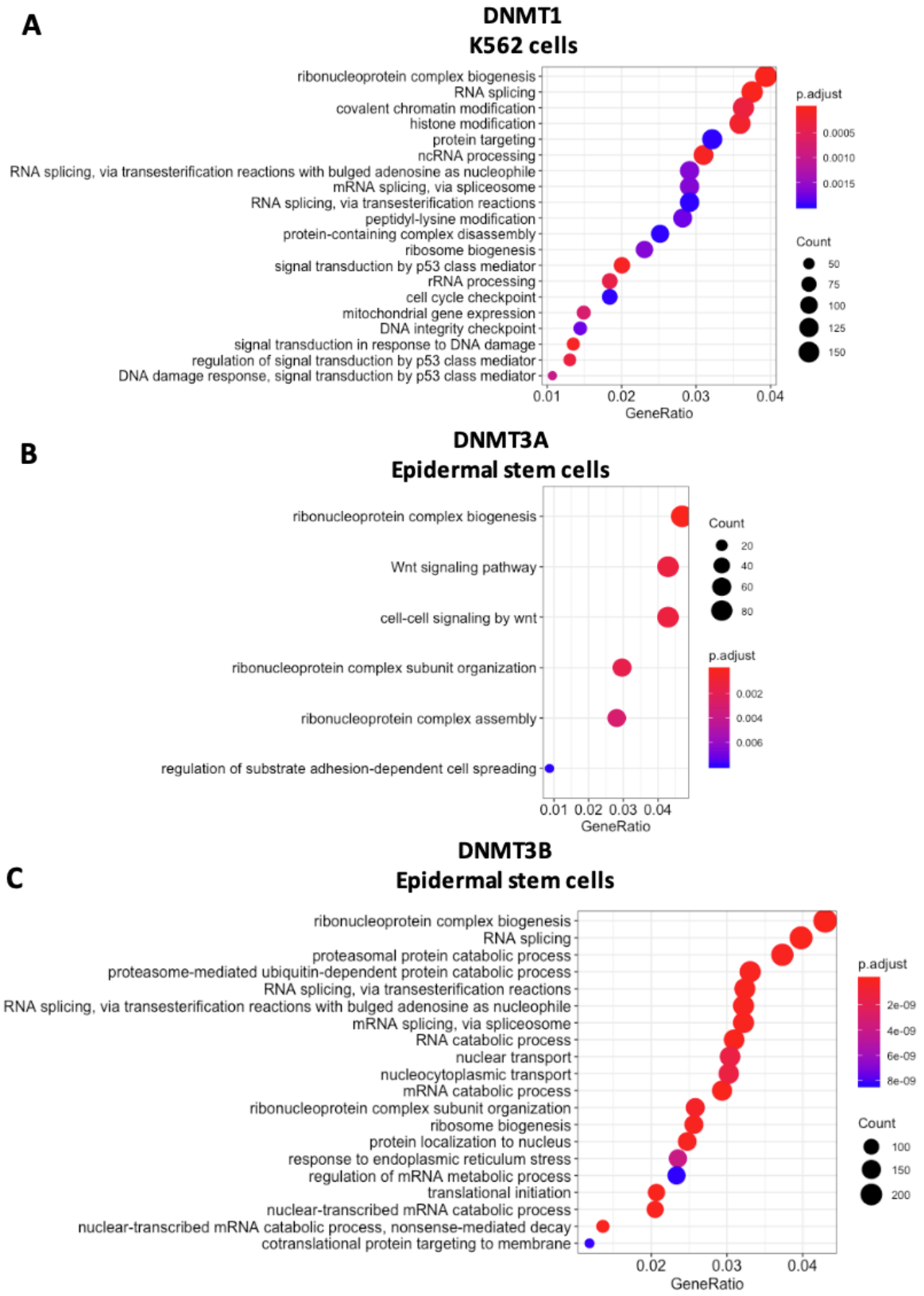


Figure 5.5 The transcriptional role of DNMTs on target genes.

5.3.7 The common binding targets and potential co-regulatory role of FOXM1 and DNMTs

FOXM1 (in MCF-7 cell line) shared 53.1% of identical targets with DNMT1 (in K562 cell line) within 1kb of promoter regions, whereas it was 34.0% with DNMT3A (in epidermal stem cell line), and a striking 60.5% with DNMT3B (in epidermal stem cell line) (Figure 5.6). This supports the hypothesised co-regulatory role of FOXM1 and DNMTs on the promoter regions though transcriptional modulation of their mutual downstream targets. FOXM1 is a known oncogenic TF that promotes tumorigenesis though transcriptional regulation of target gene expressions, whereas DNMTs methylate promoter regions of target tumour suppressors, lead to gene silencing and subsequent cancer initiation and progression. ChIP-seq peaks of the overlapped gene regions were identified within 1kb of transcription start site (TSS). This suggests that the co-regulatory role of FOXM1 and DNMTs mainly lies in the regulation of the target promoter regions and potential downstream function modulation.

GO analyses of the common binding targets between FOXM1 in MCF-7 cells and DNMTs all were enriched in RNA splicing and ribonucleoprotein complex biogenesis (Figure 5.5). Specifically, overlapped targets of FOXM1 in MCF-7 cells and DNMT1 in K562 cells were enriched in translation initiation and regulation of the ER stress signalling pathway. As for DNMT3A, GO terms were focused on mRNA processing and regulation of the apoptotic signalling pathway. Lastly, overlapped targets with DNMT3B in epidermal stem cells were involved in ER stress regulation (Figure 5.7).

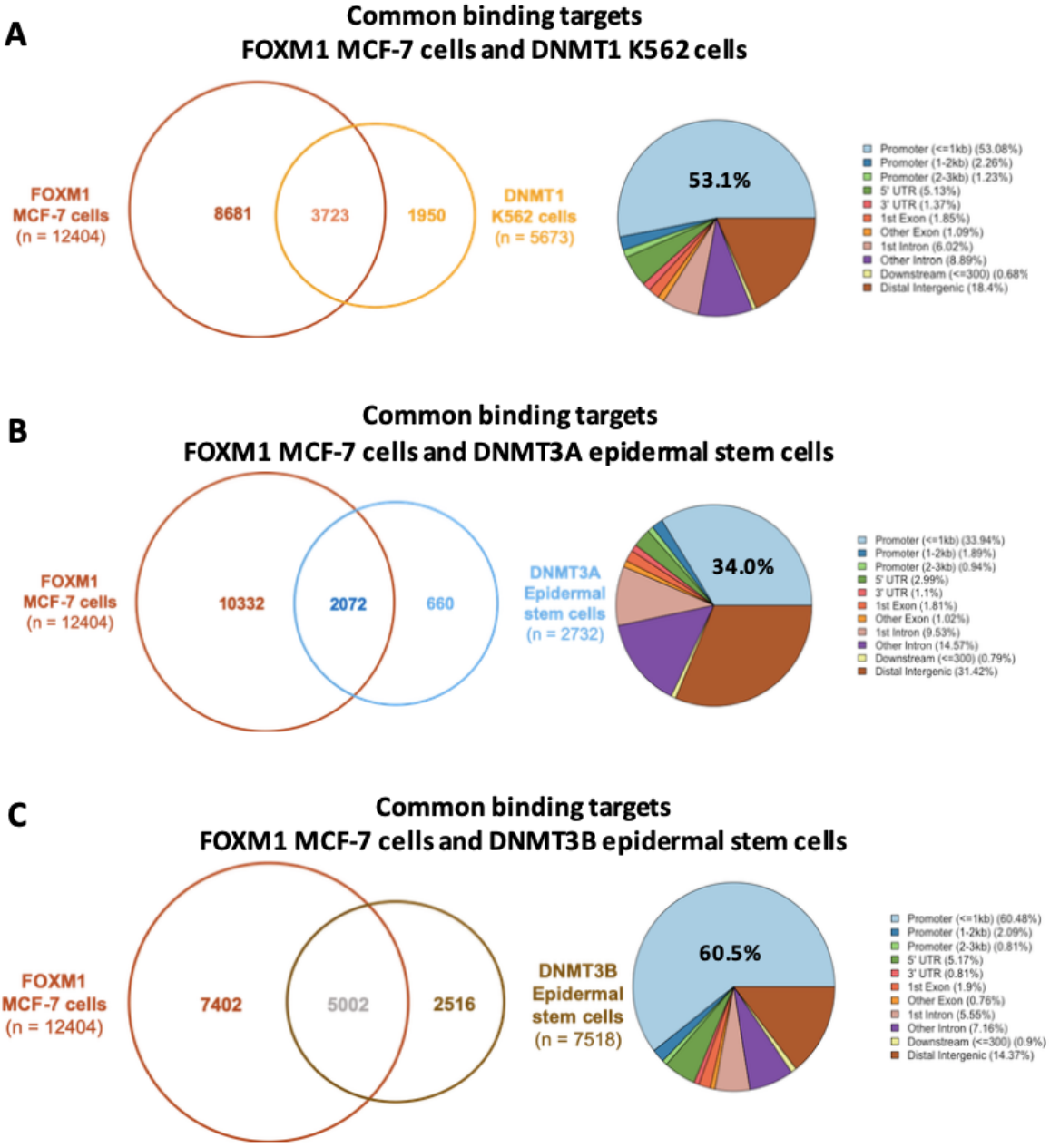


Figure 5.6 The potential co-regulatory role of common targets between FOXM1 in breast carcinoma MCF-7 cells and DNMTs.

The potential co-regulatory role of common targets between FOXM1 in breast carcinoma MCF-7 cells and DNMTs. (A) DNMT1, (B) DNMT3A and (C) DNMT3B.

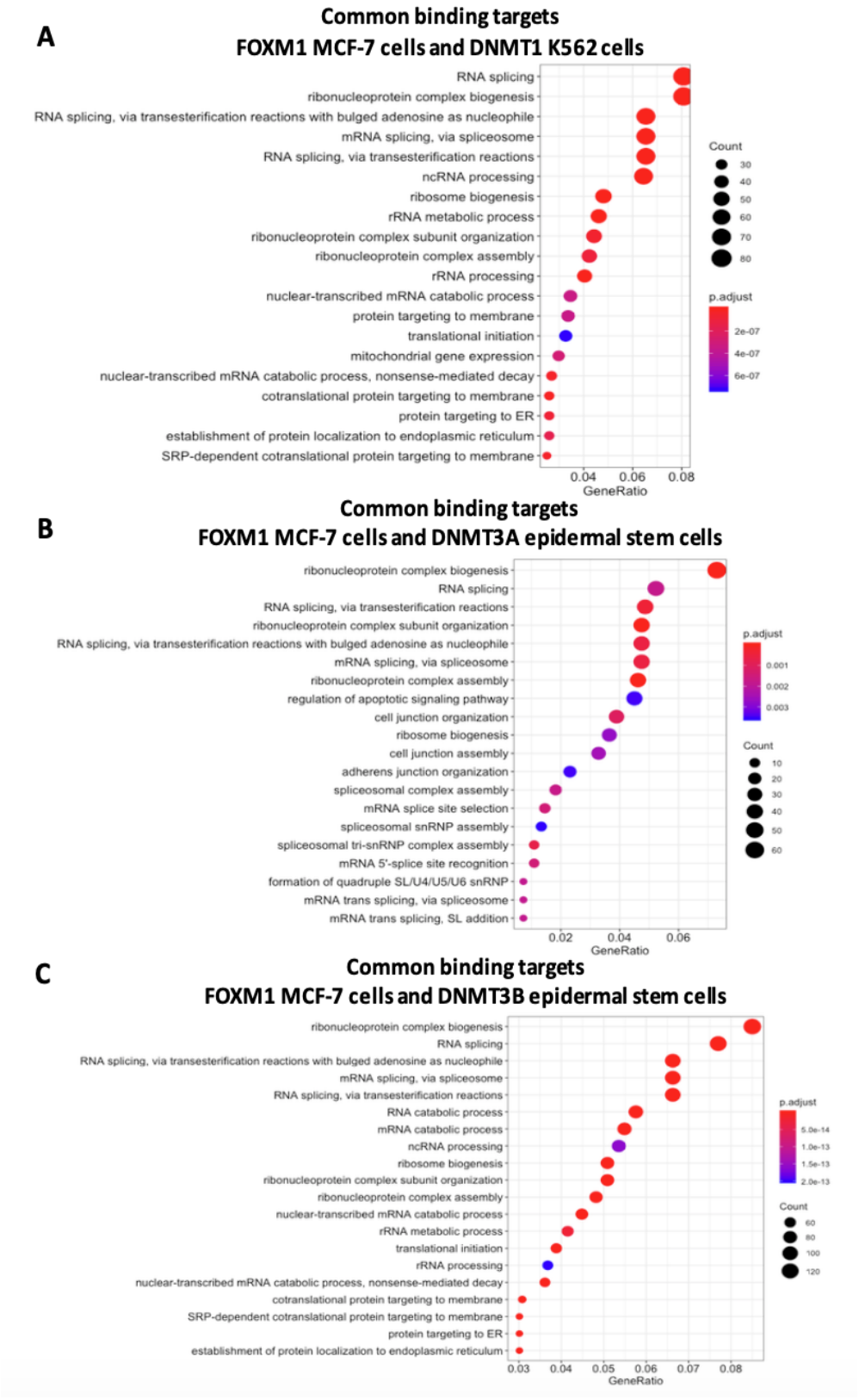


Figure 5.7 The GO analyses of the common binding targets between FOXM1 in MCF-7 cells and DNMTs.

GO analyses for (A) DNMT1 in K562 cells, (B) DNMT3A and (C) DNMT3B, both in epidermal stem cells.

5.3.8 FOXM1 in breast cancer basal subtype MDA-MB-231 cells and DNMTs have common binding targets and potential co-regulatory roles

The overlapped targets between FOXM1 in MDA-MB-231 cells and DNMT1, DNMT3A, and DNMT3B all had a striking majority of binding within 1kb of TSS (Figure 5.8). Overall, there were lesser overlapped targets between FOXM1 in MDA-MB-231 cells and DNMTs, possibly because FOXM1 has lesser identified binding targets in MDA-MB-231 cells. Fundamentally, GO analyses showed that the overlapped targets shared similar biological roles to common binding targets identified for FOXM1 in MCF-7 cells (Figure 5.9). This suggests that the co-regulatory role/interaction between FOXM1 and DNMTs is specific to each DNMTs but all conserved across both breast cancer subtypes MCF-7 cells and MDA-MB-231 cells. Finally, the common target regions between FOXM1 and DNMTs are all involved in ribonucleoprotein complex biogenesis, RNA splicing, and ER translocation, suggesting post-transcriptional roles for these common binding targets (Figure 5.9).

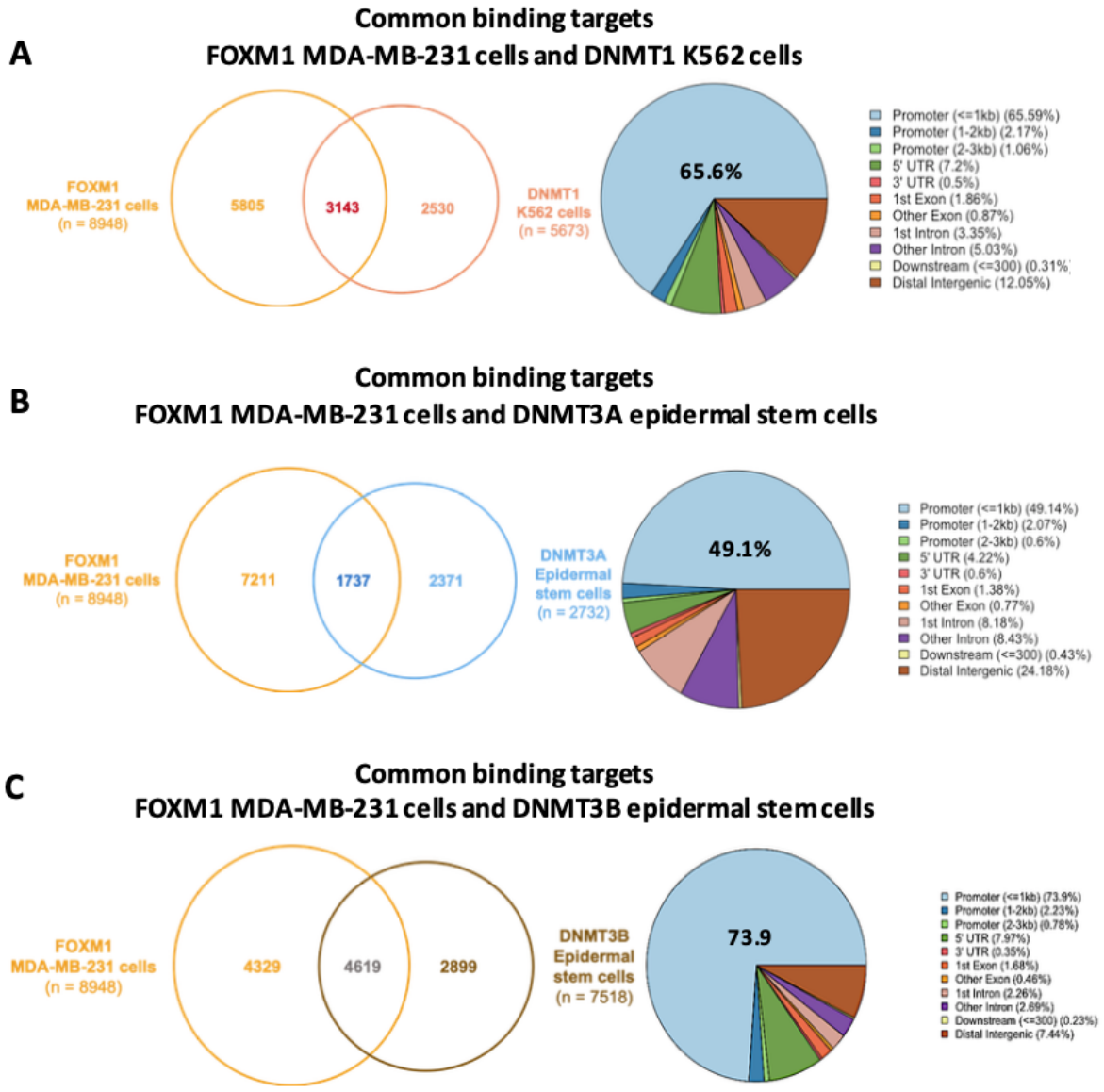


Figure 5.8 The potential co-regulatory role of common targets between FOXM1 in breast carcinoma MDA-MB-231 cells and DNMTs.

(A) DNMT1, (B) DNMT3A and (C) DNMT3B.

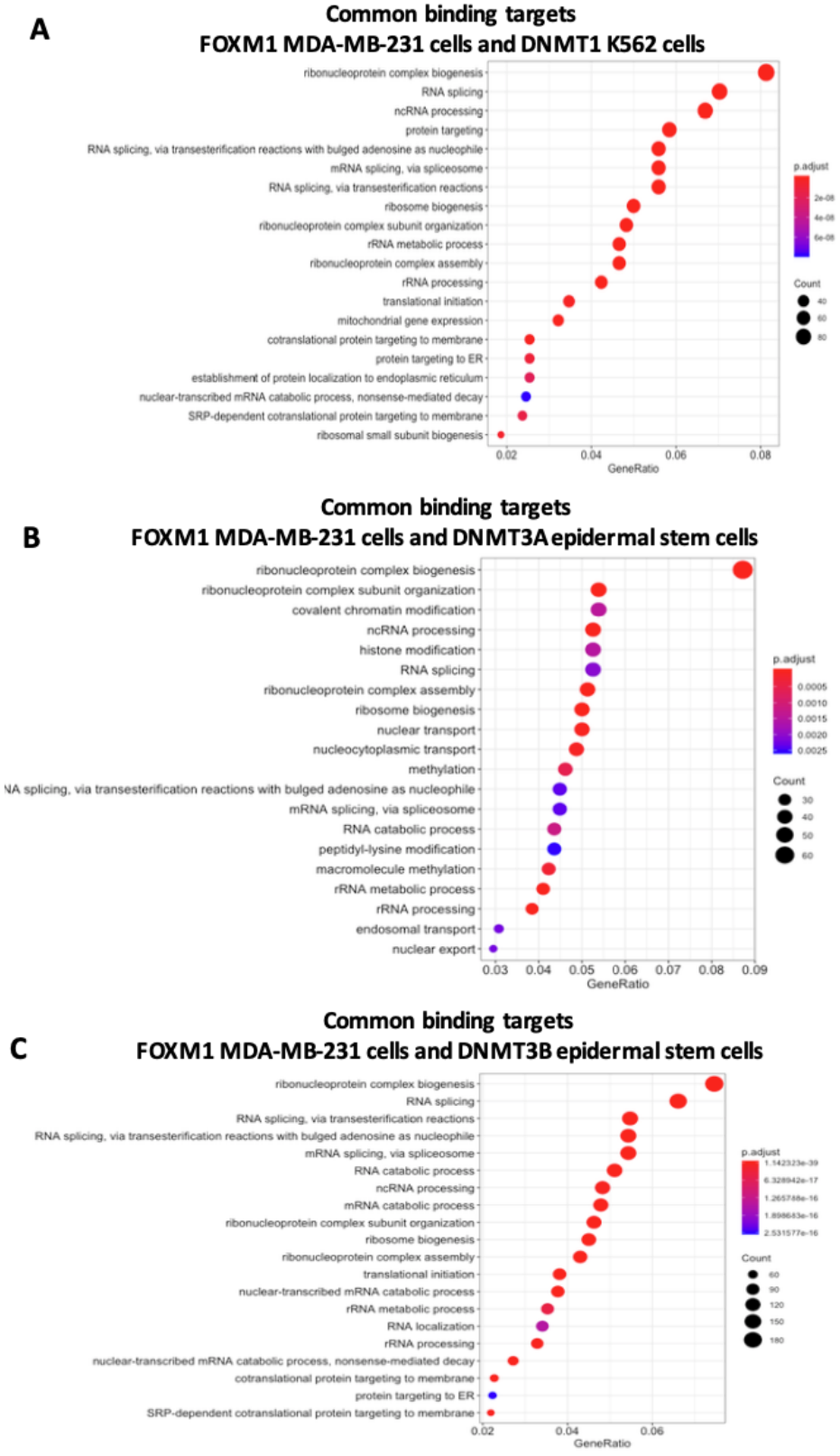


Figure 5.9 The GO analyses of the common binding targets between FOXM1 in MDA-MB-231 cells and DNMTs.

(A) DNMT1 in K562 cells, (B) DNMT3A and (C) DNMT3B, both in epidermal stem cells.

5.4 Discussion

In this chapter, I aim to demystify the relationship of FOXM1 and DNA methyltransferases, DNMT1, DNMT3A and DNMT3B in breast cancer subtype MCF-7 and MDA-MB-231 cells. Particularly, I am interested to understand the relevance of FOXM1 in epigenetic regulation of breast cancer and chemotherapy resistance.

Epigenetic states are frequently aberrant in cancers, representing one of the hallmarks of cancer (9). Abnormal DNA methylation could result in aberrant gene silencing, which is implicated in a large number of oncogenic conditions (195). The activation of oncogenes can be attributed to stabilisation of transcriptional binding on heterochromatin conformation after chromatin opening and CpG de-methylation on gene promoters, whereas promoter methylation by the recruitment of DNMTs leads to transcriptional suppression of genes (95, 196).

As previously stated, FOXM1 is a bona-fide oncogenic TF that drives tumorigenesis through transcriptional regulation of target gene expressions; on the other hand, DNMTs methylate and inhibit promoter regions of tumour suppressor genes and this occurrence leads to cancer initiation and progression.

I showed the existence of a FOXM1/DNMT3B protein complex in native MCF-7 cells and epirubicin-resistant MCF7-EpiR cells (Figure 5.1). This FOXM1 interaction was not seen with DNMT1 (Figure 5.2). The formation of a FOXM1 and DNMT3A/B complex was investigated at the protein level in breast cancer MCF-7 and MCF7-EpiR cell lines by co-immunoprecipitation analyses, indicating that the FOXM1-DNMT3A/B protein complex has a role in breast carcinoma MCF-7 cells irrespective of epirubicin resistance. This interaction may play a role in transcriptional repression activities driven by FOXM1 on regulating target gene expressions in MCF-7 and/or MCF7-EpiR cells. This role of FOXM1 in transcriptional inhibition, albeit new, is not unheard of (Figure 1.7). FOXM1 is known to regulate *HELLS* in primary human oral keratinocytes and head and neck squamous cell carcinomas, and FOXM1 regulation of *HELLS* leads to the interaction of HELLS with DNMT1 and DNMT3B to inhibit *p16INK4A* gene expression (192). *p16INK4A* is a known tumour suppressor gene, and aberrant inhibition of it leads to tumour development (193).

A similar FOXM1 regulation of *HELLS* was observed in MCF-7 cell line, suggesting a FOXM1 transcription suppression role in breast cancer cell line (Figure 5.3). The

transcriptional role of FOXM1 via DNMT3A/B on target genes in breast cancer cell lines should be explored. The interaction between FOXM1 and DNMT3B was previously observed in the MDA-MB-231 breast cancer cell line (128). Evidently, FOXM1 inhibits *GATA-3* expression in the mammary gland by interacting and recruiting DNMT3B to gene promoter region of *GATA-3*, thereby causing methylation-induced gene silencing (128). Subsequently, this inhibition of *GATA3* expression leads to less differentiated and lower grade tumours in mammary cells; in which implicates the development of highly aggressive breast cancer with negative clinical outcomes (98, 99).

GATA3 is known to be involved in mammary cell differentiation in breast cancer, whereas HELLS regulates chromatin structure and methylation activity via interaction with DNMT3B and DNMT1 (128, 193). Both target genes were selected as transcriptional readouts in this study as they were indicated to be regulated by DNMT3B and FOXM1 (192-194). GATA3 plays an important role in tissue development, but it is also upregulated in breast cancer and its expression was positively correlated with metastasis and mammary cell differentiation. The mRNA expression of *GATA3* did not increase significantly upon the transient depletion of *FOXM1* in MCF-7 cells but was induced by epirubicin (Figure 5.3). In MCF7-EpiR cells, no strong significance can be seen from either epirubicin treatment or FOXM1 downregulation. It is commonly believed that dysregulation of biological processes happens in drug-resistant MCF7-EpiR cells. DNMT3B was resistant to epirubicin treatment, irrespective of the presence of FOXM1 expression. *GATA3* may be suppressed by FOXM1 and DNMT3B but activated by epirubicin treatment in MCF-7 cells, further experiments should be conducted to explore this.

On the other hand, mRNA level of *HELLS* is activated by FOXM1 but downregulated by epirubicin treatment (Figure 5.3). SWI/SNF helicase (*HELLS*) was previously described to be positively regulated by FOXM1 (193). HELLS is involved in promoting DNA methylation and suppression of p16INK4A tumour suppressor protein via FOXM1 (192, 194). In the resistant cells, mRNA levels of *HELLS* were not drastically decreased upon FOXM1 downregulation, but there was a slight decrease before treatment. Again, the overall dysregulated mechanism in epirubicin-resistant breast carcinoma cells potentially led to the overexpression of *HELLS i*n the cells and therefore was not visibly affected by *FOXM1* silencing.

Repetitive administration of the same chemotherapeutic drugs may lead to the development of chemoresistance, which renders the respective chemotherapies ineffective. It was shown that DNA hypermethylation is the first epigenetic market associated with the development of drug resistance (197). Additionally, epirubicin, a cytotoxic drug that disrupts nucleic acid and protein synthesis, has the potential to induce DNA methylation. Interestingly, previous studies have shown that the effectiveness of doxorubicin was enhanced when DNMT3B was depleted in breast cancer MDA-MB-453, BT549, and Hs578T cells, implicating a role of DNMT3B in the regulation of therapeutic effects of anthracyclines (198). Epirubicin led to the upregulation of DNMT3A/B and silenced the expression of downstream proteins that are required to maintain drug toxicity in tumour cells.

In my study, this protein overexpression of DNMT3A/B was not seen in MCF-7 cells, which suggests either an alternative regulation in MCF-7 cells, the increase of DNMT3A/B methylation rate, or increase in mRNA expression was not reflected on western blotting (Figure 5.3). Additionally, the expression level of molecules is cyclical in cells, which does not represent a long-lasting representation of the actual state of cells (94).

ChIP analyses have shown that pioneer factor, FOXA1 recruits DNMT1 and DNMT3A/B to its promoter region in breast cancer MCF-7, MDA-MB-231, MDA-MB-436, and SUM1, MO2 cells (97). Inhibition of FOXA1 re-activated the methylation status and lead to subsequent transcriptional suppression of tumour suppressor genes, *DAPK*, *MGMT*, *RASSF1A*, *TP53*, thereby promotes tumorigenesis and apoptotic evasion (158). Downregulation of FOXA1 upregulated the mRNA levels of *DNMT1*, *DNMT3A*, and *DNMT3B*. FOXM1 and FOXA1 target the same gene sequences and thus, have highly homogenous binding profiles, sharing similar downstream target genes (50). Therefore, it would be of further interest to investigate whether FOXM1 can recruit DNMT3A/B in MCF-7 and MCF7-EpiR cells to promoter regions. Indeed, ChIP-seq data of FOXM1 in MDA-MB-231 cells have identified a FOXM1 binding on the *DNMT3B* promoter region (Figure 5.10). This relationship between FOXM1 and DNMTs are likely contextual and vary with each type of DNMT, as was supported by the overlap of binding peaks of DNMT1/3A/3B and FOXM1, suggesting that there are common and specific binding targets to each DNMT (Figure 5.6). Future experiments

should focus on ChIP experiments to validate FOXM1 binding on *DNMT3B* gene promoter region in breast cancer cell lines.

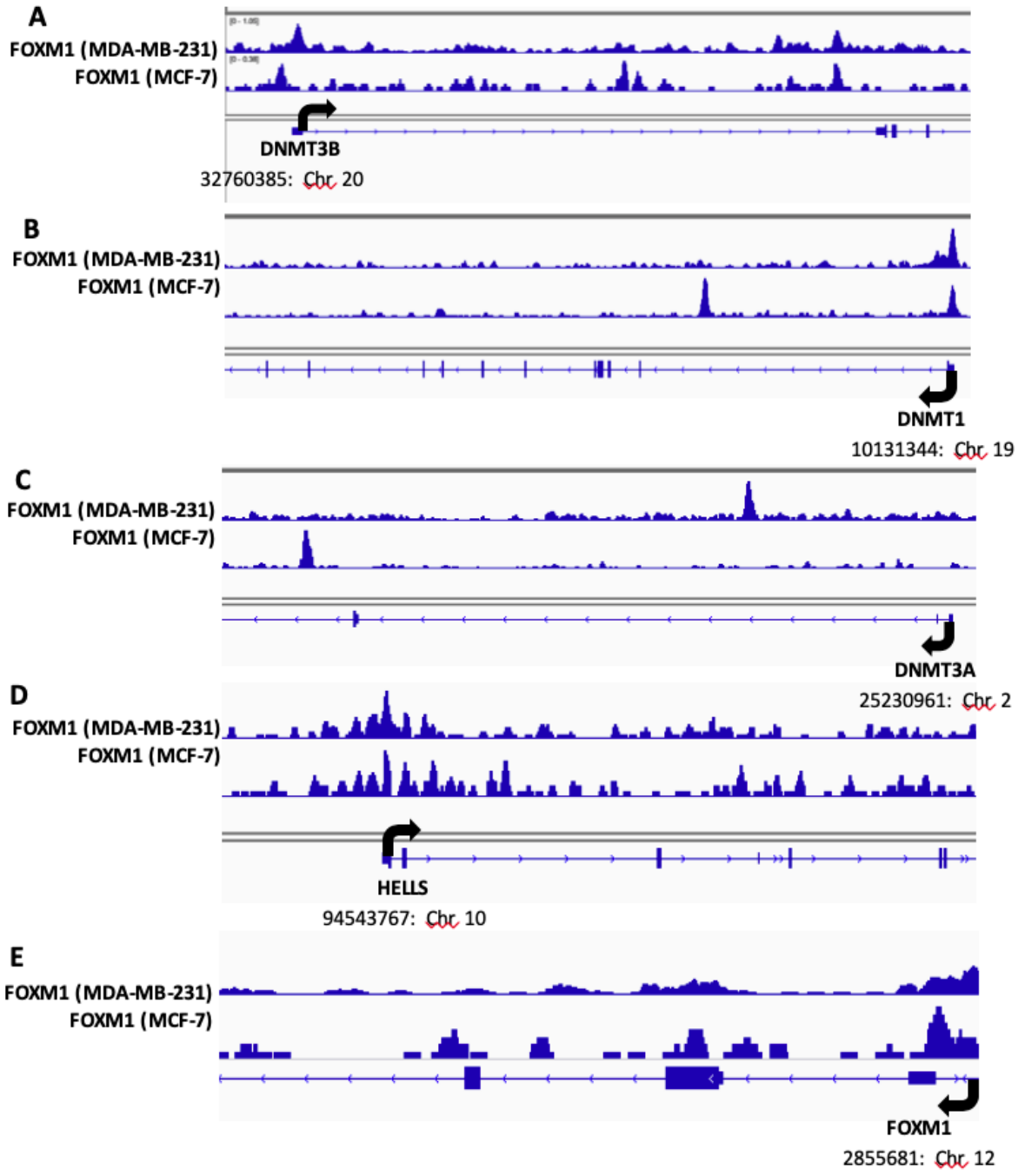


Figure 5.10 FOXM1 differentially binds on DNMT1, 3A and 3B.

No FOXM1 binding was detected on the promoter regions of all three DNMTs in MCF-7 cells, but FOXM1 was identified to bind on DNMT3B promoter region in MDA-MB-231 cells

Particularly, ChIP-seq experiments coupled with FOXM1 and/or DNMT3A/B knockdown to investigate expression level and profile changes in target genes would provide clarity on the epigenetic role of FOXM1.

The overlapped of binding peaks between FOXM1 in MCF-7 cells and MDA-MB-231 cells suggests that there are approximately 50% of FOXM1 binding peaks in MCF-7 cells which overlapped with that in MDA-MB-231 cells (Figure 5.4). As shown in **Chapter 3**, FOXM1 has common targets that are conserved across cancer cell lines and specific targets that are unique to breast cancer MCF-7 cell line. Moreover, mammalian DNMTs barely overlapped in their binding targets (4). Interestingly, majority of common binding targets between FOXM1 and DNMT1/3A/3B were detected within 1kb of the gene promoter region, this is most prominent for common targets between FOXM1 and DNMT3B in MDA-MB-231 cells (Figure 5.8).

The top 20 GO categories indicate that FOXM1 in MDA-MB-231 cells and DNMT3B has a role in ER stress (Figure 5.9). In **Chapter 4**, I showed that ER stress and UPR has a significant role in breast cancer and drug resistance (74, 171), and increased DNMT3B-mediated CpG island methylation on promoter region was detected upon stimulation by oxidative stress and ER stress in diabetic pregnancy (199). This unravels a new possibility of FOXM1-DNMT3B mediated transcriptional silencing in ER stress molecules.

Further investigation is required to explore the overexpression of FOXM1 and DNTM3A/B; alongside the role of FOXM1-DNMT3B axis in the suppression of tumour suppressor genes in breast cancer cell lines, leading to breast cancer development and progression. This may provide a novel mean to improve breast cancer treatment that lacks target specificity and reduce the risk of chemotherapy resistance. Epigenetic chemosensitisation has been known to improve the efficacy of chemotherapeutics against tumours with methylation defects, and chemotherapy resistance can be overcome with demethylation treatment in certain cases (99). This epigenetic axis may represent a reversible target for better breast cancer chemotherapy treatment. Further understanding of this relationship between FOXM1 and DNMT3B in breast cancer cell line holds crucial implications for the identification of useful targets to improve breast cancer research; and potentially lower therapeutic resistance and prevent disease recurrence.

6 Final Discussion

6.1 Overview

Breast cancer cells are capable to continue evolving and gaining additional evasive pathways for growth and cell survival in the presence of monotherapy treatment of STIs, and therefore the efficacy of these inhibitors is also limited by rapidly acquired resistance. This thesis provides the potential of co-targeting other adaptive pathways, along with ESR1 signalling antagonism (Figure 1.2).

Over the last decades, studies have shown that the overexpression of FOXM1 is common in various types of cancer, particularly in breast cancer. Upregulated FOXM1 expression has been reported in breast cancer cell lines with acquired therapy resistance, whereas FOXM1 knockdown re-sensitised these cells to therapies (56). The expression and prognostic value of FOXM1 in breast cancer cell lines were largely correlated to expressions of ESR1/PR and molecular subtypes, which indicated the potential of FOXM1 as a master regulator gene with critical functions in breast cancer (56).

I have demonstrated the multiple roles of FOXM1 under various oncogenic biological contexts in breast cancer cell line models. Particularly, I investigated FOXM1 cistromes and interactome in driving distinct chemotherapeutic, chromatin accessibility, ER stress, and epigenetic relevant contexts in breast cancer subtypes. These roles are differentiated across breast cancer MCF-7 and MDA-MB-231 cell lines, though certain fundamental roles are conserved across cell lines. These differentiations are dictated by the interaction with specific co-factors, resulting in oncogenic responses. Noteworthy, FOXM1 is a crucial TF for fundamental biological processes in normal cells; therefore, targeting specific downstream targets or co-factors seemed to be a feasible and sustainable approach rather than targeting FOXM1 itself to avoid interfering with normal biological processes (**Introduction Section 1.6.3 FOXM1**).

Collectively, these preliminary new insights of FOXM1 and potential co-factors in breast cancer MCF-7 and MDA-MB-231 breast cancer cell lines require further validation. These findings are nowhere near conclusive and functional; and clinical studies will be indispensable before the proposal of promising targets for better breast cancer prognosis.

6.2 Key limitations

6.2.1 Cancer cell lines model

Cancer cell lines have been widely utilised in research as a useful *in vitro* model to study deregulated genes and signalling pathways involved in development of cancer. The research results obtained from cell lines from drug testing and translational study are usually extrapolated to *in vivo* human tumours before committing to large scale and costly clinical trials. Different cancer cell lines have varied deregulated genes and dysfunctional signalling pathways, therefore display diverse responses to external triggers (200). It is not feasible and not recommended to draw parallels between two cancer cell lines, let alone of different cancer types. Importantly, the homogeneous cell culture environment is different from that of the original heterogeneous tumour microenvironment. Due to the resource limitation during my research, I conducted all analyses with available breast cancer cell lines and publicly accessible sequencing files. The appropriate cell lines and further *in vivo* experiments must be completed in future before crystallising any insights.

This study was conducted in breast cancer cell lines and their resistant variants that are exposed to prolonged treatment of anti-estrogens or absence of estrogen. Such cancer cell line models acknowledged mode of actions that can activate therapy resistance *in vitro* rather than those that indeed regulate resistance in breast cancer patients. The main limitation of this includes the extent to which the few ESR1-positive breast cancer cell lines represent the range of ESR1-positive phenotypes *in situ* and the negligence of epithelial-stromal and tumour-host interactions that likely mediate sensitivity *in vivo*. In short, the role of tumour environment is overlooked in cell line models and this may rule out the realistic uncertainties in patient samples. Additionally, the specificity of endocrine resistance to different anti-estrogens remains unclear.

6.2.2 Real-time qPCR

Real time quantitative PCR has remained the most popular method in molecular diagnostics, such as serving as the gold standard for the diagnosis of COVID-19, despite the development of more modern approaches. The lasting success is largely attributed to four key benefits: sensitivity, simplicity, speed, and cost. It is a relatively cost-effective and quick way to study specific genomic targets at tiny concentrations,

and the results are easy to interpret. However, it is not designed to study complex or multi-gene conditions. Secondly, there is a potential for biased outcomes, since it requires pre-selection of targets, so the outcome of assay is restricted to the specific probe selected for the assay. Thirdly, there are many factors that can affect the mRNA expression level detected by RT-qPCR, such as the stability of the mRNA or RNase present in the experiment. Therefore, a more accurate measure for transcriptional factor binding and promoter activity would be ChIP and promoter assays.

6.2.3 Western blotting

Western blotting is another technique frequently performed throughout my study. It is rather subjective technique, false-positives may happen when the used antibody reacts with a non-intended protein whereas false-negatives may develop when insufficient time were given to transfer gel to the membrane. Improper blotting, and processing procedures and durations often produce skewed, faded, or even repetitive bands, making results subjected to the interpretation of the researcher. Additionally, traditional Western blots are non-quantitative, though kits and software are available for band quantification. Molecular weight of a protein can only be estimated with a protein ladder in Western blots, rather than measured precisely as with mass spectrometry. Overall, the expression of a molecule, either mRNA or protein, is not a less accurate representation of cancer status and prognostic outcome than interactome and cistrome network of molecules of interest. These expression profiles are ‘snapshot’ taken at a timepoint when cells were harvested, rather than a longer lasting profile characteristic such as genetic and epigenetic network and interactome.

6.2.4 Co-immunoprecipitation

Co-immunoprecipitation is a convenient and cost-effective method to detect the specific protein-protein interactions in any molecular condition using the right antibodies. However, there are several disadvantages. Firstly, weaker signals from low affinity protein are not detected. Secondly, protein interactions which takes place within short period will not be detected. Importantly, this technique detects direct and indirect protein interactions and cannot distinguish them separately. There could be a chance that they are present in a protein complex, hence adding uncertainty to the understanding of target molecular mechanism. Direct protein interactions could be

explored with real affinity purification-mass spectrometry or yeast two-hybrid assay (201).

6.2.5 ChIP

ChIP was performed in majority of my study to observe the relationship between TFs and specific promoter regions of interest. Unfortunately, ChIP is an experimental technique with several limitations (202). The binding peaks do not represent binding strength or level of binding events – it is simply an indication that by using an antibody specific to our protein of interest and primers targeted to certain gene promoter regions, a signal is detected on certain sheared DNA fragments. The binding peaks can vary according to experimental procedures, physiological conditions, treatment conditions, and technical errors. Therefore, a linear comparison of protein occupancy is not feasible and not encouraged between MCF-7 and MCF7-Epi^{R} cell line. These comparisons are inevitably error-prone, and replicates can only confirm presence of these putative bindings. Furthermore, ChIP measurements do not provide any information about direct or indirect protein binding on these genomic regions, other analyses are required to determine the exact molecular mechanism. Moreover, several caveats exist in ChIP technique, the number of harvested cells and concentration of antibodies could be increased in future experiments for a cleaner (less background) and more significant result. Furthermore, a negative primer control designed to bind none of the known FOXM1 target promoter regions should be included to increase the experimental confidence.

6.2.6 Experimental replicates

Arguably the biggest limitation of my study was the lack of biological replicates, and sometimes technical replicates, for majority of the presented experiments. Both types of replicate measurements are key to producing accurate, reliable, and meaningful results and statistics to address different questions on data reproducibility. Biological replicates are parallel measurements of biologically distinct samples that capture random biological variations, which may itself be a subject of study or a background noise (203). Technical replicates are repeated measurements of the same sample that represent independent measures of the random background noises associated with experimental protocols, processes, materials and reagents, or equipment (203).

For biologically distinct conditions, getting the mean of technical replicates can minimise the impact of measurement errors, but getting additional biological replicates is often the preferable way for improving the efficiency of statistical analyses. For genome sequencing, base calls are made from several DNA reads of the same genetic locus. These reads are technical replicates that factor the uncertainty in the sequencer readout but will not expose errors in the library itself. Errors in library construction can be limited by building technical replicate libraries from the same sample. If additional resources are available, one could potentially return to the source tissue and prepare multiple samples or aliquots to repeat the whole sequencing workflow. Such replicates would be technical if the samples or aliquots are from the same source tissue or biological if obtained from different aliquots of biologically distinct materials (204). Owing to historically high costs per sequencing assays, it is not common for the field of genome sequencing to perform replicates. As the call for accuracy increases and the cost of sequencing decreases, this is expected to change. If technical replicate measurements are misinterpreted as biological variations, false positives would occur. If biological measurements have unusually small error bars, they may merely reveal measurement error, not biological variation. Additionally, technical replicates do not yield statistically independent measures, these are known as pseudo-replicates and not representative of the whole population of interest.

To summarise, both types of replicates are essential to ensure the robustness of experimental designs and meaningful results. Biological replicates are preferable to technical replicates for inference about the mean and variance of a population of interest. However, the cost-benefit trade-off of collecting additional replicates should be weighed, where biological replicates are usually more costly than the technical ones. For future references, it is noted that experimental best practice includes planning for replication. A good blocking practice suggests that samples from different biological sources should be mixed across sequencing runs or experiments to restrict the effect of any systematic run-to-run variability. As our scale for higher-throughput assays advances, it is critical to make sure efficient study designs and valid statistical analyses stay.

6.3 Overall future directions

Due to the severe limitations in resources and lack of supervisions in my study, the protein-DNA interactions were investigated using publicly available ChIP-sequencing and microarray experiments performed and kindly shared on open database by other research groups. The original source of all sequencing files are cited in **Section 2.11.2.1 Cistrome Data Browser (Cistrome DB) Data Mining**. Using these public sequencing files, I conducted further bioinformatics analyses to validate my hypotheses on the identification of new FOXM1 downstream targets or interactions with potential co-factors. As a result, I proposed these preliminary insights and inferred their relevance in breast cancer cell lines using my best knowledge on the data. However, it is by no means that these insights are conclusive. The priority for anyone to validate these preliminary insights, is to repeat all sequencing experiments in this thesis in targetted breast cancer cell lines and later in breast cancer tissue samples to ensure fair comparisons between TFs and the relevance in breast cancer patients.

Furthermore, the qPCR and western blotting experiments conducted are mostly three technical replicates and lacking of biological repeats (as stated in figure legends). It is apparent that technical replicates are not applicable to draw experimental confidence as they could represent outliers and therefore cannot be representative of the conditions tested. It is compulsory to repeat all of these experiments using three biological replicates to validate these molecular conditions before drawing any conclusions.

Not forgetting, cell line models provide only a snapshot of molecular characteristics in *in vitro* environment. In future, the application of integrative tactics that investigate gene expression in the context of global genomic, proteomic or functional data to understand breast cancer mechanism and therapy resistance will require next generation high throughput techniques to allow genomic scale screens, large-scale proteomics and parallel transcriptome, epigenome and genome sequencing of both experimental cell lines and clear-annotated clinical samples. This holistic ‘systems biology’ platform will evaluate cellular responses and signalling pathways as a whole rather than ‘screenshots’ of cellular contexts.

7 References

1. Observatory TGC. Breast fact sheets. 2020.
2. Yang G, Nowsheen S, Aziz K, Georgakilas AG. Toxicity and adverse effects of Tamoxifen and other anti-estrogen drugs. Pharmacol Ther. 2013;139(3):392-404.
3. Byler S, Goldgar S, Heerboth S, Leary M, Housman G, Moulton K, et al. Genetic and epigenetic aspects of breast cancer progression and therapy. Anticancer Res. 2014;34(3):1071-7.
4. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell. 2006;10(6):515-27.
5. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. Mol Oncol. 2011;5(1):5-23.
6. Dai X, Xiang L, Li T, Bai Z. Cancer Hallmarks, Biomarkers and Breast Cancer Molecular Subtypes. J Cancer. 2016;7(10):1281-94.
7. Bertos NR, Park M. Breast cancer - one term, many entities? J Clin Invest. 2011;121(10):3789-96.
8. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res. 2015;5(10):2929-43.
9. Stefansson OA, Moran S, Gomez A, Sayols S, Arribas-Jorba C, Sandoval J, et al. A DNA methylation-based definition of biologically distinct breast cancer subtypes. Mol Oncol. 2015;9(3):555-68.
10. Hua H, Zhang H, Kong Q, Jiang Y. Mechanisms for estrogen receptor expression in human cancer. Experimental Hematology & Oncology. 2018;7(1).
11. Nichols M, Rientjes JM, Stewart AF. Different positioning of the ligand-binding domain helix 12 and the F domain of the estrogen receptor accounts for functional differences between agonists and antagonists. EMBO J. 1998;17(3):765-73.
12. Osborne CK, Schiff R. Mechanisms of endocrine resistance in breast cancer. Annu Rev Med. 2011;62:233-47.
13. Gutierrez MC, Detre S, Johnston S, Mohsin SK, Shou J, Allred DC, et al. Molecular changes in tamoxifen-resistant breast cancer: relationship between estrogen receptor, HER-2, and p38 mitogen-activated protein kinase. J Clin Oncol. 2005;23(11):2469-76.
14. Herynk MH, Fuqua SA. Estrogen receptor mutations in human disease. Endocr Rev. 2004;25(6):869-98.
15. Jeselsohn R, De Angelis C, Brown M, Schiff R. The Evolving Role of the Estrogen Receptor Mutations in Endocrine Therapy-Resistant Breast Cancer. Curr Oncol Rep. 2017;19(5):35.
16. Pejerrey SM, Dustin D, Kim JA, Gu G, Rechoum Y, Fuqua SAW. The Impact of ESR1 Mutations on the Treatment of Metastatic Breast Cancer. Horm Cancer. 2018;9(4):215-28.
17. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. Nat Genet. 2013;45(12):1439-45.
18. Fanning SW, Mayne CG, Dharmarajan V, Carlson KE, Martin TA, Novick SJ, et al. Estrogen receptor alpha somatic mutations Y537S and D538G confer breast cancer endocrine resistance by stabilizing the activating function-2 binding conformation. Elife. 2016;5.
19. Beaver JA, Park BH. The BOLERO-2 trial: the addition of everolimus to exemestane in the treatment of postmenopausal hormone receptor-positive advanced breast cancer. Future Oncol. 2012;8(6):651-7.
20. Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F, Gonzalez-Angulo AM, et al. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. Clin Cancer Res. 2014;20(7):1757-67.
21. Green KA, Carroll JS. Oestrogen-receptor-mediated transcription and the influence of co-factors and chromatin state. Nat Rev Cancer. 2007;7(9):713-22.

22. Gelsomino L, Gu G, Rechoum Y, Beyer AR, Pejerrey SM, Tsimelzon A, et al. ESR1 mutations affect anti-proliferative responses to tamoxifen through enhanced cross-talk with IGF signaling. Breast Cancer Res Treat. 2016;157(2):253-65.
23. Musgrove EA, Sutherland RL. Biological determinants of endocrine resistance in breast cancer. Nat Rev Cancer. 2009;9(9):631-43.
24. McCarthy N. Genes, genes everywhere. Nat Rev Cancer. 2012;12(8):507.
25. Butt AJ, McNeil CM, Musgrove EA, Sutherland RL. Downstream targets of growth factor and oestrogen signalling and endocrine resistance: the potential roles of c-Myc, cyclin D1 and cyclin E. Endocr Relat Cancer. 2005;12 Suppl 1:S47-59.
26. Safe S, Kim K. Non-classical genomic estrogen receptor (ER)/specificity protein and ER/activating protein-1 signaling pathways. J Mol Endocrinol. 2008;41(5):263-75.
27. Osborne CK, Bardou V, Hopp TA, Chamness GC, Hilsenbeck SG, Fuqua SA, et al. Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. J Natl Cancer Inst. 2003;95(5):353-61.
28. Shou J, Massarweh S, Osborne CK, Wakeling AE, Ali S, Weiss H, et al. Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer. J Natl Cancer Inst. 2004;96(12):926-35.
29. Chakraborty AK, Welsh A, Digiovanna MP. Co-targeting the insulin-like growth factor I receptor enhances growth-inhibitory and pro-apoptotic effects of anti-estrogens in human breast cancer cell lines. Breast Cancer Res Treat. 2010;120(2):327-35.
30. Arpino G, Wiechmann L, Osborne CK, Schiff R. Crosstalk between the estrogen receptor and the HER tyrosine kinase receptor family: molecular mechanism and clinical implications for endocrine therapy resistance. Endocr Rev. 2008;29(2):217-33.
31. Giuliano M, Trivedi MV, Schiff R. Bidirectional Crosstalk between the Estrogen Receptor and Human Epidermal Growth Factor Receptor 2 Signaling Pathways in Breast Cancer: Molecular Basis and Clinical Implications. Breast Care (Basel). 2013;8(4):256-62.
32. Massarweh S, Osborne CK, Creighton CJ, Qin L, Tsimelzon A, Huang S, et al. Tamoxifen resistance in breast tumors is driven by growth factor receptor signaling with repression of classic estrogen receptor genomic function. Cancer Res. 2008;68(3):826-33.
33. Welsh AW, Lannin DR, Young GS, Sherman ME, Figueroa JD, Henry NL, et al. Cytoplasmic estrogen receptor in breast cancer. Clin Cancer Res. 2012;18(1):118-26.
34. Wang B, Jiang H, Ma N, Wang Y. Phosphorylated-p38 mitogen-activated protein kinase expression is associated with clinical factors in invasive breast cancer. Springerplus. 2016;5(1):934.
35. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. Breast Cancer Res. 2011;13(6):227.
36. Katzenellenbogen BS, Frasor J. Therapeutic targeting in the estrogen receptor hormonal pathway. Semin Oncol. 2004;31(1 Suppl 3):28-38.
37. Howell A, Osborne CK, Morris C, Wakeling AE. ICI 182,780 (Faslodex): development of a novel, "pure" antiestrogen. Cancer. 2000;89(4):817-25.
38. Early Breast Cancer Trialists' Collaborative G. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;365(9472):1687-717.
39. Chavez KJ, Garimella SV, Lipkowitz S. Triple negative breast cancer cell lines: one tool in the search for better treatment of triple negative breast cancer. Breast Dis. 2010;32(1-2):35-48.
40. Masoud V, Pages G. Targeted therapies in breast cancer: New challenges to fight against resistance. World J Clin Oncol. 2017;8(2):120-34.
41. Weigel D, Jurgens G, Kuttner F, Seifert E, Jackle H. The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. Cell. 1989;57(4):645-58.
42. Fyodorov DV, Zhou B-R, Skoultchi AI, Bai Y. Emerging roles of linker histones in regulating chromatin structure and function. Nature Reviews Molecular Cell Biology. 2018;19(3):192-206.

43. Benayoun BA, Caburet S, Veitia RA. Forkhead transcription factors: key players in health and disease. Trends in Genetics. 2011;27(6):224-32.
44. Clark KL, Halay ED, Lai E, Burley SK. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature. 1993;364(6436):412-20.
45. Monteiro LJ, Khongkow P, Kongsema M, Morris JR, Man C, Weekes D, et al. The Forkhead Box M1 protein regulates BRIP1 expression and DNA damage repair in epirubicin treatment. Oncogene. 2013;32(39):4634-45.
46. Wierstra I. The transcription factor FOXM1 (Forkhead box M1): proliferation-specific expression, transcription factor function, target genes, mouse models, and normal biological roles. Adv Cancer Res. 2013;118:97-398.
47. Chen X, Muller GA, Quaas M, Fischer M, Han N, Stutchbury B, et al. The forkhead transcription factor FOXM1 controls cell cycle-dependent gene expression through an atypical chromatin binding mechanism. Mol Cell Biol. 2013;33(2):227-36.
48. Vladimir V. Kalinichenko CSB, Tanya V. Kalin. Transcription Factors: Forkhead Box (FOX). Reference Module in Biomedical Sciences. 2020.
49. Sanders DA, Gormally MV, Marsico G, Beraldi D, Tannahill D, Balasubramanian S. FOXM1 binds directly to non-consensus sequences in the human genome. Genome Biol. 2015;16:130.
50. Golson ML, Kaestner KH. Fox transcription factors: from development to disease. Development. 2016;143(24):4558-70.
51. Littler DR, Alvarez-Fernandez M, Stein A, Hibbert RG, Heidebrecht T, Aloy P, et al. Structure of the FoxM1 DNA-recognition domain bound to a promoter sequence. Nucleic Acids Res. 2010;38(13):4527-38.
52. Lam EW, Brosens JJ, Gomes AR, Koo CY. Forkhead box proteins: tuning forks for transcriptional harmony. Nat Rev Cancer. 2013;13(7):482-95.
53. Wang K, Guan C, Fang C, Jin X, Yu J, Zhang Y, et al. Clinical significance and prognostic value of Forkhead box A1 expression in human epithelial ovarian cancer. Oncol Lett. 2018;15(4):4457-62.
54. Lee CS, Friedman JR, Fulmer JT, Kaestner KH. The initiation of liver development is dependent on Foxa transcription factors. Nature. 2005;435(7044):944-7.
55. Kalin TV, Ustiyan V, Kalinichenko VV. Multiple faces of FoxM1 transcription factor: lessons from transgenic mouse models. Cell Cycle. 2011;10(3):396-405.
56. Liao GB, Li XZ, Zeng S, Liu C, Yang SM, Yang L, et al. Regulation of the master regulator FOXM1 in cancer. Cell Commun Signal. 2018;16(1):57.
57. Li L, Wu D, Yu Q, Li L, Wu P. Prognostic value of FOXM1 in solid tumors: a systematic review and meta-analysis. Oncotarget. 2017;8(19):32298-308.
58. Kalinichenko VV, Major ML, Wang X, Petrovic V, Kuechle J, Yoder HM, et al. Foxm1b transcription factor is essential for development of hepatocellular carcinomas and is negatively regulated by the p19ARF tumor suppressor. Genes Dev. 2004;18(7):830-50.
59. Yau C, Wang Y, Zhang Y, Foekens JA, Benz CC. Young age, increased tumor proliferation and FOXM1 expression predict early metastatic relapse only for endocrine-dependent breast cancers. Breast Cancer Res Treat. 2011;126(3):803-10.
60. Nho RS, Hergert P. FoxO3a and disease progression. World J Biol Chem. 2014;5(3):346-54.
61. Eijkelenboom A, Burgering BM. FOXOs: signalling integrators for homeostasis maintenance. Nat Rev Mol Cell Biol. 2013;14(2):83-97.
62. Paik JH, Kollipara R, Chu G, Ji H, Xiao Y, Ding Z, et al. FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. Cell. 2007;128(2):309-23.
63. Alasiri G, Fan LY, Zona S, Goldsbrough IG, Ke HL, Auner HW, et al. ER stress and cancer: The FOXO forkhead transcription factor link. Mol Cell Endocrinol. 2018;462(Pt B):67-81.
64. Tzivion G, Dobson M, Ramakrishnan G. FoxO transcription factors; Regulation by AKT and 14-3-3 proteins. Biochim Biophys Acta. 2011;1813(11):1938-45.

65. Osei-Sarfo K, Gudas LJ. Retinoids induce antagonism between FOXO3A and FOXM1 transcription factors in human oral squamous cell carcinoma (OSCC) cells. PLoS One. 2019;14(4):e0215234.
66. Fumagalli D, Wilson TR, Salgado R, Lu X, Yu J, O'Brien C, et al. Somatic mutation, copy number and transcriptomic profiles of primary and matched metastatic estrogen receptor-positive breast cancers. Ann Oncol. 2016;27(10):1860-6.
67. Sanders DA, Ross-Innes CS, Beraldi D, Carroll JS, Balasubramanian S. Genome-wide mapping of FOXM1 binding reveals co-binding with estrogen receptor alpha in breast cancer cells. Genome Biol. 2013;14(1):R6.
68. Zou Y, Tsai W-B, Cheng C-J, Hsu C, Chung YM, Li P-C, et al. Forkhead box transcription factor FOXO3a suppresses estrogen-dependent breast cancer cell proliferation and tumorigenesis. Breast Cancer Research. 2008;10(1):R21.
69. Wang Y, Ung MH, Xia T, Cheng W, Cheng C. Cancer cell line specific co-factors modulate the FOXM1 cistrome. Oncotarget. 2017;8(44):76498-515.
70. Park YY, Jung SY, Jennings NB, Rodriguez-Aguayo C, Peng G, Lee SR, et al. FOXM1 mediates Dox resistance in breast cancer by enhancing DNA repair. Carcinogenesis. 2012;33(10):1843-53.
71. Halasi M, Gartel AL. A novel mode of FoxM1 regulation: positive auto-regulatory loop. Cell Cycle. 2009;8(12):1966-7.
72. Yang N, Wang C, Wang Z, Zona S, Lin SX, Wang X, et al. FOXM1 recruits nuclear Aurora kinase A to participate in a positive feedback loop essential for the self-renewal of breast cancer stem cells. Oncogene. 2017;36(24):3428-40.
73. Wang M, Kaufman RJ. The impact of the endoplasmic reticulum protein-folding environment on cancer development. Nat Rev Cancer. 2014;14(9):581-97.
74. McGrath EP, Logue SE, Mnich K, Deegan S, Jager R, Gorman AM, et al. The Unfolded Protein Response in Breast Cancer. Cancers (Basel). 2018;10(10).
75. Yang Y, Liu L, Naik I, Braunstein Z, Zhong J, Ren B. Transcription Factor C/EBP Homologous Protein in Health and Diseases. Front Immunol. 2017;8:1612.
76. Oyadomari S, Mori M. Roles of CHOP/GADD153 in endoplasmic reticulum stress. Cell Death Differ. 2004;11(4):381-9.
77. Nishitoh H. CHOP is a multifunctional transcription factor in the ER stress response. J Biochem. 2012;151(3):217-9.
78. Morishima N, Nakanishi K, Nakano A. Activating transcription factor-6 (ATF6) mediates apoptosis with reduction of myeloid cell leukemia sequence 1 (Mcl-1) protein via induction of WW domain binding protein 1. J Biol Chem. 2011;286(40):35227-35.
79. Ohoka N, Yoshii S, Hattori T, Onozaki K, Hayashi H. TRB3, a novel ER stress-inducible gene, is induced via ATF4-CHOP pathway and is involved in cell death. EMBO J. 2005;24(6):1243-55.
80. Ghosh AP, Klocke BJ, Ballestas ME, Roth KA. CHOP potentially co-operates with FOXO3a in neuronal cells to regulate PUMA and BIM expression in response to ER stress. PLoS One. 2012;7(6):e39586.
81. Wang X-Z. Identification of novel stress-induced genes downstream of chop. The EMBO Journal. 1998;17(13):3619-30.
82. Zou XJ, Yang L, Yao SL. Endoplasmic reticulum stress and C/EBP homologous protein-induced Bax translocation are involved in angiotensin II-induced apoptosis in cultured neonatal rat cardiomyocytes. Exp Biol Med (Maywood). 2012;237(11):1341-9.
83. Bruchmann A, Roller C, Walther TV, Schafer G, Lehmusvaara S, Visakorpi T, et al. Bcl-2 associated athanogene 5 (Bag5) is overexpressed in prostate cancer and inhibits ER-stress induced apoptosis. BMC Cancer. 2013;13:96.
84. Dewson G, Kluck RM. Mechanisms by which Bak and Bax permeabilise mitochondria during apoptosis. J Cell Sci. 2009;122(Pt 16):2801-8.
85. Jauhiainen A, Thomsen C, Strombom L, Grundevik P, Andersson C, Danielsson A, et al. Distinct cytoplasmic and nuclear functions of the stress induced protein DDIT3/CHOP/GADD153. PLoS One. 2012;7(4):e33208.

86. Iurlaro R, Muñoz-Pinedo C. Cell death induced by endoplasmic reticulum stress. The FEBS Journal. 2016;283(14):2640-52.
87. Zheng YZ, Cao ZG, Hu X, Shao ZM. The endoplasmic reticulum stress markers GRP78 and CHOP predict disease-free survival and responsiveness to chemotherapy in breast cancer. Breast Cancer Res Treat. 2014;145(2):349-58.
88. Balasubramanian MN, Butterworth EA, Kilberg MS. Asparagine synthetase: regulation by cell stress and involvement in tumor biology. Am J Physiol Endocrinol Metab. 2013;304(8):E789-99.
89. Wang J, Wang L, Ho CT, Zhang K, Liu Q, Zhao H. Garcinol from Garcinia indica Downregulates Cancer Stem-like Cell Biomarker ALDH1A1 in Nonsmall Cell Lung Cancer A549 Cells through DDIT3 Activation. J Agric Food Chem. 2017;65(18):3675-83.
90. Mihailidou C, Chatzistamou I, Papavassiliou AG, Kiaris H. Improvement of chemotherapeutic drug efficacy by endoplasmic reticulum stress. Endocr Relat Cancer. 2015;22(2):229-38.
91. Ye H, Chen M, Cao F, Huang H, Zhan R, Zheng X. Chloroquine, an autophagy inhibitor, potentiates the radiosensitivity of glioma initiating cells by inhibiting autophagy and activating apoptosis. BMC Neurol. 2016;16(1):178.
92. Kaminska K, Nalejska E, Kubiak M, Wojtysiak J, Zolna L, Kowalewski J, et al. Prognostic and Predictive Epigenetic Biomarkers in Oncology. Mol Diagn Ther. 2019;23(1):83-95.
93. Subramaniam D, Thombre R, Dhar A, Anant S. DNA methyltransferases: a novel target for prevention and therapy. Front Oncol. 2014;4:80.
94. Pouliot MC, Labrie Y, Diorio C, Durocher F. The Role of Methylation in Breast Cancer Susceptibility and Treatment. Anticancer Res. 2015;35(9):4569-74.
95. Parrella P. The value of epigenetic biomarkers in breast cancer. Biomark Med. 2018;12(9):937-40.
96. Oh TG, Wang SM, Muscat GE. Therapeutic Implications of Epigenetic Signaling in Breast Cancer. Endocrinology. 2017;158(3):431-47.
97. Mirza S, Sharma G, Parshad R, Gupta SD, Pandya P, Ralhan R. Expression of DNA methyltransferases in breast cancer patients and to analyze the effect of natural compounds on DNA methyltransferases and associated proteins. J Breast Cancer. 2013;16(1):23-31.
98. Mathe A, Wong-Brown M, Locke WJ, Stirzaker C, Braye SG, Forbes JF, et al. DNA methylation profile of triple negative breast cancer-specific genes comparing lymph node positive patients to lymph node negative patients. Sci Rep. 2016;6:33435.
99. Roll JD, Rivenbark AG, Jones WD, Coleman WB. DNMT3b overexpression contributes to a hypermethylator phenotype in human breast cancer cell lines. Mol Cancer. 2008;7:15.
100. Laisné M, Gupta N, Kirsh O, Pradhan S, Defossez P-A. Mechanisms of DNA Methyltransferase Recruitment in Mammals. Genes. 2018;9(12):617.
101. Vijayaraghavalu S, Dermawan JK, Cheriyath V, Labhasetwar V. Highly synergistic effect of sequential treatment with epigenetic and anticancer drugs to overcome drug resistance in breast cancer cells is mediated via activation of p21 gene expression leading to G2/M cycle arrest. Mol Pharm. 2013;10(1):337-52.
102. Cheishvili D, Chik F, Li CC, Bhattacharya B, Suderman M, Arakelian A, et al. Synergistic effects of combined DNA methyltransferase inhibition and MBD2 depletion on breast cancer cells; MBD2 depletion blocks 5-aza-2'-deoxycytidine-triggered invasiveness. Carcinogenesis. 2014;35(11):2436-46.
103. Lee AV, Oesterreich S, Davidson NE. MCF-7 cells--changing the course of breast cancer research and care for 45 years. J Natl Cancer Inst. 2015;107(7).
104. Harrod A, Fulton J, Nguyen VTM, Periyasamy M, Ramos-Garcia L, Lai CF, et al. Genomic modelling of the ESR1 Y537S mutation for evaluating function and new therapeutic approaches for metastatic breast cancer. Oncogene. 2017;36(16):2286-96.
105. Yu S, Kim T, Yoo KH, Kang K. The T47D cell line is an ideal experimental model to elucidate the progesterone-specific effects of a luminal A subtype of breast cancer. Biochem Biophys Res Commun. 2017;486(3):752-8.

106. Wierstra I, Alves J. Despite its strong transactivation domain, transcription factor FOXM1c is kept almost inactive by two different inhibitory domains. Biol Chem. 2006;387(7):963-76.
107. *Oliveros JC. Venny. An interactive tool for comparing lists with Venn's diagrams.* https://bioinfogp.cnb.csic.es/tools/venny/index.html. *2007-2015*.
108. Cole CG, Mccann OT, Collins JE, Oliver K, Willey D, Gribble SM, et al. Finishing the finished human chromosome 22 sequence. Genome Biology. 2008;9(5):R78.
109. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31(14):2382-3.
110. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology. 2012;16(5):284-7.
111. Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, et al. BEDOPS: high-performance genomic feature operations. Bioinformatics. 2012;28(14):1919-20.
112. Neph S, Reynolds AP, Kuehn MS, Stamatoyannopoulos JA. Operating on Genomic Ranges Using BEDOPS. Methods Mol Biol. 2016;1418:267-81.
113. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, et al. Genome-wide analysis of estrogen receptor binding sites. Nat Genet. 2006;38(11):1289-97.
114. Chen Y, Xu L, Mayakonda A, Huang ML, Kanojia D, Tan TZ, et al. Bromodomain and extraterminal proteins foster the core transcriptional regulatory programs and confer vulnerability in liposarcoma. Nat Commun. 2019;10(1):1353.
115. Brunelle M, Nordell Markovits A, Rodrigue S, Lupien M, Jacques PE, Gevry N. The histone variant H2A.Z is an important regulator of enhancer activity. Nucleic Acids Res. 2015;43(20):9742-56.
116. Lei JT, Gou X, Seker S, Ellis MJ. ESR1 alterations and metastasis in estrogen receptor positive breast cancer. J Cancer Metastasis Treat. 2019;5.
117. Maruthanila VL, Elancheran R, Kunnumakkara AB, Kabilan S, Kotoky J. Recent development of targeted approaches for the treatment of breast cancer. Breast Cancer. 2017;24(2):191-219.
118. Millour J, Constantinidou D, Stavropoulou AV, Wilson MS, Myatt SS, Kwok JM, et al. FOXM1 is a transcriptional target of ERalpha and has a critical role in breast cancer endocrine sensitivity and resistance. Oncogene. 2010;29(20):2983-95.
119. Madureira PA, Varshochi R, Constantinidou D, Francis RE, Coombes RC, Yao KM, et al. The Forkhead box M1 protein regulates the transcription of the estrogen receptor alpha in breast cancer cells. J Biol Chem. 2006;281(35):25167-76.
120. Bergamaschi A, Madak-Erdogan Z, Kim YJ, Choi YL, Lu H, Katzenellenbogen BS. The forkhead transcription factor FOXM1 promotes endocrine resistance and invasiveness in estrogen receptor-positive breast cancer by expansion of stem-like cancer cells. Breast Cancer Res. 2014;16(5):436.
121. Cicatiello L, Mutarelli M, Grober OM, Paris O, Ferraro L, Ravo M, et al. Estrogen receptor alpha controls a gene network in luminal-like breast cancer cells comprising multiple transcription factors and microRNAs. Am J Pathol. 2010;176(5):2113-30.
122. Chen M, Zhang Z, Meng ZY, Zhang XJ. [ATAC-seq and its applications in complex disease]. Yi Chuan. 2020;42(4):347-53.
123. Reske JJ, Wilson MR, Chandler RL. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. Epigenetics Chromatin. 2020;13(1):22.
124. Robinson JT, Thorvaldsdottir H, Wenger AM, Zehir A, Mesirov JP. Variant Review with the Integrative Genomics Viewer. Cancer Res. 2017;77(21):e31-e4.
125. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-6.
126. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14(2):178-92.

127. Tang Q, Chen Y, Meyer C, Geistlinger T, Lupien M, Wang Q, et al. A comprehensive view of nuclear receptor cancer cistromes. Cancer Res. 2011;71(22):6940-7.
128. Carr JR, Kiefer MM, Park HJ, Li J, Wang Z, Fontanarosa J, et al. FoxM1 regulates mammary luminal cell fate. Cell Rep. 2012;1(6):715-29.
129. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
130. Liu T. Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. Methods Mol Biol. 2014;1150:81-95.
131. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.
132. Kyrchanova O, Georgiev P. Mechanisms of Enhancer-Promoter Interactions in Higher Eukaryotes. International Journal of Molecular Sciences. 2021;22(2):671.
133. Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, et al. Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. Nature. 2015;518(7540):556-9.
134. Pugacheva EM, Kubo N, Loukinov D, Tajmul M, Kang S, Kovalchuk AL, et al. CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. Proceedings of the National Academy of Sciences. 2020;117(4):2020-31.
135. Van Der Lelij P, Lieb S, Jude J, Wutz G, Santos CP, Falkenberg K, et al. Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts. eLife. 2017;6.
136. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res. 2015;43(W1):W39-49.
137. Yao Z, Macquarrie KL, Fong AP, Tapscott SJ, Ruzzo WL, Gentleman RC. Discriminative motif analysis of high-throughput dataset. Bioinformatics. 2014;30(6):775-83.
138. Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts A, Jason, Mahony S, Pugh F, B., et al. The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Molecular Cell. 2016;62(1):79-91.
139. Jeffreys SA, Powter B, Balakrishnar B, Mok K, Soon P, Franken A, et al. Endocrine Resistance in Breast Cancer: The Role of Estrogen Receptor Stability. Cells. 2020;9(9):2077.
140. Gasser E, Johannssen HC, Rülicke T, Zeilhofer HU, Stoffel M. Foxa1 is essential for development and functional integrity of the subthalamic nucleus. Scientific Reports. 2016;6(1):38611.
141. Fu X, Jeselsohn R, Pereira R, Hollingsworth EF, Creighton CJ, Li F, et al. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. Proceedings of the National Academy of Sciences. 2016;113(43):E6600-E9.
142. Ohneda K, Yamamoto M. Roles of Hematopoietic Transcription Factors GATA-1 and GATA-2 in the Development of Red Blood Cell Lineage. Acta Haematologica. 2002;108(4):237-45.
143. Gold DA, Baek SH, Schork NJ, Rose DW, Larsen DD, Sachs BD, et al. RORα Coordinates Reciprocal Signaling in Cerebellar Development through Sonic hedgehog and Calcium-Dependent Pathways. Neuron. 2003;40(6):1119-31.
144. Wang D, Russell JL, Johnson DG. E2F4 and E2F1 Have Similar Proliferative Properties but Different Apoptotic and Oncogenic Properties In Vivo. Molecular and Cellular Biology. 2000;20(10):3417-24.
145. Gabet Y, Baniwal SK, Leclerc N, Shi Y, Kohn-Gabet AE, Cogan J, et al. Krox20/EGR2 deficiency accelerates cell growth and differentiation in the monocytic lineage and decreases bone mass. Blood. 2010;116(19):3964-71.
146. Kobayashi T, Kageyama R. Hes1 regulates embryonic stem cell differentiation by suppressing Notch signaling. Genes to Cells. 2010;15(7):689-98.
147. Xie R, Medina R, Zhang Y, Hussain S, Colby J, Ghule P, et al. The histone gene activator HINFP is a nonredundant cyclin E/CDK2 effector during early embryonic cell cycles. Proceedings of the National Academy of Sciences. 2009;106(30):12359-64.

148. Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. GATA-3 Maintains the Differentiation of the Luminal Cell Fate in the Mammary Gland. Cell. 2006;127(5):1041-55.
149. Curina A, Termanini A, Barozzi I, Prosperini E, Simonatto M, Polletti S, et al. High constitutive activity of a broad panel of housekeeping and tissue-specific cis-regulatory elements depends on a subset of ETS proteins. Genes & Development. 2017;31(4):399-412.
150. Walker W, Zhou Z-Q, Ota S, Wynshaw-Boris A, Hurlin PJ. Mnt–Max to Myc–Max complex switching regulates cell cycle entry. Journal of Cell Biology. 2005;169(3):405-13.
151. Zhang F, Song H, Zeng M, Wu F, Li Y, Pan Y, et al. A deep learning framework for gene ontology annotations with sequence - and network-based information. IEEE/ACM Trans Comput Biol Bioinform. 2020;PP.
152. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics. 2005;21(16):3439-40.
153. Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, et al. BioMart--biological queries made easy. BMC Genomics. 2009;10:22.
154. Beissbarth T, Speed TP. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics. 2004;20(9):1464-5.
155. Yau C, Meyer L, Benz S, Vaske C, Scott G, Egan B, et al. FOXM1 cistrome predicts breast cancer metastatic outcome better than FOXM1 expression levels or tumor proliferation index. Breast Cancer Res Treat. 2015;154(1):23-32.
156. Nizovtseva EV, Studitsky VM. Mechanisms of Enhancer Action. eLS. 2019:1-7.
157. Muerdter F, Boryń ŁM, Arnold CD. STARR-seq — Principles and applications. Genomics. 2015;106(3):145-50.
158. Zheng L, Qian B, Tian D, Tang T, Wan S, Wang L, et al. FOXA1 positively regulates gene expression by changing gene methylation status in human breast cancer MCF-7 cells. Int J Clin Exp Pathol. 2015;8(1):96-106.
159. Dang V, Chi. MYC on the Path to Cancer. Cell. 2012;149(1):22-35.
160. Hegde NS, Sanders DA, Rodriguez R, Balasubramanian S. The transcription factor FOXM1 is a cellular target of the natural product thiostrepton. Nat Chem. 2011;3(9):725-31.
161. Kongsema M, Wongkhieo S, Khongkow M, Lam EW, Boonnoy P, Vongsangnak W, et al. Molecular mechanism of Forkhead box M1 inhibition by thiostrepton in breast cancer cells. Oncol Rep. 2019;42(3):953-62.
162. Gormally MV, Dexheimer TS, Marsico G, Sanders DA, Lowe C, Matak-Vinkovic D, et al. Suppression of the FOXM1 transcriptional programme via novel small molecule inhibition. Nat Commun. 2014;5:5165.
163. Wonsey DR, Follettie MT. Loss of the Forkhead Transcription Factor FoxM1 Causes Centrosome Amplification and Mitotic Catastrophe. Cancer Research. 2005;65(12):5181-9.
164. Johnston SR, Martin LA, Head J, Smith I, Dowsett M. Aromatase inhibitors: combinations with fulvestrant or signal transduction inhibitors as a strategy to overcome endocrine resistance. J Steroid Biochem Mol Biol. 2005;95(1-5):173-81.
165. Al-Lazikani B, Banerji U, Workman P. Combinatorial drug therapy for cancer in the post-genomic era. Nat Biotechnol. 2012;30(7):679-92.
166. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell. 2010;141(1):69-80.
167. Pu M, Messer K, Davies SR, Vickery TL, Pittman E, Parker BA, et al. Research-based PAM50 signature and long-term breast cancer survival. Breast Cancer Research and Treatment. 2020;179(1):197-206.
168. Urra H, Dufey E, Avril T, Chevet E, Hetz C. Endoplasmic Reticulum Stress and the Hallmarks of Cancer. Trends in Cancer. 2016;2(5):252-62.
169. Dalton LE, Clarke HJ, Knight J, Lawson MH, Wason J, Lomas DA, et al. The endoplasmic reticulum stress marker CHOP predicts survival in malignant mesothelioma. British Journal of Cancer. 2013;108(6):1340-7.

170. Wang H-Q, Du Z-X, Zhang H-Y, Gao D-X. Different Induction of GRP78 and CHOP as a Predictor of Sensitivity to Proteasome Inhibitors in Thyroid Cancer Cells. Endocrinology. 2007;148(7):3258-70.
171. Sisinni L, Pietrafesa M, Lepore S, Maddalena F, Condelli V, Esposito F, et al. Endoplasmic Reticulum Stress and Unfolded Protein Response in Breast Cancer: The Balance between Apoptosis and Autophagy and Its Role in Drug Resistance. Int J Mol Sci. 2019;20(4).
172. Alasiri G, Fan LY-N, Zona S, Goldsbrough IG, Ke H-L, Auner HW, et al. ER stress and cancer: The FOXO forkhead transcription factor link. Molecular and Cellular Endocrinology. 2018;462:67-81.
173. Alasiri G, Jiramongkol Y, Zona S, Fan LY-N, Mahmud Z, Gong G, et al. Regulation of PERK expression by FOXO3: a vulnerability of drug-resistant cancer cells. Oncogene. 2019;38(36):6382-98.
174. Alasiri G, Jiramongkol Y, Trakansuebkul S, Ke H-L, Mahmud Z, Intuyod K, et al. Reciprocal regulation between GCN2 (eIF2AK4) and PERK (eIF2AK3) through the JNK-FOXO3 axis to modulate cancer drug resistance and clonal survival. Molecular and Cellular Endocrinology. 2020;515:110932.
175. Yao S, Fan LY-N, Lam EW-F. The FOXO3-FOXM1 axis: A key cancer drug target and a modulator of cancer drug resistance. Seminars in Cancer Biology. 2018;50:77-89.
176. Eijkelenboom A, Mokry M, de Wit E, Smits LM, Polderman PE, van Triest MH, et al. Genome-wide analysis of FOXO3 mediated transcription regulation through RNA polymerase II profiling. Mol Syst Biol. 2013;9:638.
177. Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. Nucleic Acids Res. 2019;47(D1):D729-D35.
178. Yu JSE, Colborne S, Hughes CS, Morin GB, Nielsen TO. The FUS-DDIT3 Interactome in Myxoid Liposarcoma. Neoplasia. 2019;21(8):740-51.
179. Dong D, Ko B, Baumeister P, Swenson S, Costa F, Markland F, et al. Vascular targeting and antiangiogenesis agents induce drug resistance effector GRP78 within the tumor microenvironment. Cancer Res. 2005;65(13):5785-91.
180. Zhao F, Siu MK, Jiang L, Tam KF, Ngan HY, Le XF, et al. Overexpression of forkhead box protein M1 (FOXM1) in ovarian cancer correlates with poor patient survival and contributes to paclitaxel resistance. PLoS One. 2014;9(11):e113478.
181. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A. 2010;107(50):21931-6.
182. Karadedou CT, Gomes AR, Chen J, Petkovic M, Ho KK, Zwolinska AK, et al. FOXO3a represses VEGF expression through FOXM1-dependent and -independent mechanisms in breast cancer. Oncogene. 2012;31(14):1845-58.
183. Karadedou CT, Gomes AR, Chen J, Petkovic M, Ho KK, Zwolinska AK, et al. Correction: FOXO3a represses VEGF expression through FOXM1-dependent and -independent mechanisms in breast cancer. Oncogene. 2019;38(25):5111-2.
184. Delpuech O, Griffiths B, East P, Essafi A, Lam EW-F, Burgering B, et al. Induction of Mxi1-SRα by FOXO3a Contributes to Repression of Myc-Dependent Gene Expression. Molecular and Cellular Biology. 2007;27(13):4917-30.
185. Delpuech O, Griffiths B, East P, Essafi A, Lam EW, Burgering B, et al. Induction of Mxi1-SR alpha by FOXO3a contributes to repression of Myc-dependent gene expression. Mol Cell Biol. 2007;27(13):4917-30.
186. Ponnusamy L, Mahalingaiah PKS, Chang Y-W, Singh KP. Reversal of epigenetic aberrations associated with the acquisition of doxorubicin resistance restores drug sensitivity in breast cancer cells. European Journal of Pharmaceutical Sciences. 2018;123:56-69.
187. Ocaña A, Pandiella A. Identifying Breast Cancer Druggable Oncogenic Alterations: Lessons Learned and Future Targeted Options. Clinical Cancer Research. 2008;14(4):961-70.

188. Nandi D, Cheema PS, Jaiswal N, Nag A. FoxM1: Repurposing an oncogene as a biomarker. Seminars in Cancer Biology. 2018;52:74-84.
189. Lu X-F, Zeng D, Liang W-Q, Chen C-F, Sun S-M, Lin H-Y. FoxM1 is a promising candidate target in the treatment of breast cancer. Oncotarget. 2018;9(1):842-52.
190. Dumitrescu RG. Interplay Between Genetic and Epigenetic Changes in Breast Cancer Subtypes. Methods in Molecular Biology: Springer New York; 2018. p. 19-34.
191. Karunarathna U, Kongsema M, Zona S, Gong C, Cabrera E, Gomes AR, et al. OTUB1 inhibits the ubiquitination and degradation of FOXM1 in breast cancer and epirubicin resistance. Oncogene. 2016;35(11):1433-44.
192. Waseem A, Ali M, Odell EW, Fortune F, Teh MT. Downstream targets of FOXM1: CEP55 and HELLS are cancer progression markers of head and neck squamous cell carcinoma. Oral Oncol. 2010;46(7):536-42.
193. Teh MT, Gemenetzidis E, Patel D, Tariq R, Nadir A, Bahta AW, et al. FOXM1 induces a global methylation signature that mimics the cancer epigenome in head and neck squamous cell carcinoma. PLoS One. 2012;7(3):e34329.
194. Hwang S, Mahadevan S, Qadir F, Hutchison IL, Costea DE, Neppelberg E, et al. Identification of FOXM1-induced epigenetic markers for head and neck squamous cell carcinomas. Cancer. 2013;119(24):4249-58.
195. Holm K, Hegardt C, Staaf J, Vallon-Christersson J, Jonsson G, Olsson H, et al. Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. Breast Cancer Res. 2010;12(3):R36.
196. Mayran A, Drouin J. Pioneer transcription factors shape the epigenetic landscape. J Biol Chem. 2018;293(36):13795-804.
197. Nyce J, Leonard S, Canupp D, Schulz S, Wong S. Epigenetic mechanisms of drug resistance: drug-induced DNA hypermethylation and drug resistance. Proc Natl Acad Sci U S A. 1993;90(7):2960-4.
198. Sandhu R, Rivenbark AG, Coleman WB. Enhancement of chemotherapeutic efficacy in hypermethylator breast cancer cells through targeted and pharmacologic inhibition of DNMT3b. Breast Cancer Res Treat. 2012;131(2):385-99.
199. Wei D, Loeken MR. Increased DNA methyltransferase 3b (Dnmt3b)-mediated CpG island methylation stimulated by oxidative stress inhibits expression of a gene required for neural tube and neural crest development in diabetic pregnancy. Diabetes. 2014;63(10):3512-22.
200. Ferreira D, Adega F, Chaves R. The Importance of Cancer Cell Lines as in vitro Models in Cancer Methylome Analysis and Anticancer Drugs Testing. Oncogenomics and Cancer Proteomics - Novel Approaches in Biomarkers Discovery and Therapeutic Targets in Cancer: InTech; 2013.
201. Tian B, Zhao C, Gu F, He Z. A two-step framework for inferring direct protein-protein interaction network from AP-MS data. BMC Systems Biology. 2017;11(S4).
202. Collas P. The current state of chromatin immunoprecipitation. Mol Biotechnol. 2010;45(1):87-100.
203. Blainey P, Krzywinski M, Altman N. Replication. Nature Methods. 2014;11(9):879-80.
204. Robasky K, Lewis NE, Church GM. The role of replicates for error mitigation in next-generation sequencing. Nature Reviews Genetics. 2014;15(1):56-62.

8 Appendix

SDS-PAGE minigel

x1 gel

	resolving gel					stacking
	5%	7%	10%	12%	14%	5%
dH2O-ml	5.68	5.02	4.02	3.35	2.68	3.67
1.5M **Tris** pH 8.8	2.5	2.5	2.5	2.5	2.5	-
1.5M **Tris** pH 6.8	-	-	-	-	-	0.42
30% **Acrylamide**/ 0.8% **Bis** mix (37.5:1)	1.67	2.33	3.33	4.00	4.67	0.83
10% **SDS**-µl	100	100	100	100	100	50
25% **APS**-µl	40	40	40	40	40	20
TEMED-µl	10	10	10	10	10	10
total-ml	10	10	10	10	10	5

Resolving gel

0.375M Tris pH 8.8, acrylamide, 0.1% (w/v) SDS, 0.1% (w/v) APS,

0.1% (v/v) TEMED

Stacking gel

0.126M Tris pH 6.8, acrylamide, 0.1% (w/v) SDS, 0.1% (w/v) APS,

TEMED

SDS-PAGE Electrophoresis Buffer (1x)

25mM Tris base, 250mM glycine, 0.1% (w/v) SDS

10x stock

Transfer Buffer (1x)

24mM Tris base, 193mM glycine, 20% (v/v) methanol

10x stock

Appendix 8.1 Recipes for gels and buffers of SDS-PAGE.

TAE Buffer (1X)

40mM Tris·acetate, 1mM EDTA [50x stock (57.1ml glacial acetic acid in 1L)]

0.5 M EDTA, pH 8

Adjust pH to 8 with NaOH~20g

TE, pH 8

10mM Tris-HCl, pH 8, 1mM EDTA

TBS-Tween Buffer pH 7.5

40mM Tris-HCl, 300mM NaCl, 0.05% (v/v) Tween 20 (20x stock)

Buffer I

0.25% Triton X-100; 10mM EDTA; 0.5mM EGTA; 10mM HEPES pH 6.5

Buffer II

200mM NaCl; 10mM EDTA; 0.5mM EGTA; 10mM HEPES pH 6.5

Lysis Buffer (ChIP)

1% SDS; 10mM EDTA; 50mM Tris-HCl pH 8.1

TSE I

0.1% SDS; 1% Triton X-100; 2mM EDTA; 20mM Tris-HCl pH 8.1; 150mM NaCl

Scraping Buffer

100mM Tris-HCl pH 9.4; 0.1% SDS

Elution Buffer

0.1M $NaHCO_3$; 1% SDS

Appendix 8.2 Buffer recipes.