Molecular Investigation of Anticancer Genes

Imperial College London

Department of Surgery and Cancer

Thesis submitted by

Motasim Masood

for the degree of

Doctor of Philosophy

Declaration of Originality

Unless stated, acknowledged and/or referenced otherwise in the text, this thesis is the result of my own work.

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

The lack of specificity of the therapeutic agents against cancer is one of the prime obstacles in developing effective therapies for the disease. The specificity of chemotherapeutic agents against cancer cells is primarily quantitative rather than qualitative. Despite the earlier success of targeted therapies, it is now clear that among an unselected population of patients, the response rate is usually limited to 10-20%. Further avenues need to be explored for the development of cancer-specific therapies. A recent concept towards such efforts is that of anticancer genes. When ectopically overexpressed, anticancer genes specifically destroy tumour cells without harming normal cells. The examples include TRAIL, MDA7, PAR4, apoptin and Orctl3. In addition, 22 novel anticancer genes were identified through systematic functional genetic screening. This thesis aimed at (i) deciphering the identity of these genes, (ii) investigating their regulation and (iii) exploring the mechanisms behind their cancer specific actions. Sanger sequencing and subsequent analyses revealed that these 22 novel anticancer genes included non-coding RNAs, novel transcripts and protein-coding genes, two among the latter are TMEFF2 and FBLN5. The expression levels of both these genes exhibit a positive correlation with the survival of breast cancer patients and both are positively regulated by FOXO3 in MCF-7 cells. FBLN5 and TMEFF2 cause cell death selectively in the COS-7 cells and not in their untransformed counterparts (CV-1 cells). Mechanistic gene expression profiling and elaborate comparison pathways analyses uncovered that FBLN5 undertakes this action putatively through the downregulation of the downstream targets of MYC and inhibition of oxidative phosphorylation. Additionally, TMEFF2 causes caspase-dependent and p53-independent apoptosis in MCF-7 cells. Gene expression profiling revealed that the TMEFF2 induces potent interferon signalling along with the upregulation of TRAIL in MCF-7 cells. TMEFF2 gives a perinuclear punctate cytoplasmic appearance in MCF-7 cells and does not co-localize with either mitochondria or α-tubulin.

Dedicated to,

Professor Stefan W. Grimm, 1963-2014



Photograph Courtesy of Dr Ming Shih-Hwang.

Reprinted by permissions from Springer Nature: Springer Nature. CELL DEATH & DIFFERENTIATION. Stefan Grimm, 1963-2014, a tragic loss for the scientific community. A-L Mahul-Mellier, L Huc, A Lemarié, S Irshad, M Hwang *et al.* 2015;22(2):247. doi:10.1038/cdd.2014.196. Copyright © 2015, Springer Nature. License number (Copyright Clearance Center order number) 4656660239195.

Acknowledgements

After the untimely departure of my mentor and friend Prof Stefan W Grimm, I would have never been able to finish the lab work let alone write a thesis if it had not been for the help of many wonderful people. I cannot thank them all enough. First of all, my friend and safety manager Paul Coote helped in so many ways that I cannot even enumerate. I would like to thank Miss Ottilie Swinyard (who did her masters project in our lab) for her dedication, creative input and mastering the art of western blotting. I would like to thank my fellow lab members, Dr Nesreen Magdy for tolerating me, Susan Van Noorden for her blissful presence and Rean Muhammad for her prayers, her encouragement and her sincere support. I would also like to convey my gratitude to Dr Adam Cawte for his skilful work with the imaging, his technical input, provision of shRNA clones and his utmost and selfless help with the project. I am grateful to my friend and former fellow PhD student, Dr Qize Ding, for his help with two experiments ($Dioc_6$ + PI staining and MTT assay), his intellectual input and fun times in and out of the lab. I would like to thank Prof Stefan Grimm for securing a generous grant for this project from the University of Dammam through his collaboration with Prof Amein Al-Ali and also for offering me the opportunity to work on this project.

At the second unfortunate turn of events during this PhD, again many came to help. Susan Farrell, Prof Bob Brown, Dr Timothy Ebbels, Rebecca Smith, Jennifer Simeon, Dr Amir Hakim, Dr Paul Seldon and most importantly the college tutor Dr Mick Jones. I am deeply grateful to Beth Barker, Dr Christine O'Farrely and Prof Paul Ramchandani for looking after my mental health; interestingly, they are specialists in child and adolescent psychiatry. Our lab support staff Cecilia George kept my spirits high and ensured a timely provision of the crucial labware. I would like to thank my dear friend Dr Suchira Gallage and his associates in Jesús Gil's group, Dr Andrew Inns and Amy Barkway for their help with automated high throughput images, technical input and some badly needed company. I am also grateful to Dr Jason Bennett for his intellectual input and assistance with the experiments, and I am thankful to his associates in Guido Franzoso's group, Dr Daria Capece and Dr Federica Begalli for their help, support and emergency supply of chemicals. I am also thankful to Mr Mike Newton from security for checking on me in the lab during the frequent odd hours. I am grateful to Dr Wayne Michelle and Dr Anabel Varela Carvar for giving me extra shifts of graduate teaching assistant work so I could sustain myself financially. I would like to convey my gratitude to the Edith Evelyn Wali Mohammad Memorial Fund Grants for helping me through the last few months of lab work. I would

like to convey my deepest gratitude to Dr Naveed Anwar from Thermo Fisher Scientific for pulling all the possible strings to get me the high-end microarrays at an unprecedented discount. In this regard I am also grateful to Dr David Chambers from UKB for his expertise with microarray processing and for the discount provided. The world is full of wonderful people – I only met Dr David Mosen Ansorena once and he gave me access to an expensive bioinformatic tool, the Ingenuity Pathway Analysis (IPA) for free, it helped in some amazing ways.

Professor Stefan Grimm was (and is) missed direly. It is impossible to encompass his wonderful qualities in a few words but briefly, Stefan was genuinely sincere, honest, upright, caring, considerate and understanding. He had the ideal work ethic and his door was always open on purpose. He was exceptionally supportive when experiments did not work and because of his utmost scientific integrity, it was very difficult to convince him of the positive results, but the best part was that even in the most heated scientific argument, he would never allow feelings to mask the judgement and would always come up with a creative solution. He would always keep certain boundaries of civilities, never intrude in anyone's personal space and yet it was amazingly comfortable to communicate with him. There was not a single day during this PhD that I didn't wish I could pick his brains and whenever the results proved his ideas, I yearned to run to his office and show him the results (for him to kill the excitement ☺ – it was very difficult to convince him of the positive results).

I am also deeply indebted to all of my family. I am unspeakably thankful to my parents; during his very tough battle with cancer, my father kept me from knowing about his agony, so that I would not get distracted during my PhD and he also kept supporting me financially to his very best. He passed away without ever telling me that he wanted so badly to see me in the final days. It is a true blessing to be able to rest under the shadow of my mother's wisdom and under the wings of her prayers. My lovely wife Silvia looked after me, helped me, supported me, tolerated me, stood by me, put food on the table (both financially and literally) and helped every step of the way, including writing this thesis by correcting my horrid punctuation and wicked grammar. I will never be able to understand what I have done to deserve her.

Last, but not least, I would like to convey my heartiest gratitude to my two supervisors. I wish I could put in writing the million ways in which Prof Mona El-Bahrawy helped and supported me like a mother. And about Dr Ernesto Yagüe, I am simply out of words, but I can say with my hand on my heart that I have never known a better person in my life – thank you, Sir!

TABLE OF CONTENTS

ABSTRACT **3**

ACKNOWLEDGEMENTS **5**

1 INTRODUCTION **15**

1.1 ANTICANCER GENES: MECHANISM OF ACTION 17

1.2 ADVANTAGES OF ANTICANCER GENES OVER CONVENTIONAL THERAPY 18

1.3 ANTICANCER GENES *VS* TUMOUR SUPPRESSOR GENES 19

1.4 CELL DEATH 20

1.4.1 Apoptosis *22*

1.4.2 Mitochondrial permeability transition (MPT)-driven necrosis *26*

1.4.3 Necroptosis *27*

1.4.4 Ferroptosis *28*

1.4.5 Pyroptosis *28*

1.4.6 Parthanatos *29*

1.4.7 Entotic cell death *29*

1.4.8 Lysosome-dependent cell death *30*

1.4.9 Autophagy-dependent cell death *31*

1.4.10 Immunogenic cell death *31*

1.4.11 Cellular senescence, mitotic catastrophe and terminal differentiation *32*

1.5 PREVIOUSLY KNOWN ANTICANCER GENES 35

1.5.1 Apoptin *35*

1.5.2 E4orf4 ..... *37*

1.5.3 Rodent parvovirus NS1 ..... *38*

1.5.4 MDA-7/IL-24 ..... *39*

1.5.5 PAR4 ..... *40*

1.5.6 Orctl3 ..... *41*

1.5.7 TRAIL *42*

1.5.8 HAMLET *45*

1.6 EXPECTATIONS AND CHALLENGES 48

1.7 INTRODUCTION TO THE ROBOTIC SINGLE CDNA INVESTIGATION (RISCI) 49

1.8 FUNCTIONAL GENETIC SCREEN FOR ANTICANCER GENES 52

1.9 AIMS 54

1.9.1 Brief Synopsis *54*

2 MATERIALS AND METHODS **55**

2.1 PLASMIDS 55

2.2 SANGER SEQUENCING 56

2.3 Sequence analyses 56
2.4 Kaplan-Meier survival estimate 56
2.5 Cell culture 58
2.6 Plasmid Transfection 58
2.6.1 Transient Transfection 58
2.6.2 Stable Transfection 60
2.7 siRNA Transfection 61
2.8 $DiOC_6$ + PI Staining - Flow Cytometry- Cell Death Assay 61
2.9 Colorimetric Methylthiazole Tetrazolium (MTT) Assay 62
2.10 Immunoblotting 62
2.11 Quantitative Real time PCR (qRT-PCR) 64
2.11.1 RNA extraction 64
2.11.2 cDNA Synthesis 65
2.11.3 qRT-PCR Reaction 65
2.11.4 Data analysis 65
2.12 Caspase-Glo® 3/7, caspase activity assay 67
2.13 TMRE + DRAQ7™ Staining - Flow Cytometry - Cell Death Assay 67
2.13.1 Caspase inhibition using z-VAD-FMK 68
2.13.2 TP53 inhibition using Pifithrin-Alpha 68
2.14 Live cell Imaging 68
2.15 Immunofluorescence 69
2.16 Microarray Analyses 70
2.16.1 Microarray Processing 70
2.16.2 Data analysis 70
2.17 Phenotype Inspection cell death assay 73
2.18 Automated Cell Counting Assay 74
2.19 Gene Sub-Cloning 74
2.20 Crystal Violet staining 75
3 IDENTIFICATION AND SELECTION 76
3.1 Background 76
3.2 Results 78
3.2.1 Sequence Verification 78
3.2.2 Gene Identities 82
3.2.3 Prognostic Value 98
3.2.4 Anticancer genes in MCF-7 cells 106
3.2.5 Selection of a subset for further investigation 112
3.2.6 Validation of overexpression 112

3.3 Discussion 114

4 ANTICANCER GENES AND FOX PROTEINS 118

4.1 Background 118

4.1.1 Forkhead box (FOX) proteins 118

4.1.2 FOX proteins and Anticancer genes? 120

4.2 Results 121

4.2.1 Regulation of Anticancer genes by FOX proteins 121

4.3 Discussion 128

5 TMEFF2 IN BREAST CANCER 130

5.1 Background 130

5.1.1 TMEFF2, an Overview 131

5.2 Results 156

5.2.1 TMEFF2 has an onco-suppressive role in breast cancer 156

5.2.2 TMEFF2 overexpression causes apoptosis in MCF-7 cells 159

5.2.3 TMEFF2-induced apoptosis is independent of p53 activity 174

5.2.4 Intracellular localization of TMEFF2 in MCF-7 cells 179

5.2.5 Preparation for transcriptomics analysis 184

5.2.6 Differential gene expression analysis of TMEFF2 overexpressing MCF-7 cells 187

5.2.7 Pathway analysis 195

5.2.8 Attempts of functional validation 218

5.3 Discussion 221

6 FBLN5 AS AN ANTICANCER GENE 233

6.1 FBLN5, a Brief Overview 233

6.2 Results 236

6.2.1 FBLN5 and TMEFF2 caused cell death in COS-7 but not in CV-1 cells 236

6.2.2 Comparative levels of TMEFF2 and FBLN5 overexpression between CV-1 and COS-7 cells 243

6.2.3 Intracellular localization of FBLN5 in CV-1 and COS-7 cells 245

6.2.4 FBLN5 Comparative Transcriptomic Analyses 249

6.3 Discussion 273

7 FINAL DISCUSSION: SUMMARY, PROSPECTS AND CONTRIBUTION TO KNOWLEDGE 279

8 REFERENCES 282

9 APPENDIX 314

LIST OF FIGURES

Figure 1.1: Regulation of Apoptosis Overview, ... 26
Figure 1.2: Traditional screening approach: Screening through batch transfection and sib selection. 51
Figure 1.3: Novel approach to cDNA library screening, Robotic single cDNA investigation (RISCI). 52
Figure 1.4: Screening the anticancer genes. 53
Figure 3.1: Short representative peak views for the Sanger sequencing results. 80
Figure 3.2: Ensembl BLAT results for ACG1 83
Figure 3.3: Ensembl BLAT results for ACG2.. 84
Figure 3.4: Ensembl BLAT results for ACG3. 84
Figure 3.5: Ensembl BLAT results for ACG4.. 85
Figure 3.6: Ensembl BLAT results for ACG5.. 85
Figure 3.7: Ensembl BLAT results for ACG6. 86
Figure 3.8: Ensembl BLAT results for ACG7. 87
Figure 3.9: Ensembl BLAT results for ACG8. 87
Figure 3.10: Ensembl BLAT results for ACG9. 88
Figure 3.11: Ensembl BLAT results for ACG10. 88
Figure 3.12: Ensembl BLAT results for ACG11 89
Figure 3.13: Ensembl BLAT results for ACG12. 89
Figure 3.14: Ensembl BLAT results for ACG13. 90
Figure 3.15: Ensembl BLAT results for ACG14. 90
Figure 3.16: Ensembl BLAT results for ACG15. 91
Figure 3.17: Ensembl BLAT results for ACG16 91
Figure 3.18: Ensembl BLAT results for ACG17. 92
Figure 3.19: Ensembl BLAT Results for ACG18. 92
Figure 3.20: Ensembl BLAT results for ACG19. 93
Figure 3.21: Ensembl BLAT results for ACG20. 94
Figure 3.22: Ensembl BLAT results for ACG21. 94
Figure 3.23: Ensembl BLAT results for ACG22. 95
Figure 3.24: Kaplan-Meier plots showing the correlations between *ZNF436_AS1* expression and patient survival 99
Figure 3.25: Kaplan-Meier plots showing the correlations between *SMLR1* expression and patient survival 100
Figure 3.26: Kaplan-Meier plots showing the correlations between *TMEFF2* expression and patient survival: 101
Figure 3.27: Kaplan-Meier plots showing the correlations between *HYAL2* expression and patient survival 102
Figure 3.28: Kaplan-Meier plots showing the correlations between *NEIL2* expression and patient survival 103
Figure 3.29: Kaplan-Meier plots showing the correlations between *FBLN5* expression and patient survival 104
Figure 3.30: Kaplan-Meier plots showing the correlations between *YPEL4* expression and patient survival 105
Figure 3.31: Cell death in MCF-7 cells upon experimental overexpression of novel anticancer genes. 107
Figure 3.32: Relative viability of MCF-7 cells upon overexpression of novel anticancer genes – MTT colorimetric assay. 111

Figure 3.33: Experimental overexpression of anticancer genes in MCF-7 cells. 113
Figure 4.1: The FOXM1-FOXO3 axis 121
Figure 4.2: qRT-PCR results for the validation of overexpression and knockdown of FOXM1, FOXO3, FOXK2 and FOXO3-A3-always active mutant (overexpression only) in MCF-7 cells. 124
Figure 4.3: Immunoblots for the confirmation of (a) overexpression and (b) knockdown of FOXM1, FOXO3, FOXK2 and FOXO3-A3-always active mutant (overexpression only) in MCF-7 cells. 125
Figure 4.4: Effects of overexpression and knockdowns of FOXM1, FOXO3, FOXK2 and FOXO3-A3-always active mutant (overexpression only) on the mRNA levels of anticancer genes in MCF-7 cells. 127
Figure 5.1: TMEFF2 protein features. 132
Figure 5.2: Kaplan-Meier plots showing the correlation between TMEFF2 mRNA expression and breast cancer patient survival. 158
Figure 5.3: (b) Early and late apoptosis in TEMFF2 overexpressing MCF-7 cells upon inhibition of caspases. 160
Figure 5.3 (a): Caspase 3/7 activity upon TMEFF2 overexpression in MCF-7 cells 160
Figure 5.3: (c) TMRE (x-axis) and DRAQ7™ (y-axis) staining upon TMEFF2 overexpression in MCF-7 Cells:. 162
Figure 5.4: Immunoblots for key apoptotic factors upon TMEFF2 overexpression. 164
Figure 5.5: Time-lapse micrographs of the MCF-7 cells: Imaging Field 1 169
Figure 5.5: Time-lapse micrographs of the MCF-7 cells: Imaging Field 2 173
Figure 5.6: TMEFF2 induced apoptosis in MCF-7 cells upon p53 knockdown. 176
Figure 5.7 (a): TMEFF2 induced apoptosis in MCF-7 cells upon p53 inhibition. 177
Figure 5.7 (b): TMRE (x-axis) and DRAQ7™ (y-axis) staining upon TMEFF2 overexpression and treatment with p53 inhibitor. 178
Figure 5.8: Intracellular localization of the experimentally overexpressed TMEFF2 in MCF-7 cells. 181
Figure 5.9: TMEFF2, a-Tubulin and ATPB (mitochondrial) Immunofluorescence in MCF-7 cells. 184
Figure 5.10: Choosing the optimum time point for harvesting cells for the transcriptomics analysis. 185
Figure 5.11: Schematic representation of the methodology employed in sample preparation for the transcriptomic analysis of TMEFF2 overexpressing MCF-7 cells. 186
Figure 5.12: Confirmation of TMEFF2 overexpression in triplicate samples for microarray transcriptomic analysis. 187
Figure 5.13 (a): Summary of differentially expressed transcripts in MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'. 189
Figure 5.13 (b): Estimation of differentially expressed transcripts as part of the analysis available transcriptome. 190
Figure 5.14: Volcano plot and scatter plot for differential gene expression upon TMEFF2 overexpression in MCF-7 cells. 191
Figure 5.15: Loci of differentially expressed transcripts in TMEFF2 overexpressing MCF-7 cells. 192
Figure 5.16: Heatmaps of a subset of differentially expressed genes upon TMEFF2 overexpression in MCF-7 cells: 193
Figure 5.17: Schematic representation of the three-tier approach employed for the pathway analysis. 197
Figure 5.18: Canonical Pathways outlook upon TMEFF2 overexpression in MCF-7 cells. 198
Figure 5.19: Interferon Signalling canonical pathway outlook upon TMEFF2 overexpression in MCF-7 cells. 204
Figure 5.20: Apoptotic pathway (WikiPathway) upon TMEFF2 overexpression in MCF-7 cells. 208

Figure 5.21: (further to table 5.15) Network view of IRF7 as the predicted upstream regulator for the transcriptomic changes observed upon TMEFF2 overexpression in MCF-7 cells. 212
Figure 5.22: (further to table 5.15) Network view of TNF as the predicted upstream regulator for the transcriptomic changes observed upon TMEFF2 overexpression in MCF-7 cells. 213
Figure 5.23: (further to table 5.15) Network view of TRAIL as the predicted upstream regulator for the transcriptomic changes observed upon TMEFF2 overexpression in MCF-7 cells. 214
Figure 5.24: IPA Downstream Effect Analysis – Diseases & Biological Functions Analysis – for the gene expression changes observed upon TMEFF2 overexpression in MCF-7 cells. 215
Figure 5.25: qRT-PCR results for the verification of microarray data. 217
Figure 5.26: TMEFF2-induced apoptosis upon knockdown of STAT1, FADD, TNFRSF10A and TNFRSF10B using TMEFF2 + shRNA co-transfection. 219
Figure 5.27: TMEFF2-induced apoptosis upon stable knockdowns of STAT1, TRAIL and FADD. 220
Figure 6.1: Difference of cell death induced in CV-1 and COS-7 by overexpression of FBLN5 and TMEFF2. 239
Figure 6.2: Effects of FBLN5 and TMEFF2 overexpression on cell viability in COS-7 and CV1 cells. 241
Figure 6.3: Modified clonogenicity assay for the effects of FBLN5 and TMEFF2 overexpression on long-term cell viability in COS-7 and CV1 cells. 242
Figure 6.4: Comparison of TMEFF2 overexpression between COS-7 and CV-1 cells. (a.) 244
Figure 6.5: Comparison of FBLN5 overexpression between COS-7 and CV-1 cells. 245
Figure 6.6: Intracellular localization of FBLN5 in CV-1 and COS-7 cells. 249
Figure 6.7: Schematic representation of the comparative transcriptomic analysis in CV-1 and COS-7 cells upon FBLN5 overexpression. 251
Figure 6.8: Confirmation of FBLN5 overexpression in triplicate samples for microarray transcriptomic analysis. 252
Figure 6.9: Summary of differentially expressed transcripts. 255
Figure 6.10: Volcano plots and scatter plots for the differential gene expression analysis. 257
Figure 6.11: Heatmaps of a subset of differentially expressed genes between COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells (left) and CV-1 cells transfected with FBLN5 *vs* wild type CV- cells (right). 258
Figure 6.12: Comparison of differentially expressed genes between CV-1 cell (blue) and COS-7 (yellow) upon FBLN5 overexpression. 259
Figure 6.13: The heatmap representing a subset of differentially expressed genes between COS-7 cells transfected with FBLN5, wild type COS-7 cells, CV-1 cells transfected with FBLN5 and wild type CV- cells. 260
Figure 6.14: Canonical Pathways Comparison Analysis. 263
Figure 6.15: Upstream Regulator Comparison Analysis. 267
Figure 6.16: Network view of MYC as the predicted upstream regulator for the transcriptomic changes upon transformation of CV-1 cells (to COS-7 cells). 268
Figure 6.17: Network view of MYC as the predicted upstream regulator for the transcriptomic changes observed upon FBLN5 overexpression in COS-7 cells. 269
Figure 6.18: Network view of MYC as the predicted upstream regulator for the transcriptomic changes observed upon FBLN5 overexpression in CV-1 cells. 270

Figure 6.19: A subset of results from IPA Downstream Effects Comparison Analysis. 272

LIST OF TABLES

Table 1.1: An overview of the previously reported anticancer genes. 47
Table 2.1: Transfection parameters for MCF-7 cells, CV-1 cells and COS-7 cells in different formats. 59
Table 2.2: List of antibodies used for immunoblotting. 63
Table 2.3: List of primers used for qRT-PCR. 67
Table 3.1: List of anticancer genes used in this study. 77
Table 3.2: Results for NCBI nucleotide blast two-sequence alignment tool, BLASTN 2.2.32+ for ACG1 to ACG22. 81
Table 3.3. Correct sequence accession numbers for the genes which did not align with their corresponding NITE library reference sequences. 81
Table 3.4: Identities of anticancer genes. 96
Table 5.1: Reported functions of TMEFF2 154
Table 5.2: Hyperlinks for the time-lapse movies of the MCF-7 cells. 173
Table 5.3: Reported Intracellular localization of the TMEFF2. 180
Table 5.4: Summary of differentially expressed transcripts in MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'. 188
Table 5.5: List of ~60 most upregulated and ~30 most downregulated genes in MCF-7 cells upon TMEFF2 overexpression. 194
Table 5.6: Molecules involved in 'Interferon Signalling' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 200
Table 5.7: Molecules involved in 'Activation of IRF by Cytosolic Pattern Recognition Receptors' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 201
Table 5.8: Molecules involved in 'Role of Pattern Recognition Receptors in Recognition of Bacteria and Virus' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 201
Table 5.9: Molecules involved in 'Role of RIG1-like Receptors in Antiviral Innate Immunity' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 202
Table 5.10: Molecules involved in 'Role of Hypercytokinemia / hyperchemokinemia in Pathogenesis of Influenza' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control 202
Table 5.11: Molecules involved in 'Antigen Presentation Pathway' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 203
Table 5.12: Molecules involved in 'Apoptosis Signalling' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 205

Table 5.13: Molecules involved in 'Death Receptor Signalling' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 206

Table 5.14: Molecules involved in 'Apoptosis Signalling' IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control. 207

Table 5.15: IPA Upstream Regulator Analysis for the gene expression changes observed upon TMEFF2 overexpression in MCF-7 cells. 209

Table 5.16: Comparison of qRT-PCR and microarray data for the selective differentially expressed genes between MCF-7 cells transfected with TMEFF2 and MCF-7 cells transfected with Luciferase 'control' 217

Table 6.1: Summary of differentially expressed transcripts in (a) CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells, (b) COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells, (c) wild type COS-7 cells *vs* wild type CV-1 cells. 254

Table 6.2: A comparison of molecules involved in 'Cyclins and Cell cycle Regulation' IPA canonical pathway extracted from figure 6.14. 264

Table 6.3: A comparison of molecules involved in 'Oxidative Phosphorylation' IPA canonical pathway extracted from figure 6.14. 265

1 Introduction

Despite the leaps forward in our understanding of neoplastic transformation, development, sustenance and spread in the recent past, comparatively little progress has been made when it comes to treatment. The majority of anticancer drugs in use today were developed before or during the last quarter of 20th century when our knowledge of the biochemical and genetic basis of cancer pathophysiology was largely limited [1]. A combination of chemotherapy, radiotherapy and surgery still remains the primary therapeutic regime in the treatment of cancer. Thanks to the advancement in diagnostic techniques, combined with extensive screening programmes, we are able to detect cancer much earlier. One could infer that the apparent advancements in extending the lifespan of cancer patients could very well be a result of diagnosing the cancer earlier than we could before. Moreover, looking at incidence, survival, and mortality statistics from 1975-2010, Cho *et al*. [2] suggest that the slight improvement in the mortality rates of some cancers is attributed to the progress made in cancer diagnostics and removal of precursors rather than the therapeutics.

The lack of specificity of the therapeutic agents against cancer is currently one of the prime obstacles in effective cancer therapy. A classical definition of chemotherapy is the use of chemical compounds to kill or inhibit the growth of foreign organisms in the body while sparing the normal function of the body itself [3]. Paradoxically, for the majority of therapeutic agents currently being used to treat cancers, the specificity against cancer cells is primarily quantitative rather than qualitative. That is, cancer cells have a similar biochemical profile as the normal cells of the body, but certain processes occur at a higher rate in the cancer cells when compared to the normal cells. The only way a great majority of currently available therapeutic agents impart their cancer specificity is by targeting this higher rate, thus enhancing cytotoxicity against cancer cells as compared to normal ones [4]. This ‘crude approach’ comes with the obvious drawback of generalized cytotoxicity, leading to an increase in patient morbidity and sometimes even mortality [5]. Common side effects include hair loss, nausea, vomiting, mouth ulceration and bone marrow depression. The latter complicates matters for worse, as bone marrow depression, in combination with drug-line infections, commonly leads to neutropenic sepsis. Typically, the oncologist remains in a constant struggle to find the right therapeutic window in order to balance the advantageous effect of chemotherapy with the harms caused to the patient. In this respect, the lack of specificity imposes the biggest challenge in the administration of a therapeutically effective dose of cytotoxic drugs. Even if a balance is struck in

favour of the patient, the selection pressure applied on the constantly evolving tumour and the development of resistance as a result are not an uncommon occurrence. Of the few fortunate ones who survive the ordeal of chemotherapy with successful remission of the tumours, the continued systematic ailments of virtually every organ of the body remain as barriers in patients' well-being in the long run. At the moment, we do not have a choice other than to accept these as the price for controlling a fatal disease. In short, as Grimm and Noteborn [6] put it, '*every cell can be eradicated at any given time. For cancer treatment, the most crucial aspect is to limit this destruction to tumour cells'* (p.88).

However, one must accept that it is not all bad news. Chemotherapeutic agents offer some degree of specificity; for example, cisplatin is particularly effective against testicular cancer while carboplatin is the prime treatment option for ovarian cancer [7]. Also, chemotherapy does have its advantages; going back to the study by Cho *et al.* [2], a slight reduction in the incidence-based mortality of the distant breast cancer could be credited to more efficacious treatment regimens. Furthermore, Morgan *et al.* [8] attributed a minor improvement of 2.1% (2.3% in Australia) in the overall 5-years survivability of cancer patients to the curative and adjuvant chemotherapy. However, when compared to advances made in the understanding of cancer pathophysiology, these improvements are negligible. The low molecular weight drugs which fall under the umbrella of chemotherapy belong to five broad categories: alkylating agents (such as cisplatin), antimetabolites (such as methotrexate), topoisomerase inhibitors (e.g. etoposide), cytotoxic antibiotics (e.g. doxorubicin), and alkaloids (e.g. Taxol). All of these induce, directly or indirectly, the regulated cell death (apoptosis) and there can only be a limited number of targets to induce apoptosis. The recent advancements in novel drug discovery and development have been directed to increasing the specificity rather than efficacy of anticancer drugs [9]. The development of tyrosine kinase inhibitor imatinib [10] for the treatment of chronic myeloid leukaemia and the monoclonal antibody against Human Epidermal Growth Factor Receptor 2 (HER2) trastuzumab for metastatic breast cancer in the 1990s, potentiated the efforts towards targeted anticancer therapy. It was later followed by developing small molecule compounds and monoclonal antibodies against a variety of targets. The most salient ones include the epidermal growth factor receptor (EGFR), vascular endothelial growth factor receptor(s) (VEGFR) and platelet derived growth factor receptors (PDGFR) [11]. Conversely, the epigenetic enzymes like histone deacetylase (HDAC) and histone methyltransferase (HMT) are emerging as potentially promising targets for targeted cancer therapies. Despite the earlier success of targeted therapies, it is now clear

that among an unselected population of patients, the response rate of targeted therapies is usually limited to 10-20% [11]. A substantial proportion of tumours is *ab initio* refractory towards such therapies. Furthermore, owing to the rapid development of resistance and the presence of cancer stem cells (CSCs), the clinical response towards targeted therapies as a whole (in unselected population) is typically 6-12 months [11]. Lately, the focus has been shifting towards the innovative immunotherapies and personalized medicine (i.e., genetic profiling of tumours). Nonetheless, further avenues need to be explored for the development of cancer specific therapies.

A recent concept towards such efforts is that of anticancer genes, a term coined by Grimm and Noteborn [6]. They defined anticancer genes as '*when ectopically overexpressed they specifically destroy tumour cells without harming normal cells'* (p.88). The most well-known example of an anticancer gene is TRAIL. Several other such genes have also been reported. Some have viral origins like apoptin, parvovirus NS1 and E4orf44. Brevinin-2R was isolated from frog. In addition, several anticancer genes from human and other mammalian genomes have also been reported. Examples include MDA7, PAR4 and Orctl3 [7].

1.1 Anticancer genes: mechanism of action

The tumour specific function of anticancer genes is based on the principle of synthetic lethality. Synthetic lethality is defined as a type of genetic interaction where the co-existence of two genetic alterations results in cell or organismal death [7]. The phenomenon was first described in experiments involving crossing *Drosophila melanogaster*. The term 'synthetic' is used in the context of combination.

According to Weinberg [1], '*cancer is really a disease of aberrant signal processing*' (p.176). In cancer cells, the activation of oncogenes and the loss of tumour suppressor genes in combination with widespread genetic instability result in the distortion of otherwise smooth signalling circuitry, i.e., several normally unconnected networks are now connected and *vice versa*. Although these distortions provide the crucial survival and proliferation advantages to cancerous cells, they also pose certain vulnerabilities for them. One of the ways in which anticancer genes impart their tumour specific effect is by activating certain signalling pathways and, as a result, they create conflict(s) in the distorted signalling networks of cancer cells [6, 7]. Most of the time, the signalling networks activated by the anticancer genes are pre-existing; it is only in the distorted signalling circuitry of cancer cells

that their activation is lethal. Besides, some anticancer genes exhibit neomorphic functions, specifically in cancer cells, which are very different from their physiological functions in the normal cells. A good example of such phenomenon is that of Orctl3. In a normal cell, Orctl3 functions as a cation transporter, but when overexpressed in HRAS transformed CV-1 cells, Orctl3 inhibits stearoyl-CoA desaturase (SCD-1) [12], thus disrupting the fatty acid metabolism pathway crucial for the rapidly dividing cells.

1.2 Advantages of anticancer genes over conventional therapy

Tumours are comprised of a very heterogeneous population of cells. One of the advantages of anticancer genes like apoptin is their ability to impart multiple signalling effects in the cell, i.e., conflicts could be introduced at several points throughout the distorted signalling circuitry [13]. Therefore, in comparison to the small molecular weight compounds, the activation of targets can have a pleotropic effect, hence imparting synthetic lethal interference on multiple levels in a heterogeneous population of tumour cells. Anticancer genes present another distinct advantage, as in comparison to the small molecular weight compounds whose actions are generally inhibitory, the anticancer genes can activate their targets. In the case that the target lies on the cell membrane, not only does it present the benefit of a more upstream activation of a signalling cascade, it also circumvents the problem of getting the drug into the cell. It is important to mention here that the mode of action of anticancer genes is not limited only to activation; as mentioned earlier, Orctl3 inhibits stearoyl-CoA desaturase (SCD-1) in HRAS-transformed CV-1 cells. In addition, the majority of therapeutic agents used in anticancer therapy induce apoptosis. In cancer cells where the apoptotic machinery is generally defective, the chemotherapy ends up killing more normal than cancer cells. Anticancer genes like E4orf4 can cause necrotic cell death and the overexpression of the rodent parvovirus NS1 leads to lysosome-dependent cell death. The added complexity of anticancer genes compared to low molecular weight compounds is itself beneficial. As cancer cells can develop resistance through rapid efflux of the small molecular weight compound [14], this would be very difficult in the case of the highly complex protein products of anticancer genes (no data is available so far to back this hypothesis). Last but not least, the complexity of protein products of anticancer genes makes their actions more specific as compared to the pharmacological low molecular weight compounds.

1.3 Anticancer genes *vs* tumour suppressor genes

Since the mode of action of anticancer genes is *a priori* unknown, a convenient hypothesis on how anticancer genes impart their tumour-specific effect is through the reconstitution of downregulated network(s) which govern cell-death. As discussed earlier, this is not always the case, which brings us to an obvious and most popular question – what is the difference between anticancer genes and tumour suppressor genes? Anticancer genes kill cancer cells actively when overexpressed. Tumour suppressor genes, by contrast, act passively by working against tumour formation at their endogenous expression levels [6]. Moreover, the tumour suppressor genes are recessive in nature and malignancy is aided only when both of the alleles are inactivated; as Weinberg puts it [1], '*the very existence of a tumour suppressor gene becomes apparent only when it is absent*' (p.271). The loss of heterozygosity (LOH) is a hallmark of tumour suppressor genes. Anticancer genes, on the other hand, do not have to be necessarily removed or downregulated in order to achieve successful tumorigenesis and/or metastasis. This distinction is apparent in the case of viral anticancer genes or the ones from other non-human species, but when it comes to human or even mammalian genes, the distinction can become diffuse. With recent advances in epigenetics, we have come to know that LOH is not the only method through which cancer cells disable the tumour suppressor genes. For example, epigenetic alteration of the *bona fide* tumour suppressor gene RUNX3 is the only known source of its silencing in gastric cancer [1]. Besides, even the viral anticancer gene apoptin reinstates the level of ceramide, which is otherwise reduced in the malignant cells; in other words, anticancer genes can activate tumour suppressor activity. However, it is only upon the selective phosphorylation in malignant cells that apoptin undertakes this action [6]. Moreover, tumour suppressor genes can also act as anticancer genes when they encounter their synthetic lethality partner(s); PAR4 is a good example of this phenomenon [15]. Conversely, the common criticism faced during deciphering the anticancer genes from the tumour suppressor genes is that a different term is simply invented to bypass the strict criteria that a gene has to fulfil in order to be designated as a tumour suppressor gene. In my view, an even stricter criterion has to be fulfilled in order to be designated as an anticancer gene – its ectopic overexpression has to cause cell death in cancer cells while leaving the normal cells unharmed. As shown by Stempien-Otero *et al.* [16], overexpression of p53 causes apoptosis in normal endothelial cells. A related critical question regarding the synthetic lethality being the core mode of action of anticancer genes is the differing gene overexpression levels between the normal and cancer cells. Such criticism does hold weight because transformed or cancer cells divide

faster in cell culture, tend to take up transfection reagents more easily and usually have an amplified transcription and translational machinery when compared to the normal and/or immortalized cell lines. However, there could be counter arguments: numerous studies regarding anticancer genes involve transforming the normal cells using stably expressing oncogenes and comparing the effects of overexpressed anticancer genes between transformed cells and the parental normal counterpart. In this case, firstly, the relative expression levels are not very different and, secondly, the differential expression level is actively taken into context and proven not to be the reason behind the tumour specific effects of anticancer genes (the studies regarding Orctl3 are a good example where the Orclt3 fused with the green fluorescent protein (GFP) was used to alleviate the possibility of differing protein expression [17]). On the other hand, the apoptotic mechanisms in cancer and transformed cells tend to be defective or downregulated, which should in theory compensate for the functional effects of relatively higher expression levels of anticancer genes, i.e., a substantially lower expression level of pro-apoptotic factor should be enough to kill a normal cell where the apoptotic machinery is intact. In addition to this, there is a possibility that as a result of distorted signalling circuitry, the cancer cell might fail to destabilize the mRNA or protein of a particular gene, thus resulting in its accumulation and the subsequent functional effect, whereas in the normal cell, the overexpression of the same gene would trigger feedback mechanisms (or similar), resulting in its downregulation through transcriptional block, mRNA destabilization or protein degradation. Such phenomenon appears to be in part responsible for the differential killing of transformed cells by TMEFF2 as shall be discussed in the last chapter. This phenomenon still falls under the umbrella of synthetic lethality.

1.4 Cell death

Programmed and regulated cell death is a crucial process during both embryonic and adult life. During embryonic development, cells are normally produced in excess and later on removed in an orderly fashion; for example, neurons are produced in excess of 50% and only the ones which manage to accomplish the right connections survive while the rest are eliminated [18]. A rather more captivating example is the formation of tetrapod fingers and toes where the tissue between the digits is trimmed off. Similarly, during the adult life regulated cell death serves numerous crucial roles in a multitude of physiological processes including tissue homeostasis, immunity, inflammation and onco-suppression. Aberrations in regulated cell death lead to several pathologies: abnormally high cell death leads to diseases such as neurogenerative disorders and myocardial infarction, whereas malfunctions in

regulated cell death machinery result in uncontrolled cell growth ultimately leading to ailments like cancer and some autoimmune disorders [19].

In 1973 Schweichel and Merker [20] categorized cell death in three different phenotypic forms based on the morphological changes exhibited in the process of dying:

I. Apoptosis or type I cell death: cell exhibits chromatin condensation (pyknosis), nuclear fragmentation (karyorrhexis), cytoplasmic shrinkage and plasma membrane blebbing. At the end of the process, the dead cell forms small vesicles (apoptotic bodies) for the degradation in the phagocytic cells.
II. Autophagy or type II cell death: this type of cell death is marked by an extensive cytoplasmic vacuolization which grows in size and number, ultimately forming small bodies for phagocytic uptake.
III. Necrosis or type III cell death: necrosis involves distension rather than shrinkage of cytoplasm. As a result, the plasma membrane collapses releasing the cell content into the environment. Moderate chromatin condensation and swelling of cytoplasmic organelles are also observed but the marked distinction from type I and type II cell death is the absence of apoptotic bodies and the subsequent phagocytic uptake.

These morphological definitions are still in use but since 2005, the Nomenclature Committee on Cell Death (NCCD) gathers periodically to reach a consensus in describing forms of cell death taking into account the cell morphologies, genetic and biochemical processes involved in the process and distinctions between the causative and accessory aspects of cell death. Over the years, NCCD guidelines appear to have transitioned gradually from relying on morphological features to cell death signalling. In 2009 NCCD guidelines [21] were mostly based on the morphologies. In 2012 [22] some emphasis was laid on biochemical changes and the 2018 definitions are primarily based on the signal transduction pathways involved in the process. The definitions provided by NCCD in 2018 are included in box 1.1. I find the NCCD guidelines to be the most comprehensive source of up-to-date information regarding cell death. Therefore, I am including a brief synopsis of the relevant sections from the 2018 NCCD guidelines [23] herein.

Regulated cell death (RCD) is an umbrella term for the self-destructions unique to multicellular life forms but also found in unicellular eukaryotes living in colonies as well as in *E. coli*. Compared to

accidental cell death (ACD), RCD relies on dedicated cellular machinery which is capable of being modulated. When RCD is undertaken in the absence of environmental perturbation, i.e., as a result of built-in physiological signalling, it is referred to as programmed cell death (PCD). The environmental or intracellular stressors normally trigger a coping mechanism where an adaptive response is generated to restore the cellular homeostasis. Should such adaptive response fail to elicit the desired effect, the subsequent RCD enables the dying cells to expose or release the molecules that alert the organism or colony about potential danger through signals commonly known as damage-associated molecular patters (DAMPs).

1.4.1 Apoptosis

Greek for shedding of leaves and originally a morphological term, apoptosis is the most well-known and well-researched form of RCD. An overview of apoptosis signalling is provided in figure 1.1. Apoptosis is classically categorized into intrinsic and extrinsic apoptosis.

1.4.1.1 Intrinsic apoptosis

Intrinsic apoptosis is initiated by a variety of stimuli including DNA damage, replication stress, cytoskeletal alterations, mitotic defects, reactive oxygen species (ROS) overload, endoplasmic reticulum stress (ER stress) and growth factor withdrawal. Mitochondria occupy the centre stage when it comes to intrinsic apoptosis signalling and the rest of apoptotic signalling circuitry forms an intricate web around this central organelle. Among these signalling molecules are the family members of B-cell lymphoma gene-2 (BCL2) proteins. BCL2 family comprises of several proteins sharing the BCL2 homology (BH) domains. These are classified into three main categories: (i) the anti-apoptotic BCL2 proteins include BCL2 itself, B-cell lymphoma-extra large (BCL-X_L, also known as BCL2L1), myeloid cell leukaemia sequence 1 (MCL1), BCL2-like 2 (BCL2L2) and BCL2 related protein A1 (BCL2A1). These proteins contain four BH domains (BH1, BH2, BH3 and BH4) and a transmembrane domain. (ii) The pro-apoptotic BCL2 proteins include BCL2 Associated X (BAX), BCL2 Antagonist/Killer 1 (BAK) and BCL2 related ovarian killer (BOK). They contain three BH domains (BH1, BH2 and BH3) and a distinct transmembrane domain which anchors these onto the outer mitochondrial membrane. (iii) The pro-apoptotic BCL2 proteins which contain only the single BH3 domain include BH3 interacting domain death agonist (BID), BCL2-Like 11 (BCL2L11, also known as BIM), BCL2 interacting killer (BIK), BCL2 associated agonist of cell death (BAD), BCL2 modifying factor (BMF), harakiri BCL2

interacting protein (HRK), NOXA and PUMA. Among these, all but BID has a transmembrane domain. The BCL2 family proteins are in turn regulated by a plethora of signalling molecules involving several cellular processes which modulate them at their transcriptional, translational and post-translational levels. The execution of intrinsic apoptosis depends on the balance between the activities of pro- and anti-apoptotic BCL2 family members. When the pro-apoptotic side wins out, the crucial event in the intrinsic apoptosis is the mitochondrial outer membrane permeabilization (MOMP), and the subsequent release of cytochrome *c* into the cytosol. The latter stabilizes apoptotic peptidase activating factor 1 (APAF1) polypeptides into a protein complex called the apoptosome. The apoptosome converts procaspase-9 into active caspase-9. Caspase-9 belongs to the family of proteases known as effector caspases (cysteine-dependent aspartate-directed proteases). The activated caspase-9 later catalyses the proteolytic conversion of procaspase-3 and procasepase-7 into their active counterparts. These are known as executioner caspases and have widespread targets throughout the cell – the digestion of which lead to the morphological features characteristic of apoptosis. Even after the caspase activation, the cell is capable of averting apoptosis through an inhibitor of apoptosis (IAP) family member, X-linked inhibitor of apoptosis (XIAP), which binds and inactivates the caspases. Such anti-apoptotic factors are in turn neutralized by the IAP-binding mitochondrial protein SMAC/DIABLO, which is also released into the cytoplasm, along with the cytochrome *c*, upon MOMP. In the generalised scheme of events, MOMP is executed by the formation of a pore in the outer mitochondrial membrane by pro-apoptotic BCL2 proteins, BAX and BAK. Notably, MOMP ultimately leads to loss of mitochondrial membrane potential. Moreover, a special form of intrinsic apoptosis is triggered by loss of integrin-dependent attachment to the extracellular matrix, commonly known as anoikis. Cancer cells employ a number of mechanisms to counter anoikis in order to become metastatic. Therefore, the NCCD [23] defines intrinsic apoptosis as '*a form of RCD initiated by perturbations of the intracellular or extracellular microenvironment, demarcated by MOMP and precipitated by executioner caspases, mainly CASP3*' (p.489). It is important to mention here that there are certain exceptions to the general scheme of signalling of intrinsic apoptosis. These include (but are not limited to): (i) prolonged inhibition of caspases promotes switching to other modes of RCD; (ii) some cells can survive MOMP affecting a limited number of mitochondria; (iii) BAK and BAX are not always required for MOMP, some specific lipids like ceramide can also permeate the outer mitochondrial membrane.; and (iv) BAK and BAX have been suggested to permeabilise ER membranes subsequently resulting in intracellular Ca^{++} release.

1.4.1.2 Extrinsic Apoptosis

In comparison to intrinsic apoptosis, extrinsic apoptosis is initiated by perturbations of only the extracellular environment. In order to drive the extrinsic apoptosis, the cell has generally two types of receptors: death receptors and dependence receptors. The activation of death receptors is dependent upon binding of their cognate ligand(s). These include FAS (APO-1) receptors and the tumour necrosis factor receptor (TNFR) superfamily. The latter includes TNFR1 and the TRAIL receptors DR4 and DR5. In a generalised scheme of events, the FAS and TRAIL receptors trimerize and undergo conformational changes upon binding of their receptors. The conformational change brings about the recruitment of the FAS-associated *via* death domain protein (FADD), which in turn brings in caspase-8 ultimately leading to its autoactivation. This multiprotein complex comprising ligands, the transmembrane receptors and the associated proteins is called the death-inducing signalling complex (DISC). The active caspase-8 leads to proteolytic activation of downstream executioner caspases. The apoptotic modulation in this pathway comes through c-FLIP, an inactive homologue of caspase-8 which competes with caspase-8 for binding the DISC. In the case of TNFR1 signalling, things are slightly more complex. The binding of ligands brings about the recruitment of an adapter protein, TRADD, which further recruits a host of proteins resulting in the formation of multiprotein complexes I and II. The formation of complex I generally culminates in ubiquitinoylation of IKβ, leading to the activation of NF-κB and the subsequent anti-apoptotic, pro-proliferative cell state. The typical complex I proteins include TNF receptor-associated factor 2 (TRAF2), TRAF5, c-IAP1, c-IAP2, receptor interacting serine/threonine kinase 1 (RIPK1), linear ubiquitin chain assembly complex (LUBAC), SHANK associated RH domain interactor (SHARPIN), HOIL, HOIP and mitogen-activated protein kinase 7 (MAP3K7). On the other hand, the formation of complex II involves the recruitment of FADD and the subsequent caspase-8 activation. In the majority of cancer cells and some normal cells like pancreatic β cells and hepatocytes, the caspase-8 activation is not sufficient in bringing about the activation of executioner caspases and the role of BH3-only domain protein BID comes to light. Caspase-8 causes the cleavage driven conversion of BID to truncated BID (tBID). The latter in turn translocases to outer mitochondrial membrane and activates BAK and BAX, subsequently leading to MOMP and eventually culminating in the activation of executioner caspases. As already mentioned, this is a very generalized understanding of external apoptotic signalling. A salient divergence in this scheme is the relatively recent research into the pro-apoptotic role of RIPK1. Previously, the understanding was that RIPK1 contributes to the activation of NF-κB following its polyubiquitinylating

by c-IAP1, c-IAP2 and LUBAC. It is now apparent that the aberrations in the RIPK1 ubiquitinoylation and/or its impaired phosphorylation (IKKα-IKKβ mediated and NEMO aided RIPK1 phosphorylation) can lead to the caspase-8 activation through the formation of RIPK1-FADD-caspase-8 complex (see Ref [24] for details). This aberrant ubiquitinoylation and impaired phosphorylation of RIPK1 can also result in an increase in its kinase activity, consequently leading to RIPK1 driven necroptosis (discussed later). Further, in some cells, the toll-like receptor proteins family can induce caspase-8 activation [23].

The activation of dependence receptors relies on the absence of their ligands rather than the presence, i.e., the dependence receptors activate a distinct lethal cascade when the level of their ligands falls below a certain threshold. Moreover, the dependence receptors normally promote proliferation, survival, and differentiation under the physiological conditions. The mechanism of lethal signalling by dependence receptors is not fully elucidated in most cases, but there are some interesting examples: (i) the sonic hedgehog receptor patched 1 (PATCH1), interacts with DRAL to bring about the assembly of caspase-9 activation complex, (ii) the Nitrin 1 receptor UNC5B mediates the dephosphorylation of DAPK1 by PP2A and the subsequent p53 dependent apoptosis. In light of the above, the NCCD [23] defines extrinsic apoptosis as '*a type of RCD initiated by perturbations of the extracellular microenvironment that are detected by plasma membrane receptors, propagated by caspase-8 (with the optional involvement of MOMP), and precipitated by executioner caspases, mainly caspase-3'* (p.489).

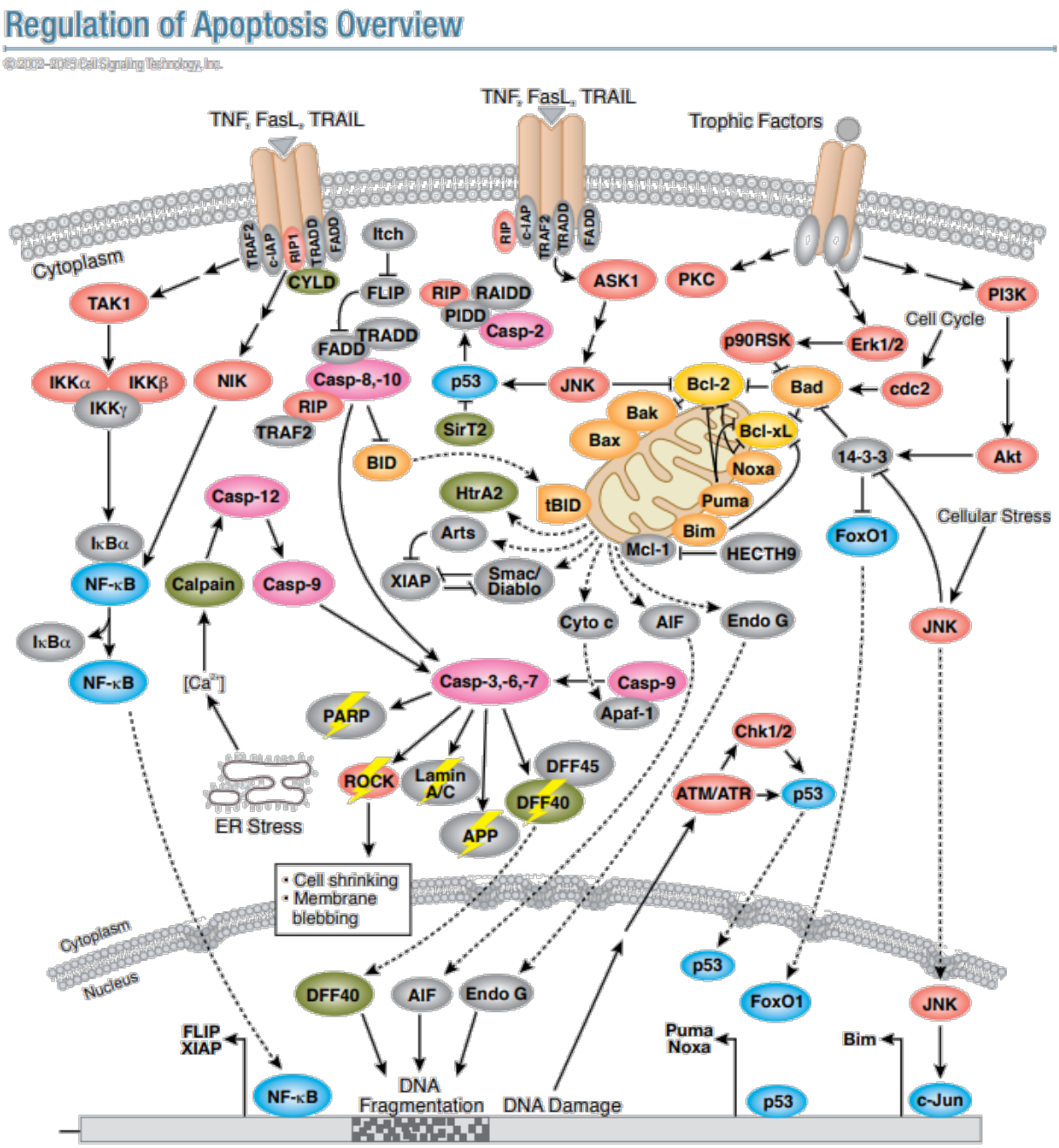


Figure 1.1: **Regulation of Apoptosis Overview**, Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com).

1.4.2 Mitochondrial permeability transition (MPT)-driven necrosis

The permeability transition pore complex (PTPC) is a supramolecular complex which resides at the junction of the inner mitochondrial membrane (IMM) and the outer mitochondrial membrane (OMM). The opening of PTPC results in an abrupt loss of IMM impermeability to small solutes leading to a sudden dissipation of mitochondrial membrane potential, consequently culminating in necrotic

RCD. The regulation of PTPC control is an area of active research; however, no clear pathways have been reported. To this date, the only crucial factor observed in the process is the peptidylprolyl isomerase F commonly known as cyclophilin D. The noteworthy physical interactors of PTPC which have been shown to regulate the MPT-driven necrosis include both the pro- and anti-apoptotic members of BCL2 family and p53, with the latter having been reported to drive MPT-driven necrosis through its interaction with cyclophilin D [23].

Conversely, in 2012 NCCD's definitions of cell death [22], extrinsic apoptosis was defined to be only caspase dependent, and intrinsic apoptosis was further categorised into caspase-dependent and caspase-independent. If apoptosis could be abolished by inhibition of caspases, it was defined as extrinsic apoptosis that could also potentially fall into the intrinsic criteria. The caspase-independent intrinsic apoptosis was reported to be executed by proteins, poly (ADP-ribose) polymers to apoptosis inducing factor mitochondria associated 1 (AIF) and migration inhibitory factor (MIF). MPT-driven death was also included in intrinsic apoptosis. Later, in 2018, NCCD abolished these caspase-dependent/independent definitions and now it is only intrinsic and extrinsic apoptosis, both of which are caspase-dependent. MPT-driven necrosis and necroptosis are now separate forms of cell death. Moreover, the caspase-independent cell death involving the MIF and AIF through activation by PRAP1 is now categorized separately as parthanatos (discussed below) [22].

1.4.3 Necroptosis

Necroptosis is another well-modulated form of RCD (at least in specific settings) where the dying cells manifest necrotic morphology. Previously considered as an alternative adaptive response to the failure of stress responses, it is now clear that necroptosis is implicated in several physiological and pathological processes. Necroptosis is triggered by specific plasma membrane and intracellular receptors which include (but are not limited to) TNFR1, toll-like receptor 3 (TLR3), TLR4 and Z-DNA binding protein 1 (ZBP1). The execution of necroptosis depends on the mixed lineage kinase domain-like pseudokinase (MLKL), culminating in its oligomerization and translocation to the plasma membrane where it binds specific phosphatidylinositol phosphates and cause plasma membrane permeabilization. Upon simultaneous activation of TNFR1 and caspase-8 inhibition, RIPK1 (which canonically contributes to the activation of NF-κB) phosphorylates RIPK3, and in turn RIPK3 phosphorylates and activates MLKL. In addition, aberrations in RIPK1 ubiquitinoylation and phosphorylation have also been shown to have the same effect (i.e., activation of RIPK3 and MLKL).

The activation of MLKL involves the formation of an amyloid-like complex comprising of RIPK1, RIPK3 and MLKL– designated as a necrosome. The negative regulators of necrosomes include protein phosphatase Mg2+/Mn2+ dependent 1B (PPM1B), aurora kinase A (AURKA) and A20 (also known as TNFAIP3). In addition to RIPK1, RIPK3 can also be activated by TRIF upon TLR3, TLR4 and ZBP1 activation. These regulatory mechanisms are a generalized scheme of affairs; the exact signalling factors and processes involved in necroptosis are very complex, not well understood and beyond the scope of this thesis. TNFR1 activation is by far the best characterized regulator of necroptosis. NCCD [23] defines necroptosis as '*a type of RCD triggered by perturbations of extracellular or intracellular homeostasis that critically depends on MLKL, RIPK3, and (at least in some settings) on the kinase activity of RIPK1*' (p.489).

1.4.4 Ferroptosis

It is a form of RCD which is independent of caspases, necrosome or autophagy and manifests necrotic morphology. Ferroptosis is caused by specific intracellular perturbations characterised by ROS generation, culminating in a severe iron-dependent lipid peroxidation. Several regulators of ferroptosis are emerging, the most salient being the reduced glutathione (GSH)-dependent enzyme glutathione peroxidase 4 (GPX4), an endogenous inhibitor of ferroptosis [23].

1.4.5 Pyroptosis

Mainly related to innate immunity, it is a caspase-dependent form of RCD which manifests very peculiar morphology, i.e., a form of chromatin condensation that is different from apoptosis, accompanied by cell swelling and plasma membrane permeabilization. Pyroptosis was first thought to be triggered by pro-inflammatory caspase-1 and being primarily involved in monocyte and macrophage RCD. Later, it emerged to be manifested by several cell types and it has a key role in innate immunity against intracellular pathogens. Moreover, pyroptosis is also involved in lethal septic shock induced by LPS. At the molecular level, LPS from invading pathogens bind the effector caspases 11, 4 and 5 resulting in their oligomerization and activation. Pyroptosis is executed upon proteolytic cleavage and the subsequent activation of gasdermin A (GSDM). Activated GSDM molecules translocate to the plasma membrane where they bind specific phosphoinositides. This binding triggers GSDM oligomerization resulting in the formation of a pore and the subsequent plasma membrane permeabilization. The bacterial plasma membrane is also permeabilized through the same

mechanism by immune cells. A few noteworthy exceptions to this generalized scheme of events include the following: (i) caspase-1 and caspase-3 are also capable of activating GSDM in some instances; (ii) pyroptosis is often related to IL-1β and IL18 secretion; and (iii) interferon α and γ receptor signalling also contributes to pyroptosis through transactivation of caspase-11 and expression of IFN-inducible GTPase (IRGB10) which increases cytosolic LPS by mediating the lysis of vacuoles containing Gram-negative bacteria [23].

1.4.6 Parthanatos

It is a caspase-independent form of RCD driven by hyperactivation of poly(ADP-ribose) polymerase1, (PARP1). The resulting poly(ADP-ribose) polymers and poly(ADP-ribosyl)ated proteins accumulate around the mitochondrial membrane and cause mitochondrial membrane potential dissipation and MOMP. In this process a key factor, i.e., the apoptosis-inducing factor mitochondria associated 1 (AIF), which normally resides in the mitochondrial intermembrane space, is released into the cytoplasm. AIF translocates to the nucleus and causes widespread DNA fragmentation and chromatin condensation. Macrophage migration inhibitory factor (MIF) catalyses the process of DNA cleavage by AIF. Additionally, PARP1 hyperactivation causes bioenergetic collapse (because of NAD^+ and ATP depletion), where hexokinase 1 (HK1) in association with the poly(ADP-ribose) polymers inhibits glycolysis. Interestingly, TRAIL-driven activation of RIPK1 and RIPK3 has been suggested to stimulate the enzymatic activity of PARP1. Parthanatos is reported to occur as a consequence of prolonged exposure to alkylating agents, inflammatory damage, hypoxia, hypoglycaemia, and oxidative stress [23].

1.4.7 Entotic cell death

A very interesting phenomenon of entosis is a type of cell cannibalism where a viable cell is engulfed by a non-phagocytic cell of the same (homotypic) or different (heterotypic) type. In the (frequent) case that the internalization results in the demise of the internalized cell, it constitutes a special form of RCD referred to as entotic cell death. The elimination of internalized cell death relies on an autophagy-related process where the internalized cells containing vesicles are fused with lysosomes, degraded, and the nutrients recovered by the engulfing cell. Entosis is primarily triggered by the detachment of epithelial cells and loss of integrin signalling. Entotic cell death has been proposed as an onco-suppressive mechanism whereby the rapidly dividing cells within a tumour compete for

space and nutrients. On the other hand, entosis can also favour tumour progression by enhancing aneuploidization and polyploidization. Moreover, KRAS and RAC1 activation has been shown to confer advantage to the engulfing cell through myosin downregulation. Additionally, in conditions of glucose withdrawal, the cells with elevated levels of 5' adenosine monophosphate-activated protein kinase (AMPK) succumb more readily to entosis by cells with comparatively low AMPK activity. This phenomenon not only favours the engulfing cells with nutrients but also enhances the selection of cells which are more adaptable to the hostile tumour environment.

The signalling of entosis involves the cytoskeletal proteins and their regulators as well as the adhesion proteins, their regulators and their downstream signalling. Lastly, entosis does not always end in the demise of the internalized cell, and in certain circumstances the internalized cell remains not only viable but also capable of proliferation inside the host or upon its escape [23].

1.4.8 Lysosome-dependent cell death

It is a form of RCD characterized by lysosomal membrane permeabilization (LMP) and the consequent release of proteolytic enzymes of the cathepsin family into the cytoplasm. In terms of signalling, the regulation of lysosome-dependent cell death is not fully elucidated. According to some reports, LMP occurs after MOMP while in other settings lysosomes permeabilize before mitochondria through the pore forming activity of BAX. Other factors causing LMP include (but are not limited to) (i) ROS, through enhanced lipid peroxidation or activation of lysosomal Ca^{++} channels, (ii) lysosomotropic agents, (iii) DNA damage regulated autophagy modulator 1 (DRAM1) in HIV1-infected T cells, and (iv) signal transducer and activator of transcription 3 (STAT3) during involution of mammary glands by upregulating cathepsin B and L. In either case, once in the cytoplasm, cathepsins catalyse the proteolytic activation of pro-apoptotic BID and BAX and inactivation of anti-apoptotic XIAP, thus potentiating the intrinsic apoptotic processes. However, lysosome-dependent cell death does not necessarily involve MOMP or caspases, neither does it necessarily result in apoptotic morphology. NCCD [23] defines lysosome-dependent cell death as a '*form of RCD demarcated by primary LMP and precipitated by cathepsins, with optional involvement of MOMP and executioner caspases*' (p.489).

1.4.9 Autophagy-dependent cell death

Autophagy is a physiological and adaptive process which involves specific marking of cellular cargo, its engulfment in autophagosome vesicles and subsequent degradation of cargo through the fusion of autophagosomes with lysosomes. Numerous autophagic proteins have been reported which are typically designated as ATGs (autophagy-related genes). In specific circumstances (mainly during embryonal/organismal development), cells employ autophagic machinery for the purpose of programmed cell death. NCCD advises caution in the use of the term (autophagy-dependent cell death), as the detection of adaptive autophagy in stressed cells could be easily confused with autophagy being the sole causative factor of cell death. Moreover, NCCD also advises against the use of this term when autophagy is a co-contributor in other forms of RCD. In some circumstances, autophagy-dependent cell death can also proceed independently of several otherwise crucial ATGs. In addition, during *D. melanogaster* development, autophagic and apoptotic apparatuses work hand in hand to bring about the RCD. NCCD [23] defines autophagy-dependent cell death as '*a form of RCD that depends on the autophagic machinery (or components thereof). To avoid confusion, this term should be consistently avoided (i) in the absence of robust experimental evidence mechanistically linking RCD to (ideally more than one) components of the autophagy apparatus, as well as (ii) when pharmacological or genetic manipulations of the molecular machinery for autophagy impact on other RCD subroutines'* (p.489).

Conversely, in the 2012 NCCD guidelines [22], autophagic cell-death could be designated in the case that the cell death is abolished upon the pharmacological or genetic inhibition of autophagy. Moreover, NCCD recommended using LC3 lipidation or SQSTM1 (p62) degradation as concurrent autophagy assays [23].

1.4.10 Immunogenic cell death

The RCD is designated to be immunogenic if it results in the activation of an adaptive immune response through releasing a series of DAMPs. The phenomenon has recently attracted substantial attention in the context of cancer therapy. Six DAMPs have been implicated in the immunogenicity of RCD, namely, calreticulin, ATP, high-mobility group box 1 (HMGB1), type I interferons, cancer-cell derived nucleic acids and annexin A1 (ANXA1). Very briefly, calreticulin translocates from ER to outer leaflet of plasma membrane through specific signalling events, acts as an "eat me" signal for

phagocytic immune cells and is involved in T cell priming. ATP is released downstream of caspase activation; the process is aided by concurrent autophagy and ATP acts as a “find me” signal for the immune cells. CD39 expressed on regulatory T cells (immune-tolerance inducing) lower the levels of extracellular ATP. HMGB1 brings about immunogenicity by binding pattern recognition receptors. Type I interferons are produced by cancer cells through (i) the activation of TLR3 receptors after exposure to endogenous dsRNA and (ii) the activation of cyclic GMP-AMP synthase (cGAS) by dsDNA exposure (including mitochondrial DNA). In addition to the stimulation and attraction of interferon alpha receptor 1 (IFNAR1) expressing immune cells, type I interferon also activates an autocrine/paracrine signalling in cancer cells whereby interferon stimulated genes (ISGs) are expressed. These include a variety of pro-inflammatory cytokines and chemoattractant(s). Moreover, the release of dsDNA and RNA molecules by dying cells is also taken up by the immune cells, thus resulting in a potent type I interferon response. Interestingly, the activated caspase-3 inhibits type I interferon response triggered upon the release of mitochondrial DNA, while some other caspases also suppress immunogenicity through several mechanisms [23].

1.4.11 Cellular senescence, mitotic catastrophe and terminal differentiation

Lastly, NCCD classifies cellular senescence, mitotic catastrophe and terminal differentiation as non-lethal processes and therefore recommends excluding them from RCD [23]. These processes, however, are not to be ignored and I shall discuss them very briefly.

Cellular senescence is a pathophysiological state where cells are devoid of their proliferative capability but remain metabolically active and viable. The senescent cells are marked by peculiar morphological and biochemical characteristics. The former includes flattening, increase in size, extensive intracellular vacuolization and the presence of senescence-associated heterochromatic foci (SHAF). The latter comprises an increase in β-galactosidase activity and a marked downregulation of multiple cyclin-dependent kinases. During embryonic development and tissue repair, cellular senescence appears to precede and sometimes contribute to RCD. Moreover, cellular senescence presents a cellular defence mechanism against sublethal cellular insults, neoplastic transformation and chemotherapeutic agents. Senescent cells are not dormant, as they secrete a variety of cytokines, chemokines, growth factors and even matrix metalloproteinases and hence affect the biology of neighbouring cells [23].

Mitotic catastrophe is a well-regulated mechanism whereby the cells undergoing incomplete mitosis are either eliminated through apoptotic/necrotic cell death or enter cellular senescence. Several factors can contribute to a defective mitosis including (but not limited to) DNA damage, defects in cell cycle checkpoint, aberrant chromosomal segregation, abnormal microtubule dynamics and problematic DNA replication. Morphologically, mitotic catastrophe is characterised by multinucleation, macronucleation and micronucleation. The exact molecular mechanisms regulating the intrinsic apoptosis during mitotic catastrophe are not clear; however, p53-dependent caspase-2 activity is reported to play an important part as in the absence of p53 mitotic defects trigger a caspase-2-independent necrotic RCD. Importantly, overcoming mitotic catastrophe is crucial for neoplastic transformation and progression (i.e., polyploid and aneuploid cells with high genetic instability would not survive otherwise). Interestingly though, the cancer cells which survive mitotic catastrophe are reported to secrete type I IFN due to the presence of cytosolic dsDNA. The main reason NCCD recommends against including mitotic catastrophe in RCD is because mitotic catastrophe does not always result in RCD and can lead to cellular senescence as well. Instead, NCCD [23] recommends using the term “mitotic death” when mitotic catastrophe leads to RCD.

Terminal differentiation is a pathophysiological process where cells, at the end of their developmental programme either enter a state of permanent growth arrest (e.g., erythrocytes, osteocytes and neurons) or undergo cell death and, in the case of keratinized epithelium, become part of the tissue [23].

Accidental cell death (ACD): Virtually instantaneous and uncontrollable form of cell death corresponding to the physical disassembly of the plasma membrane caused by extreme physical, chemical, or mechanical cues.

Anoikis: Specific variant of intrinsic apoptosis initiated by the loss of integrin-dependent anchorage.

Autophagy-dependent cell death: A form of RCD that mechanistically depends on the autophagic machinery (or components thereof).

Autosis: A specific instance of autophagy-dependent cell death that critically relies on the plasma membrane Na+/K+-ATPase.

Cell death: Irreversible degeneration of vital cellular functions (notably ATP production and preservation of redox homeostasis) culminating in the loss of cellular integrity (permanent plasma membrane permeabilization or cellular fragmentation).

Cellular senescence: Irreversible loss of proliferative potential associated with specific morphological and biochemical features, including the senescence-associated secretory phenotype (SASP). Cellular senescence does not constitute a form of RCD.

Efferocytosis: Mechanism whereby dead cells and fragments thereof are taken up by phagocytes and disposed.

Entotic cell death: A type of RCD that originates from actomyosin-dependent cell-in-cell internalization (entosis) and is executed by lysosomes.

Extrinsic apoptosis: Specific variant of RCD initiated by perturbations of the extracellular microenvironment detected by plasma membrane receptors, propagated by CASP8 and precipitated by executioner caspases, mainly CASP3.

Ferroptosis: A form of RCD initiated by oxidative perturbations of the intracellular microenvironment that is under constitutive control by GPX4 and can be inhibited by iron chelators and lipophilic antioxidants.

Immunogenic cell death: A form of RCD that is sufficient to activate an adaptive immune response in immunocompetent hosts.

Intrinsic apoptosis: Type of RCD initiated by perturbations of the extracellular or intracellular microenvironment, demarcated by MOMP, and precipitated by executioner caspases, mainly CASP3.

Lysosome-dependent cell death: A type of RCD demarcated by primary LMP and precipitated by cathepsins, with optional involvement of MOMP and caspases.

Mitochondrial permeability transition (MPT)-driven necrosis: Specific form of RCD triggered by perturbations of the intracellular microenvironment and relying on CYPD.

Mitotic catastrophe: Oncosuppressive mechanism for the control of mitosis-incompetent cells by RCD or cellular senescence. Per se, mitotic catastrophe does not constitute a form or RCD.

Mitotic death: Specific variant of RCD (most often, intrinsic apoptosis) driven by mitotic catastrophe.

Necroptosis: A modality of RCD triggered by perturbations of extracellular or intracellular homeostasis that critically depends on MLKL, RIPK3, and (at least in some settings) on the kinase activity of RIPK1.

NETotic cell death: A ROS-dependent modality of RCD restricted to cells of hematopoietic derivation and associated with NET extrusion.

Parthanatos: A modality of RCD initiated by PARP1 hyperactivation and precipitated by the consequent bioenergetic catastrophe coupled to AIF-dependent and MIF-dependent DNA degradation.

Programmed cell death (PCD): Particular form of RCD that occurs in strictly physiological scenarios, i.e., it does not relate to perturbations of homeostasis and hence does not occur in the context of failing adaptation to stress.

Pyroptosis: A type of RCD that critically depends on the formation of plasma membrane pores by members of the gasdermin protein family, often (but not always) as a consequence of inflammatory caspase activation.

Regulated cell death (RCD): Form of cell death that results from the activation of one or more signal transduction modules, and hence can be pharmacologically or genetically modulated (at least kinetically and to some extent).

Box 1.1: **The recommended operational definitions of cell death modalities by the Nomenclature Committee on Cell Death (NCCD) 2018.** Reprinted by permissions granted under Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International License [http://creativecommons.org/licenses/by-nc-sa/4.0/] from Springer Nature: Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, *et al*. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. CELL DEATH & DIFFERENTIATION. 2018;25(3):486-541. doi:10.1038/s41418-017-0012-4. Copyright © 2018, Springer Nature.

1.5 Previously Known Anticancer Genes

A brief overview of some of the previously known anticancer genes is provided in table 1.1.

1.5.1 Apoptin

Apoptin was the first reported anticancer gene [6]. It was isolated from the genome of Chicken Anaemia Virus (CAV) in the early 1990s. CAV causes cell death in myeloid progenitor cells in bone marrow and the T-lymphocyte precursors in thymus. Apoptin encodes a small 14 kDa protein which does not share sequence homology with any known cellular proteins. It has been reported to induce cell death in several malignant and transformed cells while leaving the normal cells relatively unharmed. Over 70 cell lines from a variety of tumour types including colon carcinoma, lung cancer, melanoma and lymphoma, as well as the SV40- and UV-radiation transformed cells have shown sensitivity to apoptin-induced cell death [13]. Apoptin causes caspase-dependent and p53-independent apoptosis with loss of mitochondrial membrane potential and cytochrome *c* release. Moreover, BCL2 overexpression is incapable of counteracting apoptin-induced apoptosis [25]. Mitotic catastrophe has also been reported as a mode of apoptin-induced cell death; however, the evidence remains to be conclusive. Lanz *et al.* [26] reported apoptin-induced cell death during mitosis and the following interphase, but the cells showed the typical morphological characteristics of apoptosis and these were not confined to dividing cells. As the nuclear envelope presents the main barrier in nuclear transport of transfected plasmid DNA [27], cells tend to express higher levels of transfected genes following cell division when transfected DNA can freely enter and incorporate into the nuclear environment.

The exact mode of action for the cancer specific cytotoxicity remains elusive, although there are a few interesting indications:

I. Different intracellular localization of apoptin between cancer and normal cells: apoptin is capable of forming globular multimeric complexes comprising 30-40 non-covalently attached apoptin monomers and these multimeric complexes are retained in the cytoplasm. In tumour cells and transformed cells, apoptin is mainly localized in nucleus, while in normal cells it is primarily found in the cytoplasm where it makes the multimeric complexes and becomes epitope shielded [13]. There are a few caveats to this proposal: firstly, apoptin does not act as

a transcription factor. Secondly, the experimental localization (and overexpression) of apoptin in the nuclei of normal cells does not induce apoptosis [13], and thirdly, at least one study has suggested the nuclear localization to be a result of protein expression levels rather than cell type [28].

II. Differential phosphorylation of apoptin: apoptin contains several putative phosphorylation sites. One of these, Thr-108, is reported to be phosphorylated in cancer and transformed cells but not in normal cells. Two survival pathways which are upregulated in cancer cells, namely PI3 kinase/AKT and protein kinase C (PKC), have been proposed to harbour the kinases which phosphorylate apoptin [13]. How the phosphorylated apoptin causes cell death compared to its unphosphorylated counterpart is not clear; however, in the proposed model, the phosphorylated apoptin translocates to the nucleus while the unphosphorylated protein stays in the cytoplasm where it eventually undergoes proteasomal lysis. Several proteins have been shown to colocalize with apoptin and modulate its function in various cells. These include FADD, B cell CLL/lymphoma 10 (BCL10), protein kinase G (PKG), heat shock protein 70 (HSC70), promyelocytic leukaemia protein (PML), anaphase promoting complex (APC1), death effector domain associated factor (DEDAF), peptidyl-prolyl isomerase like 3 (PIPIL3) and breast cancer associated gene (BCA3) [13]. There is no definitive answer as to how these interactions result in apoptin-induced tumour specific cell death. One particular player which keeps on 'popping up' when these interactions are studied in detail is NF-κB.

As described earlier, apoptin causes apoptosis but the relevant pathway(s) thus triggered remain(s) a matter of debate. In Soas-2 cells, for example, apoptin colocalizes with FADD but Jurkat cells, lacking FADD and caspase-8, remain sensitive to apoptin-induced apoptosis [13]. In MCF-7 breast cancer cells, apoptin triggers NUR77 shuttling to mitochondria. NUR77 is an orphan receptor which upon various apoptotic stimuli translocate to mitochondria where it activates an intrinsic apoptotic pathway by changing the conformation of anti-apoptotic BCL2 family proteins [13]. In the prostate cancer cell line overexpressing the anti-apoptotic BCL2, apoptin is reported to increase the levels of sphingolipid ceramide, the activity of which is downregulated in malignant cells and a defective ceramide pathway is also reported during chemotherapy resistance [13]. Lastly, the role of p53-associated protein p73 has also been proposed in the apoptotic functioning of apoptin but the signalling pathway linking the two proteins is not known [13].

1.5.2 E4orf4

The adenovirus early region 4 transcription unit (E4) open reading frame transcript 4 (E4orf4) is a 14 kDa polypeptide consisting of 114 amino acids. Physiologically, E4orf4 acts as a negative regulator of early virus gene expression, thus giving way to the capsid formation for the optimum packaging and release of the virus (before the onset of excessive early gene transcription). E4orf4 is also involved in (i) alternative splicing of viral *L1* RNA, (ii) activation of mTOR pathway for bypassing the nutrient checkpoint for the protein translation and (iii) downregulation of cellular antiviral immunity. It was the second reported gene of its kind – when overexpressed on its own (out of the context of viral infection), E4orf4 causes cell death in 40 human cancer cell lines while leaving many primary human cell cultures unharmed [29]. E4orf4 causes necrotic cell death with a peculiar morphology involving membrane blebbing and nucleic acid condensation. Caspases are reported to be active but dispensable for E4orf4-induced cell death. Moreover, the cell death caused by E4orf4 is p53-independent. Mitotic catastrophe has also been reported as a mode of cell death induced by E4orf4 after prolonged G2/M arrest [6]. Just like apoptin, the exact mechanism of E4orf4-induced differential cell death is far from clear, but the clues and explanations are indeed very interesting.

During the viral infection, E4orf4 is primarily localized in the nucleus, whereas during E4orf4-induced cell death, it localizes in the characteristic membrane blebs. Although the caspases are dispensable for E4orf4-induced cell death, the inhibition of Ca^{++} dependent intracellular calpain proteases greatly inhibits E4orf4-induced cell death [29]. One of the two main E4orf4-interacting proteins is tyrosine kinase proto-oncogenes of SRC family. When the SRC family kinases are constitutionally upregulated, they bind with E4orf4, pulling it out of the nucleus and into the membrane blebs where E4orf4 apparently interacts with calpains to disrupt the cytoskeleton and induce the peculiar necrotic cell death [29]. On the other hand, nuclear condensation is also observed during E4orf4-induced cell death. Here, the second major interacting partner of E4orf4 comes into the picture – protein phosphatase 2A (PP2A) complex comprises a large family of Ser/Thr phosphatases which are involved in the regulation of several cellular processes. E4orf4 binds with PPA2 (this binding site overlaps the one for SRC kinases) [30]. It is suggested that E4orf4 causes nuclear condensation upon binding with PP2A complex due to PPA2 retaining the phosphatase activity in complex with E4orf4. It is also proposed that targeting PPA2 to novel substrates may play a role in tumour-specific cell death induced by E4orf4 [29]. For example, the PPA2 complexed with E4orf4 binds the ATP-utilizing chromatin

assembly and remodelling factor (ACF). Furthermore, E4orf4 treatment in breast cancer, gastric cancer, and melanoma xenograft models, by employing various delivery methods, has been shown to successfully reduce tumour size with minimal generalized toxicity [29]. E4orf4 also causes a non-apoptotic cell death and organismal lethality in PP2A-B55 and SRC64B overexpressing *D. melanogaster*. Interestingly, E4orf4 actively inhibits apoptosis in normal *D. melanogaster*. It is important to mention here that the cell death induced by E4orf4 is mainly apparent at high expression levels of E4orf4 [29].

1.5.3 Rodent parvovirus NS1

Following the observations that certain parvoviruses hamper the growth of the xenografts implanted in laboratory animals, the rodent parvovirus attracted attention as a potential cancer-specific therapeutic agent. It was also noted that the rodent parvoviruses have the capacity to not only prevent the formation of spontaneous tumours, but they also actively inhibit the formation of chemically-induced tumours in laboratory animals [31]. The small single stranded DNA genome of parvovirus encodes 8 proteins, and the 83 kDa non-structural protein 1 (NS1) is essential for viral propagation. NS1 has been reported to cause cell death, preferentially in neoplastic cells through the canonical apoptotic pathway, cytosolic necrosis through cytoskeleton collapse and lysosomal cell death. Moreover, ectopic expression of NS1 causes mitochondrial depolarization and an increase in reactive oxygen species (ROS), subsequently leading to DNA damage response and cell cycle arrest [32]. The mode of selective killing of neoplastic cells by ectopic expression of NS1 probably provides the best example of synthetic lethality. Upon the parvovirus infection, the phosphoinositide dependent kinase 1 (PDK1)/phosphokinase C (PKC)/phosphokinase B (PKB) signalling cascade is activated, which helps the progenitor virions to avoid apoptosis as well as the stress-induced translational shut-off. The continued activation of PDK1/PKC/PKB in the later stages of infection leads to phosphorylation of NS1, which in turn activates multiple cell death mechanisms for the efficient release of virions. When ectopically overexpressed in cancer cells where the PDK1/PKC/PKB signalling cascade is constitutively active, the rapid phosphorylation of NS1 leads to the induction of cancer specific cell death. NS1 overexpression in lung embryonic fibroblast cell line MRC-5 could not induce cell death, but when these cells were engineered to have constitutively active PDK1, NS1 overexpression led to necrosis [32]. This is only one of the proposed mechanisms of NS1-induced cell

death; the protein is very multifunctional. Various domains of NS1 are capable of physical interaction with several cellular proteins and impart multiple effects [32].

The self-propagating oncolytic parvovirus H1-PV (brand name ParvOryx) underwent phase I/II clinical trials on recurrent glioblastoma multiforma patients in 2011 with promising results [33]. The results were interesting in the sense that more emphasis was placed on the immunogenicity of the virus rather than on oncolytic capacity. Although this may seem unexpected in the first instance, in fact, the tumour-specific immunogenicity is one of the prime advantages of oncolytic virus therapy. The selective neoplastic tropism of oncolytic virus attracts the immune response into target regions through the 'bystander' effect where the immune response is aimed at the released virus particles but is advantageous for combatting malignant cells in the vicinity [32].

1.5.4 MDA-7/IL-24

Melanoma differentiation associated gene-7, later named as interleukin-24, is a 23.8 kDa secreted cytokine belonging to IL-10 family. In the physiological context, the protein has an immunomodulatory function and is expressed in the tissues of the immune system. MDA-7/IL-24 binds its cognate receptors in order to activate STAT1 and STAT3 signalling [34]. But when ectopically overexpressed, MDA-7/IL-24 causes selective cell death in several cancers. MDA-7/IL-24 mainly causes canonical apoptosis but autophagy has also been reported as a mode of cell death. MDA-7/IL-24 has anti-angiogenic and anti-metastatic properties as well [35]. Moreover, in hepatocellular carcinoma, MDA-7/IL-24 causes G2/M cell cycle arrest [36].

The mechanisms underlying the cancer-specific cell death by MDA-7/IL24 are far from clear (as is the case with the other anticancer genes). In addition, even the proposed possibilities in the case of MDA-7/IL24 (induced cancer specific cell death) are rather vague. MDA-7/IL-24 causes apoptosis through activation of ER stress. MDA-7/IL-24 prevents the interaction between the ER chaperone protein BiP and PKR-like endoplasmic reticulum kinase (PERK), resulting in the autophosphorylation of the latter and the subsequent activation of apoptosis. Putatively, MDA-7/IL-24 causes selective apoptosis in cancer cells because the levels of ER stress in cancer cells are already higher than in the normal cells [34]. Thus, the overloading of already heightened ER stress could lead to the selective activation of apoptosis in cancer cells. Similarly, MDA-7/IL-24 overexpression leads to an increase in the ROS levels

[37]. As the cancer cells already have higher ROS, further increase in ROS would overcome the cells' natural antioxidant setup.

A MDA-7/IL-24 incorporated adenoviral vector (Ad.*mda-7*; INGN 241) underwent phase I clinical trial in 28 patients diagnosed with various malignancies and according to Menezes *et al.*, [34] '*Overall, Ad.mda-7's (INGN 241) entry into the clinic in a Phase I trial employing intratumoral injection was exceptionally successful demonstrating that this therapy was well-tolerated in human subjects with an ~44 % clinical response in injected lesions*' (p.146).

1.5.5 PAR4

PAR4 (prostate apoptosis response 4) is a 340 amino acid-long multi-domain protein which is expressed ubiquitously. PAR4 overexpression induces apoptosis in a variety of cancer cells in a p53-independent manner and can override the anti-apoptotic effects of BCL2 and BCL-X_L. Physiologically, PAR4 is an apoptosis-related gene and is claimed to be a *bona fide* tumour suppressor as it is deleted, mutated, or epigenetically silenced in a number of cancers [38]. Moreover, PAR4 is downregulated by oncogenes like RAS, cSRC and AKT. At the same time, PAR4 is reported to cause cell death selectively in cancerous and transformed cells while only sensitizing the immortalized and normal cells to apoptosis. The proposed model of cancer specific apoptosis by this anticancer gene is very interesting. In cancerous and transformed cells, PAR4 causes the trafficking of FASLG to the cell membrane leading to the induction of the extrinsic apoptotic pathway and simultaneously suppresses the NF-κB mediated anti-apoptotic pathway [39]. The latter depends on the nuclear translocation of PAR4. In normal and immortalized cells, PAR4 fails to translocate to the nucleus and suppress the NF-κB driven anti-apoptotic signals. The nuclear trafficking of PAR4 is regulated by its phosphorylation through the protein kinase A I (PKA-I). PKA-I is a cyclic AMP-dependent kinase which is involved in the regulation of cell growth, metabolism and differentiation and it is overexpressed in various cancer cell lines and primary tumours. The phosphorylation of PAR4 by PKA-I is proposed to be one of the factors behind selective nuclear translocation of PAR4 in transformed and cancer cell lines. Other factors involved in the PAR4-induced selective apoptosis include binding and deactivation of BCL2, PKCs and topoisomerase I (TOPO-I) [38].

The amino acids 137-195 of PAR4 constitute a selective for apoptosis of cancer (SAC) domain. Overexpression of SAC domain on its own causes selective apoptosis in malignant and transformed

cells while leaving the normal and immortalized cells unharmed [40]. The SAC domain contains the nuclear localization signal and PKA-I phosphorylation site. The selective phosphorylation by PKA-I is proposed to be the main regulating factor for its tumour-specific apoptosis-inducing properties. Moreover, the transgenic mice overexpressing SAC domain showed a significantly higher resistance to the development of induced and spontaneous tumours. Also, the murine embryonic fibroblasts (MEFs) derived from SAC mice showed a ten-fold higher rate of apoptosis and were also resistant to oncogenic transformation using cMYC and RAS [41].

Lastly, PAR4 is a secretory protein and is present / detectable in the systemic circulation. The secreted protein retains its SAC domain and its ability to cause apoptosis selectively in cancer cells when applied from the outside. In fact, the treatment of cells with SAC domain (on its own) is sufficient in inducing selective apoptosis. The secreted PAR4 binds to the cell stress protein GRP78 on the cell surface, which leads to the upregulation of ER stress protein PERK, and the subsequent activation of caspase-8 and caspase-3 in a FADD-dependent manner. The proposed mechanism behind the cancer cell specificity of PAR4 in this case is considered to be the higher number of GRP78 receptors in cancer cells in comparison to the normal cells [38].

1.5.6 Orctl3

The orphan transporter protein, Orctl3 is a transmembrane anion transporter in the physiological context. The overexpression of the mouse Orctl3 induces caspase-dependent and p53-independent apoptosis in several transformed and cancer cell lines but not in the non-transformed immortalized cells [12]. The human ORCTL3 overexpression, however, is devoid of this effect. Unlike other anticancer genes which were identified by chance, Orctl3 was discovered through implication of a systematic functional genetic screen (methodologies discussed later) employing HEK293T cells followed by normal rat kidney NRK cells and NRK cells transformed through HRas [42].

Orctl3 causes apoptosis through inhibition of stearoyl-CoA desaturase-1 (SCD1). SCD1 is an ER-resident enzyme which converts many saturated fatty acyl-CoAs (SFAs) to monounsaturated fatty acids (MUFAs). SCD1 plays a key role in membrane synthesis and fluidity. One of the key metabolic changes during cancer cell transformation is the increased synthesis of MUFAs which are required for the biosynthesis of membrane phospholipids in the rapidly dividing cells. It is no wonder, then, that SCD1 is overexpressed in transformed and cancer cells. The downregulation of SCD1 itself causes

selective apoptosis in tumour cells but not in the normal cells (notably, the Orctl3 overexpression was reported to induce a much stronger apoptosis than the one caused by inhibition of SCD1 alone [12, 17]). Moreover, the SCD1 substrates are known to be potent pro-inflammatory molecules and SCD1 is deregulated in atherosclerosis [12, 17].

Orctl3 overexpression also causes an incomplete ER stress in the sense that the ER stress marker which is used to gauge the misfolded proteins (ATF4) is upregulated, but it does not lead to subsequent upregulation of ER stress-induced apoptotic marker (CHOP). It is thought that ER stress caused by Orctl3 in transformed cells might be activating novel pathways, different from the canonical misfolded proteins response. One of these novel mechanisms was previously reported by Stefan Grimm's group where mitochondrial fission 1 protein (FIS1) and B-cell receptor associated protein 31 (BAP31) are part of a unique complex named ARCosome. In this scenario, FIS1 is recruited from the outer mitochondrial membrane to ER, where it (FIS1) cleaves BAP31 into smaller fragments, named p20Bap31. The p20Bap31 harbour a variant of death effector domain (vDED) through which they recruit and activate procaspase-8. Additionally, the ARCosome formation results in the release of Ca^{++} from the ER which subsequently actuates mitochondria for apoptosis (see Ref [43] for details). Upon Orctl3 overexpression, more p20BAP31 fragments are found in transformed cells than in the normal cells [17].

1.5.7 TRAIL

Tumour necrosis factor related apoptosis inducing Ligand (TRAIL) is probably the most well-known among the cancer-specific cell death inducers and thus makes the prime example of an anticancer gene. TRAIL (also known as TNFSF10 and APO2L) belongs to the TNFSF family and is closely related to FASLG, also known as CD95L. Like most of the other members of the TNFSF family, TRAIL is expressed as a type II transmembrane protein and is later released through the action of membrane bound metallo and serine proteases, a phenomenon known as shedding [44]. TRAIL is expressed in several tissues throughout the body including kidney, lung, liver, spleen and placenta [45]. Moreover, TRAIL is reported to be present in the plasma of a healthy adult at a concentration of approximately 100 pg/ml [44]. The physiological role of TRAIL appears to be the activation and modulation of the immune response. TRAIL expression on the surface of several immune cells is reported upon stimulation with interferon γ (IFNγ). Moreover, TRAIL also serves to maintain the immune homeostasis through prevention of aberrant T-cell activation [44].

TRAIL attracted the attention of cancer researchers when the injection of human recombinant TRAIL resulted in shrinkage of xenografts from a variety of cancers with minimal overall cytotoxicity, especially after the previous attempts with FASLG had failed due to liver toxicity within hours of treatment [45]. Several normal, immortalized, non-transformed cells remain insensitive to TRAIL but upon their transformation with RAS or cMYC, they become sensitive to TRAIL-induced cell death [45]. Moreover, TRAIL acts as an anti-tumour effector molecule at its endogenous expression level. The *Trail* knockout mice grow normally but develop spontaneous lymphomas. Furthermore, loss of *Trail* in *Trp53*$^{+/-}$ mice predisposes them to spontaneous development of diverse types of tumours. Also, when treated with Trail-neutralizing antibodies, mice become susceptible to chemically induced carcinogenesis. This anti-tumour role is at least partly reported to be a result of TRAIL expression on the natural killer (NK) cells [44]. As the name suggests, the NK cells are a part of body's natural cellular defence against viral infections and tumorigenesis and they actively kill the virally infected, malignant, and pre-malignant cells in part through the INFγ induced expression of TRAIL on their surface. Interestingly, the FASLG can only induce cell death in its membrane-bound state while the soluble form has the opposite effects. Similarly, the recombinant forms of TRAIL which are fused to the motifs that mimic the membrane-bound conformation are more potent inducers of cell death [44]. Upon binding with TRAIL receptor TRAIL-R1 (TNFRSF10A) and TRAIL-R2 (TNFRSF10B), TRAIL induces classical apoptosis through the canonical extrinsic pathway which includes the formation of a DISC complex, harbouring FADD, activation of effector caspases 8 and 10, leading to the subsequent activation of effector caspases 3 and 7. Where the DISC complex fails to activate the apoptosis, activation of tBID leads to release of cytochrome *c* from the outer mitochondrial membrane and ultimately the activation of intrinsic apoptotic cascade.

Similar to the rest of anticancer genes, the mechanisms behind the tumour specific apoptosis by TRAIL are unclear and there are several indirect indications. Firstly, differential binding of TRAIL to different receptors is proposed to be one of the mechanisms in cancer specificity. Unlike other members of the TNFSF family, TRAIL binds to five different types of receptors. TRAIL-R1 and TRAIL-R2 contain the death domain (DD), whereas the TRAIL receptors TRAIL-R3 (TNFRSF10C) and TRAIL-R4 (TNFRSF10D) lack the DD. Moreover, osteoprotegerin (OPG, also known as TNFRSF11B) acts as the soluble receptor. The non-death domain containing receptors are proposed to act as decoy receptors, scavenging the ligand away from the effector receptors. This proposition, although attractively simplistic, remains unproven *in vivo* at the physiological expression levels of these receptors.

In addition, contrary to popular belief, TRAIL has a dual role in tumorigenesis, i.e., in TRAIL sensitive cells, it induces apoptosis whereas in TRAIL resistant cells it aids cell proliferation, immunosuppression and even metastasis [44]. A key molecule in this context again happens to be NF-κB. TRAIL-R4 actively induces NF-κB upon binding of its ligand. Interestingly, TRAIL-R1 and TRAIL-R2 DISC complexes harbour receptor-interacting serine/threonine protein kinase 1 (RIPK1), a necroptosis executioner kinase. However, in the context of TRAIL signalling, RIPK1 induces NF-κB activation and the subsequent proliferation or invasion in the TRAIL-resistant cells. Just like the rest of anticancer genes, TRAIL induces a plethora of signalling pathways including the JUN N-terminal kinase (JNK) and p38 MAPK-ERK pathways. The latter is reported to induce proliferation in caspase-8 deficient small-cell lung cancer (SCLC) cells upon treatment with TRAIL. The caspase-8 deficiency/inactivity plays a significant role in TRAIL resistance as caspase-10 cannot compensate for the loss of caspase-8 in the extrinsic apoptosis induced by TRAIL [44]. Also, TRAIL induces vascularization in soft tissue sarcoma just like the vascular endothelial growth factor (VEGF). Moreover, at low expression levels, TRAIL triggers migration in non-small cell lung cancer through RIPK1, SRC and STAT3. One could imply that expression levels of TRAIL are also important in the differential response. It has also been proposed that the intracellular localization of TRAIL receptors could be the deciding factor in TRAIL sensitivity *vs* resistance. In the TRAIL resistant pancreatic cancer cells, the preferential localization of TRAIL-R2 in the nucleus leads to impaired maturation of *let-7*, a tumour-suppressor microRNA which is known to downregulate KRAS expression [46, 47]. Pertaining to the diversity of the intracellular signalling cascades initiated by TRAIL, Kumar-Sinha *et al.* [48] reported that TRAIL induces a classical type I interferon response in MCF7 breast cancer cells *in vitro*. Through a series of elaborate experiments, the study reports that the pre-treatment of cells with interferon-β (INFβ) sensitizes the breast cancer cells to TRAIL-induced apoptosis.

No other anticancer gene has enjoyed more popularity and undergone more clinical studies than TRAIL. Recombinant TRAIL (dulanermin) and TRAIL-R1 and TRAIL-R2 agonist antibodies (mapatumumab, conatumumab, tigatuzumab, TAS266 and DS-8273a) have been subjected to clinical trials; however, the outcomes have been rather disappointing [44]. The failure of recombinant TRAIL therapy could be attributed to (i) very short systemic half-life of 30 minutes, (ii) binding to all the TRAIL receptors instead of targeting the death domain receptors only, (iii) poor affinity to its receptors and (iv) non-availability of biomarkers to discern TRAIL-resistant *vs* TRAIL-sensitive tumours. The latter is particularly important and also contributes to poor outcomes concerning TRAIL receptor

agonists. Administering TRAIL or its receptor-agonists in patients with TRAIL-resistant tumours would in fact prove counterproductive. In the case of TRAIL receptor agonist antibodies, their (inherent bivalency and consequent) inability to trigger higher order receptor multimerization is also proposed to be the reason behind their diminished efficacy. In my view, the poor outcomes of TRAIL clinical trials have discouraged research into tumour specific cell-death inducers. But fortunately, work is continuing to alleviate the obstacles in TRAIL/receptor-agonist therapy [44].

1.5.8 HAMLET

The human α-lactalbumin made lethal to tumour cells (HAMLET) is a unique anticancer agent. HAMLET is a partially misfolded form of α-lactalbumin which is conjugated with the oleic acid group [49] – the α-lactalbumin is first partially misfolded using low pH or EDTA and kept in this misfolded state by attaching the oleic acid group. The application of HAMLET causes cell death differentially in the cancer cells while sparing their normal counterparts [50]. The cell death caused by HAMLET involves activation of caspases but these cysteine proteases are dispensable, rather autophagy is reported to be necessary for HAMLET-induced cell death [6]. Moreover, the HAMLET-induced cell death is independent of the p53 status of the cell and can overcome BCL2 overexpression. After encounter, HAMLET invades the sensitive cells and triggers a range of intracellular effects including mitochondrial membrane permeabilization and cytochrome *c* release, it causes an unfolded protein overload, activation of proteasome 20S and also appears to involve ER stress. HAMLET translocates to the nucleus and binds histones H3, H2B and H4 leading to impaired transcription, chromatin malfunction and nuclear condensation [51]. The mechanisms behind the tumour specificity of HAMLET are far from clear and the literature is devoid of any suggested possibilities either. However, Gustafsson *et al.* [52] reported a very successful clinical trial where ectopic application of HAMLET reduced the lesion volume of skin papilloma by more than 75% in all 20 test patients compared to only 3/20 patients in the placebo group. All lesions were reported to be resolved in 83% of patients at the end of a two-year follow up. Puthia *et al.* [53] also reported that the oral administration of HAMLET not only reduced the colorectal tumour progression and mortality in $Apc^{+/-}$ mice but also the supply of HAMLET through drinking water from the weaning time significantly prevented the development of tumours in these animals.

Notably, despite the fact that Grimm and Notborne [6] mentioned HAMLET in the list of anticancer genes, it is not a gene or its protein product. The question remains if the differential apoptosis by

ectopic overexpression is the right criterion for being designated as an anticancer gene, because even in the case of TRAIL or SAC domain of PAR4, it is not the overexpression but the treatment with the polypeptide products that causes tumour-specific cell death.

Table 1.1: **An overview of the previously reported anticancer genes.**

Anticancer gene	Origin	Type(s) of cell death	Caspases involved?	Functional p53 required?	Activated by phosphorylation?	Engaging cell death pathway	Intracellular localization in cancer cells	Related molecules and pathways	Proposed differential action in cancer *vs* normal cells
Apoptin	Viral	Apoptosis, autophagy, mitotic catastrophe.	Yes	No	Yes	Intrinsic	Nucleus	PI3K/AKT, FADD, BCL10, PKG, HSC70, PML, APC1, DEDAF, PIPL3, BCA3, NUR77 and NF-kB.	Differential phosphorylation and intracellular localization
E4orf4	Viral	Necrosis.	No	No	Yes	Intrinsic	Cytoplasm, nucleus.	SRC kinases and PP2A complex.	Differential intracellular interactions between normal and cancer cells.
Parvovirus NS1	Viral	Apoptosis, necrosis, cytoskeletal collapse, lysosome-dependant cell death.	No	No	Yes	Intrinsic	Cytoplasm.	PDK1/PKC/PKB.	Differential phosphorylation between normal and cancer cells, increase in ROS and mitochondrial depolarization in cancer cells.
MDA-7/IL24	Cellular	Apoptosis and autophagy.	Yes	No	No	Intrinsic	ER and receptor binding.	BIP and PERK.	ER stress and ROS induction, as cancer cells have higher of both already, they succumb to further induction of these.
PAR-4	Cellular	Apoptosis.	Yes	No	Yes	Extrinsic, intrinsic.	Nucleus, ER and plasma membrane.	APO-1, NF-kB, PKAI, PKC, BCL2, TOPO-I and PERK.	Trafficking of FASL to plasma membrane and inhibition of NF-kb, Induction of ER -Stress.
Orctl3	Cellular	Apoptosis.	Yes	No	No	Intrinsic	ER, plasma membrane and Golgi.	SCD-1, FIS1 and BAP31.	Inhibition of fatty acid metabolism which affects cancer cells more than normal cells, Induction of ER-stress.
TRAIL	Cellular	Apoptosis.	Yes	No	No	Intrinsic	Plasma membrane.	TRAIL receptors R1, R2, R3 and R4, RIPK1 and NF-kB.	Death receptor activation.
HAMLET	Chemical	Autophagy.	Yes	No	No	Intrinsic	Nucleus, ER, mitochondria.	Proteasome 20S, histones H3, H2B and H4.	Causes ER-stress, cytochrome c release and transcription block in cancer cells.

Adapted and modified from, Trends Mol Med, 16(2), Grimm S., Noteborn M., Anticancer genes: inducers of tumour-specific cell death signalling, page No.88-96. Copyright © 2010, with permission from Elsevier. Permissions granted through Copyright Clearance Center's RightsLink®, Order No, 4702160462681. Modifications applied: Three columns added: 'Origin', 'Related molecules and pathways' and 'Proposed differential action in cancer vs normal cells' (Please see section 1.5 and the sub-headings for the original sources of factual information added). Information in columns 'Type of cell death' (heading changed to 'Type(s) of cell death') has been changed for Apoptin, MDA-7/IL24 and Parvovirus NS1 (see section 1.5 and the subheadings for the original sources of the factual information added). Heading 'Subcellular localization in cancer cells' changed to 'Intracellular localization in cancer cells'. Rows, 'Noxa' and 'Brevinin-2R' removed. Column 'Blocked by Bcl-2' and 'References' removed.

1.6 Expectations and Challenges

The research in the field of anticancer genes holds exciting potential. The obvious answer as to how anticancer genes could be beneficial in a clinical context would be their possible use in gene therapy, where the ectopic overexpression of anticancer genes could eradicate the malignancy with minimal generalized toxicity. Ectopic application of tumour-specific agents like HAMLET is another promising possibility. In addition, studying the function(s) of anticancer genes could reveal the vulnerabilities that are specific to cancer cells. Such studies would further our knowledge regarding the mechanisms employed by cancer cells to ensure their survival and progression, hence aiding in improving current therapeutic strategies against cancer. Ideally, developing a cocktail of agonists and antagonists for cancer-specific targets could possibly lead to more tolerable, economical, and effective treatments for even the most aggressive and well-established malignancies. These expectations, however, come with enormous challenges. The first challenge is the delivery of the gene to the cancer cells. Several strategies are under development; most commonly they involve the use of viruses for gene therapy. Gene therapy with the anticancer genes is fundamentally different from the broader use of the term in the sense that in the case of anticancer genes, the therapy is intended to ectopically overexpress the gene rather than correcting a pre-existing genetic defect. Here, delivering the anticancer gene to virtually all the cancer cells is a daunting task because even if a single cancer cell survives, it would act as the progenitor of further cancerous growth. The quiescent tumour stem cells, for example, are resistant to oncolytic parvovirus. The parvovirus has single stranded DNA genome which needs to be converted to the double stranded transcription template in S-phase of cell cycle, which is not possible in quiescent cells capable to survive in G0 [32]. It is worth mentioning here that adeno-associated viruses and herpes virus-based vectors are capable of producing viral protein in quiescent tumour cells. Moreover, viruses are essentially immunogenic and without using the concomitant immuno-suppressive therapies, the use of delivery viruses is limited to essentially a one-time injection [54]. In addition, as discussed previously, the expression level of the gene could be the crucial factor in the effectivity of an anticancer gene and achieving the desired gene expression levels is a major challenge with use of viral vectors. Anticancer genes, however, offer an advantage in terms of onco-specificity of viral vectors, which means that even if there is overexpression of an anticancer gene in normal cells of various organs, only the cancer cells would be affected. Also, the relatively recent advances in gene delivery methods such as the use of fusion proteins with cell-penetrating peptides (CPP) or their

domains (e.g., HIV Tat system, reviewed in Ref [55]) and ultrasound sonoporation in combination with microbubbles [56] offer hope for effective anticancer gene/protein delivery.

When it comes to pinpointing the changes that anticancer genes impart in tumour cells and using this information to develop targets for novel anticancer therapies, the challenges are even more daunting. Firstly, cancer is a multifactorial disease involving a plethora of aberrations in several cellular processes. The inherent genetic instability and rapid cell division cycles complicate matters further with a continuous supply of mutations that aid the cell in escaping the intra- and extracellular selection pressures applied. The aberrations are not at a stable steady state, but they continuously change according to the needs. On top of it, all of the anticancer genes engage a multitude of signalling molecules; when taken together in the context of multiple cell types, heterogeneous and "heteroploidious" cell population in tumours, tissue cultures and xenografts, it is obvious that pinpointing the exact mechanism of selective cell death is a very challenging task. There are two potential solutions to this problem. Firstly, with the advent of cutting-edge single cell transcriptomics, it is now possible to decipher the changes in individual cells. The problem in this context is that anticancer genes impart their actions mostly through protein-protein interactions, which are not easily spotted through changes in the transcriptome. Fortunately, reliable bioinformatic tools are now available which can flag the areas worthy of further functional studies. Secondly, a relatively small number of cancer progenitor cells present an attractive target for such experimental undertaking. No such study has been conducted involving anticancer genes and cancer stem cells (CSCs). The very advantage of anticancer genes lies in their ability to impart multiple effects in a heterogeneous population, that is, they cause several disruptions in an already disrupted, but vulnerable, signalling circuitry. How these effects would translate in a relatively smaller population of cancer progenitor cells is yet to be seen. Therefore, finding several targets of anticancer genes and using a cocktail of agonists and antagonists for the cancer specific therapy would probably offer more promising and workable prospects.

1.7 Introduction to the Robotic Single cDNA Investigation (RISCI)

The advancements in sequencing technologies coupled with the enormous collaborative scientific efforts (e.g., genome projects) gave rise to a flood of sequence information. Assigning function to the sequenced genes, on the other hand, is technically a different matter altogether. Genetic screening, using mammalian culture cells and other model organisms, is an effective tool for discovering the

function of newly sequenced genes. Functional genetic screens rely on mutagenesis, detection of resulting biochemical changes and isolation and characterization of causative gene(s) [57]. One of the strategies to achieve mutagenesis is through the ectopic overexpression of genes. cDNA libraries generated from individual cells, organs and the whole organism transcriptome(s), when linked to a strong promoter, provide appropriate expression levels for such ectopic gene expression. Traditionally, the genetic screening begins with the introduction of batches of expression plasmids into the cells, the batch that exhibits the biochemical change in question is isolated from mammalian cells (Hirt extraction) and the DNA is purified. This purified DNA is divided into further batches and introduced into the cells. The process is repeated until the plasmid causing the biochemical change in question is identified and isolated. This approach is termed screening through batch transfection (figure 1.2 – dotted red line). However, screening through batch transfection is not suitable for genes which cause cell death upon overexpression. This drawback can be countered by employing sib selection, where instead of purifying the extrachromosomal DNA from cells, the batch of plasmids that imparted the desired phenotype is divided further into smaller pools and introduced into cells. Similar to batch transfection, the process is repeated until a single plasmid population causing the desired phenotypic change is isolated (figure 1.2 – Solid blue line). However, such traditional screening procedures are laborious, expensive, time-consuming and thus, offer limited throughput potential.

A novel solution was devised by Stefan Grimm and associates, where, instead of introducing a batch of plasmids into the mammalian cells, expression plasmids containing only one cDNA are introduced into the culture cells. First, the plasmids are multiplied in bacteria, DNA is isolated and purified and the culture cells are transfected in the 96-well plate format. In this way, the cells are transfected by the individual plasmids, separated in the 96-well plate (figure 1.3). The high-throughput is achieved through robotic automation. This approach is termed robotic single cDNA investigation (RISCI) [57].

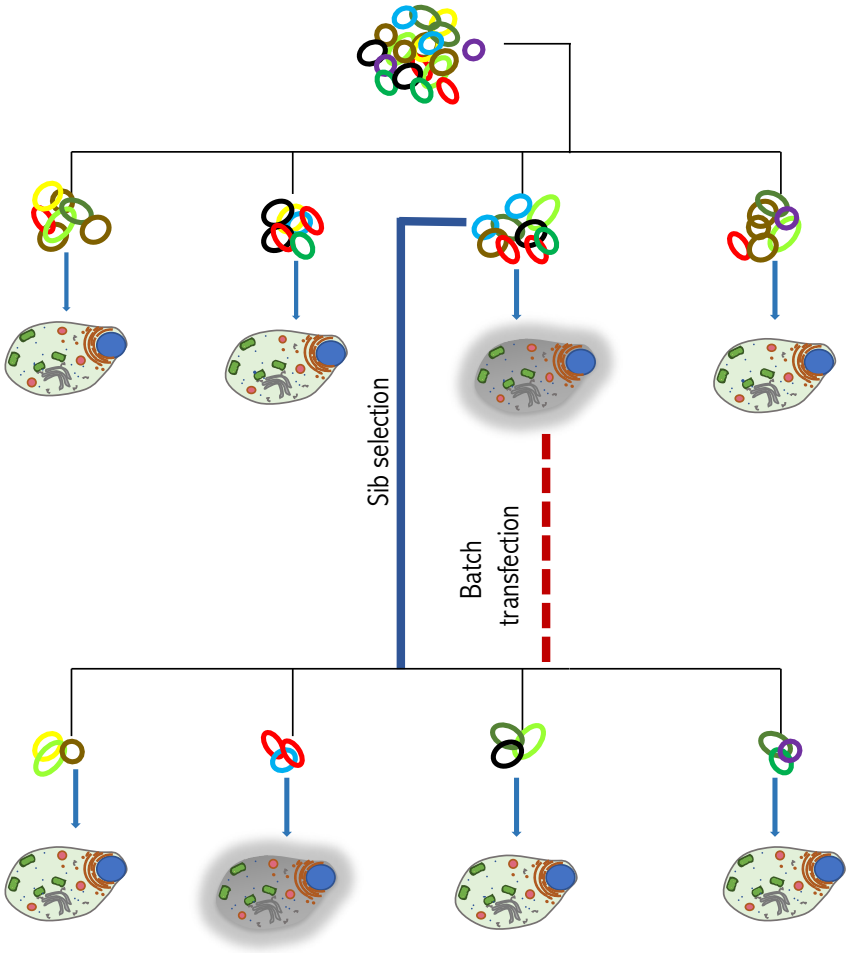


***Figure 1.2*: Traditional screening approach: Screening through batch transfection and sib selection**. Screening begins with the introduction of pools of expression plasmids into the cell. In batch transfection, the pool that exhibits the questioned biochemical change is isolated through Hirt extraction and the DNA is purified (dotted red). This pool is divided into further smaller pools and introduced into the cell. The process is repeated until the plasmid causing the biochemical change in question is identified and isolated. In sib selection (solid blue), purification of extrachromosomal DNA from mammalian cell is averted and the plasmid pool exhibiting the phenotype is divided further. The process is repeated until the gene causing the phenotypic change is isolated.

Adapted by permissions from Springer Nature: Springer Nature. NATURE REVIEWS GENETICS. The art and design of genetic screens: mammalian culture cells. Grimm S. 2004;5(3):179-189. doi:10.1038/nrg1291. Copyright © 2004, Springer Nature. License number (Copyright Clearance Center order number) 4703810496842.

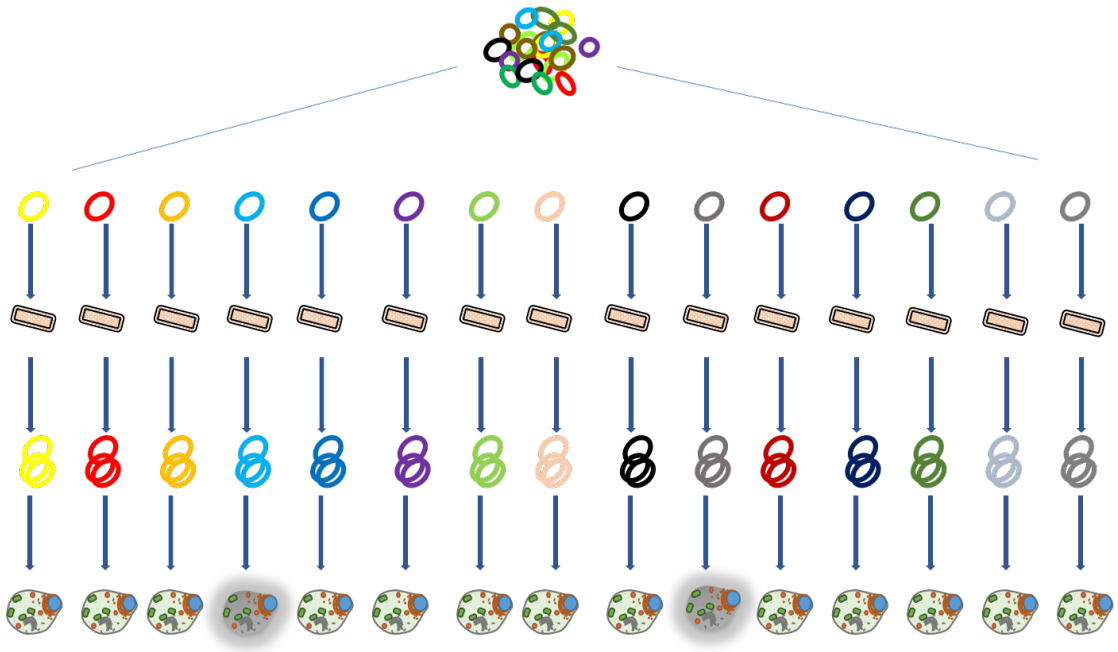


Figure 1.3: **Novel approach to cDNA library screening, Robotic single cDNA investigation (RISCI).** Employing robotic automation, one gene is introduced into the culture cells at a time. First the plasmids are grown in bacteria, DNA is isolated and purified and the culture cells are transfected in the 96-well plate format with one gene per well.

1.8 Functional Genetic Screen for Anticancer Genes

As mentioned earlier, all the known anticancer genes were discovered by chance except for Orctl3. The latter was isolated through a systematic functional genetic screen in Stefan Grimm's lab by Shazia Irshad [42]. In the first step, genes causing cell death in HEK293T cells were isolated from a mouse cDNA library. These genes were transfected into the normal rat kidney (NRK) cells. The genes which passed the second filter, and did not cause cell death in NRK cells, were further transfected into NRK cells transformed by HRAS. Orctl3 was isolated for its ability to cause cell death in HEK293T cells and HRAS-transformed NRK cells but not in the normal NRK cells [42]. Later, Ghada AbuAli capitalized on Shazia Irshad's work and identified adult monkey kidney fibroblasts, CV-1 cells, as the appropriate model system for further investigation of the selective cell-death causing capacity of Orctl3 [58]. Further, Bevan Lin and Qize Ding mixed these two approaches, where they used Shazia Irshad's approach and Ghada AbuAli's model cell culture system, and screened a further set of 22 anticancer genes based on their ability to cause cell death in HEK293T cells and not in the CV-1 cells [59]. In the

first phase, several hundreds of cell-death inducing genes in HEK293T cells were isolated through robotic single cDNA investigation (RISCI) by Bevan Lin [60]. In the second phase, a subset of 377 lethal genes were transfected into CV-1 cells and the ones which did NOT cause cell death were identified as anticancer genes (figure 1.4). It is important to mention here that this is a very simplified explanation of the entire process (see Ref [59] for details). In the final phase, for the confirmation of their anticancer activity, these anticancer genes were transfected into CV-1 cells transformed by oncogenes and mutated tumour-suppressor genes (discussed later). Out of 22 anticancer genes 19 caused cell death in transformed CV-1 cells. The schematic diagram of screening anticancer genes is provided in figure 1.4.

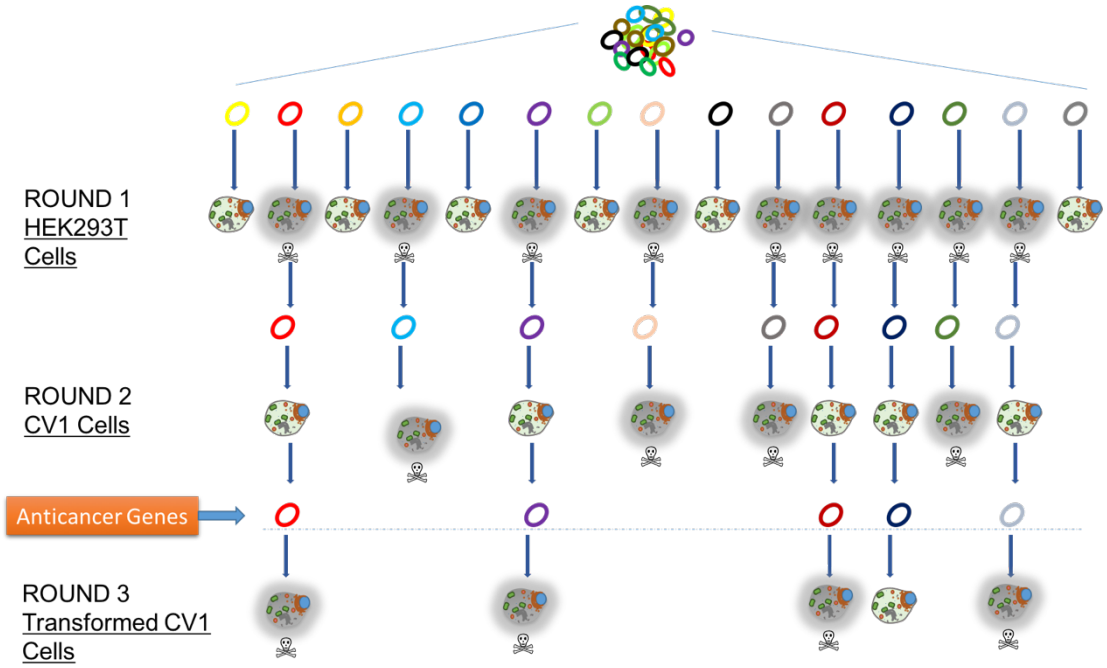


***Figure 1.4*: Screening the anticancer genes.** In the first round, cell-death inducing genes in HEK293T cells were isolated. In the second round, these genes were transfected into CV-1 cells and a subset of the ones which did NOT cause cell death were identified as anticancer genes. In the third round, anticancer genes were transfected into CV-1 cells transformed by oncogenes and mutated tumour-suppressor genes. Nineteen out of 22 genes caused cell death in transformed CV-1 cells.

1.9 Aims

As I started my PhD project with a set of 22 novel anticancer genes, hypothesizing the mechanism of action of each individual gene and building upon proving and disproving these hypotheses were deemed unfeasible. It would have been equally inappropriate to randomly select a subset of genes from the beginning. Emphasis was therefore laid on gradually selecting a workable number from the given set of genes and ultimately deciphering their mode of action. Thus, the aims of my PhD project were:

1. Deciphering the identities of the given anticancer genes (screened by Qize Ding).
2. Verifying their cell death-causing capacity in cancer cells.
3. Investigating the regulation of a subset of anticancer genes in cancer cells.
4. Identifying the type of cell death induced by a subset of these novel anticancer genes.
5. Exploring the mechanism of action of a subset of anticancer genes in cancer cells, transformed cells and normal cells.

1.9.1 Brief Synopsis

After elaborating the materials and methods in the second chapter, the third chapter is primarily focused on deciphering the identities of the given novel anticancer genes, the relationship between the expression levels and survival of cancer patients, verifying the cell death in cancer cells and choosing a subset for further investigation. In the fourth chapter I examine the relationship between a subset of anticancer genes and FOX transcription factors. In the fifth and most interesting chapter I explore the putative mechanism of cell death induction by one of the genes – TMEFF2 – in MCF-7 breast cancer cells. Finally, the sixth chapter focuses on the possible mode(s) of action of another gene, FBLN5, in causing cell death selectively in transformed cells (COS-7) while not in the normal cells (CV-1).

2 Materials and Methods

2.1 Plasmids

The plasmids for anticancer genes were obtained from the Japanese National Institute of Technology and Evaluation (NITE) human cDNA library [61] and were preserved in Glycerol stocks (30% v/v glycerol) of bacteria stored at -80 °C [60]. Bacteria containing plasmids were cultured in Terrific broth (Sigma-Aldrich – T0918) supplemented with 100 µg/mL ampicillin (Sigma-Aldrich – A9518) and Plasmid DNA was purified using PureLink™HiPure Midiprep or Maxiprep plasmid purification kits (Invitrogen® – K210004, K210006) according to manufacturer's protocol. GFP cloned into pME18SFL3 [60] and luciferase cloned into pME18SFL3 [60] were also purified from the bacterial glycerol stock as above.

Other plasmids were obtained through courtesy of sources: human wild type FOXO3 and the FOXO3-A3 always active mutant in pLPC ((Brunet *et al.* [62]; Fernandez de Mattos *et al.* [63]), the human wild type FOXM1 in pcDNA3 and human FOXK2 in pCMV5 (Nestal de Moraes *et al.* [64] [65]) and tBID in pcDNA3 (Mund *et al*. [66]). RIP3K (Myc-DDK-tagged) in pCMV6-Entry was obtained from OriGene Technologies – catalogue number RC209549. These plasmids were transformed into NEB-5α chemically competent *E. coli* (New England BioLabs C2987H) using the standard protocol provided by the manufacturer. Bacteria were cultured and plasmids were purified in the same way as described above (for RIP3K, ampicillin was replaced with kanamycin (Sigma-Aldrich – K4000) in growth media). Only for the experiment presented in figure 6.3, TMEFF2 and FBLN5 were subcloned into pcDNA3.1 (details described later). The pGIZP1 empty vector (Dharmacon™) and shRNA sequences in pGIPZ1 (Dharmacon™) were obtained from the Sigma-Aldrich and MRC LMS Genomics facility, kind courtesy of Dr Adam Cawte. These clones were provided preserved in agar slants and plasmid DNA from these clones were also purified as described above.

The concentrations of DNA preparations were determined by measuring the optical density of the samples at 260 nm using a NanoDrop™ 2000 Spectrophotometer (Thermo scientific).

2.2 Sanger Sequencing

pME18SFL3 plasmids containing anticancer genes were subjected to Sanger sequencing using the primer 5' ATCCGGTGGTGGTGCAAATC 3'. Plasmids were cycle sequenced using BigDye v3.1 (Applied Biosystems Ltd) chemistry at 94°C for 1 minute followed by 30 cycles of 94°C for 10 seconds, 55°C for 15 seconds, 60°C for 4 minutes. After cycling, products were cleaned up by EDTA-Ethanol precipitation, re-suspended in HiDi Formamide and run on a 3730xl DNA Analyser (Applied Biosystems). Sanger sequencing was performed in the MRC LMS Genomics facility.

2.3 Sequence analyses

The resulting sequences obtained for the anticancer genes in pME18SFL3 were subjected to NCBI BLAST against the sequences provided by the NITE library. Alignment was performed using NCBI nucleotide blast two-sequence alignment tool, BLASTN 2.2.32+ [67, 68]. The sequences for the genes which did not align with their corresponding library sequences were then subjected to NCBI nucleotide BLAST against the whole NCBI genomic database [68]. After sequence confirmation, the corresponding library sequences for each gene were subjected to BLAST-like Alignment Tool (BLAT) [69] using the Ensembl genomic interpretation system [70].

2.4 Kaplan-Meier survival estimate

The Kaplan-Meier survival analyses presented in figures 3.26 to 3.32 were plotted using the Kaplan-Meier Plotter webpage http://kmplot.com/analysis/ [71] by following the instructions stated. Default options were selected for all four types of cancer except for the Probe set option where the 'user selected probe set' was selected – a brief overview of the general analysis settings offered by default for each type of cancer tested are as follows (excluding detailed settings for the individual genes tested),

<u>Breast cancer</u>
Survival: RFS
Split patients by: median
Follow up threshold: all
Censore at threshold: checked
Compute median over entire database: false
Probe set option: user selected probe set
Invert HR values below 1: not checked
Remove redundant samples: checked

<u>Ovarian cancer</u>
Survival: PFS
Split patients by: median
Follow up threshold: all
Censore at threshold: checked
Compute median over entire database: false
Probe set option: user selected probe set
Invert HR values below 1: not checked
Restrictions: 'all' selected in each sub-option.

Array quality control: exclude biased arrays Average
Proportional hazards assumption: 0
Cohorts: not selected
Restrictions: 'all' selected in each sub-option.
Remove redundant samples: checked
Array quality control: exclude biased arrays
Proportional hazards assumption: 0

CA125 below lower quartile: not checked
Array quality control: exclude biased arrays

Lung cancer
Survival: OS
Split patients by: median
Follow up threshold: all
Censore at threshold: checked
Compute median over entire database: false
Probe set option: user selected probe set
Invert HR values below 1: not checked
Restrictions: 'all' selected in each sub-option.
Array quality control: exclude biased arrays

Gastric Cancer
Survival: OS
Auto select best cutoff: checked
Follow up threshold: all
Censore at threshold: checked
Compute median over entire database: false
Probe set option: user selected probe set
Invert HR values below 1: not checked
Exclude GSE62254: not checked
Restrictions: 'all' selected in each sub-option.

The http://kmplot.com/analysis/ webpage states the following for how it gathers and analyses the patient data for survival analysis*: 'The background database is manually curated. Gene expression data and relapse free and overall survival information are downloaded from GEO (Affymetrix microarrays only), EGA and TCGA. The database is handled by a PostgreSQL server, which integrates gene expression and clinical data simultaneously. To analyze the prognostic value of a particular gene, the patient samples are split into two groups according to various quantile expressions of the proposed biomarker. The two patient cohorts are compared by a Kaplan-Meier survival plot, and the hazard ratio with 95% confidence intervals and logrank P value are calculated. Each database is updated biannually*'. For the detailed functionality of this webpage and the related databases please see Refs [72, 73, 74, 75, 76].

The Kaplan-Meier survival plot for TEMFF2, presented in figure 5.2 (a) was taken from 'Human Protein Atlas' [77, 78, 79] through the webpage https://www.proteinatlas.org/ENSG00000144339-TMEFF2/pathology/breast+cancer#imid_4069248. The Kaplan-Meier survival plot presented in figure 5.2 (b) was acquired from the Kaplan-Meier Plotter Pan-cancer RNA-seq [72, 80] through the webpage, http://kmplot.com/analysis/index.php?p=service&cancer=pancancer_rnaseq which is a subsidiary webpage/platform of Kaplan-Meier Plotter http://kmplot.com/analysis/. Instead of using microarray data, this webpage utilizes data where RNA-seq was used for the measurement of mRNA expression levels. Default options were selected – which offer the following analysis settings,

RNAseq ID: TMEFF2 -
Survival: OS

Auto select best cutoff: checked
Follow up threshold: all
Censore at threshold: checked
Compute median over entire database: false
Cutoff value used in analysis: 0
Expression range of the probe: 0 - 311
Invert HR values below 1: not checked
Restrictions
Tumor type: Breast cancer
Restrictions
Stage: all
Gender: all
Race: all
Grade: all
Mutation burden: all

2.5 Cell culture

Human breast cancer cell line MCF-7, the adult African green monkey (*Cercopithecus aethiops*) kidney cell line CV-1, and the CV-1 cells transformed using SV40 origin defective DNA, i.e., COS-7 cells (CV-1, origin, SV40-clone), were cultured in the Dulbecco's Modified Eagle's Medium (DMEM)- high glucose with 4500 mg/L glucose, L-glutamine, sodium pyruvate, and sodium bicarbonate (Sigma-Aldrich – D6429) supplemented with 10 % (v/v) heat inactivated foetal bovine serum (FBS) (Life Technologies – 10500064). Cells were maintained in a humidified incubator at 37°C. CV-1 and COS-7 cells were maintained in 5% CO_2 while MCF-7 cells were maintained in 10% CO_2. Upon reaching ~ 70% confluency, cells were sub-cultured following the standard adherent cell sub-culturing procedure.

MCF-7 cells were from ECACC, CV-1 cells were from ATCC and a kind gift from Professor Yaohe Wang at the Barts Cancer Institute London, COS-7 cells were kindly provided by Dr Talat Malik from Professor Marina Motto's group at Imperial College London.

2.6 Plasmid Transfection

2.6.1 Transient Transfection

MCF-7, CV-1 and COS-7 cells were transfected using Xfect™ transfection reagent (Takara-Clonetech – 631318) by following the manufacturer's instructions. Briefly, cells were seeded 24 hours prior to transfection. Transfection was performed in cells having ~ 70% confluency. For MCF-7 cells and CV-1 cells, media was replaced with Opti-MEM™ - Reduced Serum Medium (Life Technologies – 11058021) at the time of transfection. The Opti-MEM™ was replaced with DMEM + 10% FBS, 4 hours post-

transfection. For COS-7 cells, transfection was performed in DMEM + 10% FBS and media was replaced with fresh DMEM + 10% FBS, 8 hours post-transfection. The transfection parameters including number of cells seeded, quantity of plasmid DNA, volume of transfection mix, etc. for MCF-7 cells, CV-1 cells and COS-7 cells in 10 cm dish, 6-well plate, 24-well plate, 96-well plate and ibidi® 8-well chamber are given in table 2.1.

Table 2.1: **Transfection parameters for MCF-7 cells, CV-1 cells and COS-7 cells in different formats.**

10 cm dish			
	MCF-7 cells	**CV-1 cells**	**COS-7 cells**
Number of cells seeded 24 h before transfection	1.2 x 10^6	5.0 x 10^5	4.5 x 10^5
Quantity of plasmid DNA used	30 µg	30 µg	30 µg
Final volume of Xfect transfection mix	600 µl	600 µl	600 µl
Transfection media	Opti-MEM	Opti-MEM	DMEM + 10% FBS
Transfection media volume	6 ml	6 ml	6 ml
Media changed post-transfection	4 h	4 h	8 h
Culture media volume	10 ml	10 ml	10 ml
6-well plate			
	MCF-7 cells	**CV-1 cells**	**COS-7 cells**
Number of cells/well seeded 24 h before transfection	-	7.0 x 10^4	-
Quantity of plasmid DNA used	-	5.0 µg	-
Final volume of Xfect transfection mix	-	100 µl	-
Transfection media	-	Opti-MEM	-
Transfection media volume	1 ml	1 ml	1 ml
Media changed post-transfection	-	4 h	-
Culture media volume	3 ml	3 ml	3 ml
24-well plate			
	MCF-7 cells	**CV-1 cells**	**COS-7 cells**
Number of cells/well seeded 24 h before transfection	4.0 x 10^4	3.0 x 10^4	2.0 x 10^4
Quantity of plasmid DNA used	1.0 µg	1.0 µg	1.0 µg
Final volume of Xfect transfection mix	25 µl	25 µl	25 µl
Transfection media	Opti-MEM	Opti-MEM	DMEM + 10% FBS
Transfection media volume	250 µl	250 µl	250 µl
Media changed post-transfection	4 h	4 h	8 h
Culture media volume	1 ml	1 ml	1 ml
96-well plate			
	MCF-7 cells	**CV-1 cells**	**COS-7 cells**
Number of cells/well seeded 24 h before transfection	8.5 x 10^3	2.0 x 10^3	3.0 x 10^3
Quantity of plasmid DNA used	175 ng	175 ng	175 ng
Final volume of Xfect transfection mix	10 µl	10 µl	10 µl

Transfection media	Opti-MEM	Opti-MEM	DMEM + 10% FBS
Transfection media volume	50 µl	50 µl	50 µl
Media changed post-transfection	4 h	4 h	8 h
Culture media volume	100 µl	100 µl	100 µl
ibidi® 8-well chamber			
	MCF-7 cells	**CV-1 cells**	**COS-7 cells**
Number of cells/squares seeded 24 h before transfection	$2.3x\ 10^4$	8.0×10^3	9.0×10^3
Quantity of plasmid DNA used	500 ng	500 ng	500 ng
Final volume of Xfect transfection mix	12.5 µl	12.5 µl	12.5 µl
Transfection media	Opti-MEM	Opti-MEM	DMEM + 10% FBS
Transfection media volume	125 µl	125 µl	125 µl
Media changed post-transfection	4 h	4 h	8 h
Culture media volume	500 µl	500 µl	500 µl

2.6.2 Stable Transfection

2.6.2.1 Stable transfection of p53 shRNA

Transfections with pGIPZ empty vector (Dharmacon™) and p53 shRNA in pGIPZ (Clone ID V3LHS_333920, target sense sequence 5’ CGGCGCACAGAGGAAGAGA 3’) were performed in 10 cm dishes as described previously. Forty-eight hours post transfection, the cells were added with DMEM + 10% FBS containing 1.0 µg/ml puromycin dihydrochloride (Sigma-Aldrich – P9620), media was changed (containing 1.0 µg/ml puromycin) every third day. After 10 days, the puromycin concentration was reduced to 0.5 µg/ml and the pools were maintained thus forth (in 0.5 µg/ml puromycin). Puromycin was removed from the media 24 hours before the start of any subsequent experiments.

2.6.2.2 Stable transfection of STAT1, FADD or TRAIL shRNAs

Transfections were performed in 10 cm dishes as described previously with pGIPZ empty vector (Dharmacon™) and STAT1 shRNA (Clone ID V2LHS_86643, target sense sequence 5’ GCCCTAAAGGAACTGGATA 3’), FADD shRNA (Clone ID, V3LHS_388202, target sense sequence 5’ CCGATGTCATGGAACTCAG 3’) or TNFSF10 shRNA (Clone ID V3LHS_329177, target sequence 5’ AGTTCAAGAAAAGCAACAA 3’) in pGIZP. After 48 hours of transfection, cells were added with 0.5 µg/ml puromycin dihydrochloride (Sigma-Aldrich P9620) and media was changed every second day

till ~7 days, after which the puromycin concentration was reduced to 0.2 µg/ml and pools of cells maintained thereafter (in 0.2ug/ml puromycin). Puromycin was removed from the media 10 days before the start of any subsequent experiment.

2.7 siRNA Transfection

MCF-7 cells were seeded 24 hours prior to transfection. Transfection was performed in cells at ~70% confluency. MCF-7 cells were transfected with ON-TARGET*plus*™ SMARTpool™ from Dharmacon™; these included siRNA FOXM1 (L-009762-00), siRNA FOXO3 (L-003007-00-0010), siFOXK2 (L-008354-00-0005) and the non-specific control siRNA (Non-Targeting pool D-001810-10) using Oligofectamine™ transfection reagent (Invitrogen 12252-011) by following the manufacturer's instructions. Transfection was performed in Opti-MEM™, media was changed to DMEM + 10% FBS, 6 hours post-transfection. The cells were harvested 48 hours post-transfection for the subsequent experiments.

2.8 $DiOC_6$ + PI Staining- Flow Cytometry- Cell Death Assay

MCF-7 cells were transfected in 24-well plates. After 48 hours since transfection, floating and adherent MCF-7 cells were harvested and centrifuged at 1500 rpm for 5 minutes at 4°C. Cells were re-suspended in 150µl of PBS (Dulbecco's Phosphate Buffered Saline, no calcium, no magnesium Sigma-Aldrich – 14190169) containing 40nM 3,3-dihexaoxacarbocyanine iodide ($DiOC_6$) (Life Technologies – D273) and 6 µg/ml propidium iodide (PI) (Sigma-Aldrich – P4864), incubated at 37 °C for 30 minutes and further incubated for 30 minutes at room temperature. Cells were then subjected to flow cytometry in a FACSCalibur™ (Becton, Dickinson & Company) with active excitation 488 nm laser. $DiOC_6$ staining was detected using FL-1 channel and FL-3 channel was used for PI staining. GFP in pME18SFL3 was transfected in parallel and transfection efficiency measured using flow cytometry single channel analysis (FL-1). The data analysis was performed using FlowJo® (FlowJo LLC). Readings from duplicate experiments (2 wells of a 24-well plate) were first adjusted to un-transfected cells, further normalised to the transfection efficiency, averaged and standard deviation calculated using Microsoft Excel.

2.9 Colorimetric Methylthiazole Tetrazolium (MTT) Assay

MCF-7 cells were seeded in a 96-well plate at a density of 3000 cells/well and transfected as described previously. Please note that less cells were seeded than mentioned in table 2.1 in order to avoid contact inhibition over time. Cell proliferation was assessed by first removing the media, adding 120 µl of 833 µg/ml MTT (Sigma-Aldrich – M5655) in DMEM+10%FCS to each well and incubating at 37°C and 5% CO_2 for 3.5 hours. The MTT solution was replaced with 150 µl of 4 mM HCl, 0.1% NP-40 (Sigma-Aldrich – I3021) in isopropanol. Plate was shaken at 200 rpm on an orbital shaker at room temperature for 10 minutes and OD at 590 nm was measured using a microplate reader (FLUOstar optima). Cell proliferation was measured 6, 24, 48 and 72 hours post-transfection. The 6-hours' time point was regarded as 0 hours. Transfections were performed in triplicate, i.e., three wells of a 96-well plate. The ODs obtained were averaged and normalized to the 0 hours time point.

2.10 Immunoblotting

Adherent cells were harvested by washing once with PBS followed by incubation in Trypsin-EDTA solution (Sigma-Aldrich – 59418C) for 5 minutes at 37° C. Trypsin was neutralized by adding 9 volumes of DMEM (Sigma-Aldrich – D6429) supplemented with 10% FBS (Life Technologies – 10500064). Cell suspension was centrifuged at 250 *g* at 4°C for 5 min, supernatant discarded, and cell pellets were flash frozen on dry ice and stored at -80° C. Frozen pellets were lysed on ice for 20 minutes in 3 volumes of RIPA lysis buffer (Thermo Scientific – 89900, containing 25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate and 0.1% SDS), supplemented with Halt™ protease inhibitor (Thermo Scientific – 87786 containing AEBSF, Aprotinin, Bestatin, E64, Leupeptin, Pepstatin A, EDTA) and phosphatase inhibitor (Sigma-Aldrich – P5726 containing Sodium orthovanadate, Sodium molybdate, Sodium tartrate, Imidazole). The lysate was centrifuged at 2800 *g* at 4°C for 10 minutes. The supernatant was transferred to a new pre-chilled tube and its protein concentration was assayed using Pierce™ BCA Protein Assay Kit (Thermo Scientific – 23227) following the manufacturers' instructions.

For separation of the proteins, 10 µg total protein was mixed with 1/4 volume of 4X TruPAGE™ LDS sample buffer (Sigma-Aldrich – PCG3009) supplemented with DTT (Sigma-Aldrich – PCG3005) and heated at 70°C for 10 min. Proteins were fractionated by electrophoresis through Laemmli [81] SDS-PAGE (7%, 10% or 12.5% resolving and 4% or 6% stacking gels) for 2 hours in Tris/Glycine/SDS running

buffer (Biorad – 1610772, 25 mM Tris-base, 190 mM Glycine, 0.1% (w/v) SDS) under reducing conditions using a Mini-Protean III apparatus (Bio-Rad Laboratories). Proteins were separated alongside PageRule™ pre-stained SDS-PAGE markers (Thermo Scientific – SM0671). The reagents used for preparing gels include 30% (w/v) acrylamide (Protogel – EC-890), 25% (w/v) APS (Sigma-Aldrich – A3678), TEMED (Sigma-Aldrich – T9281), 10% (w/v) SDS solution (Biorad – 1610416), resolving gel buffer 1.5 M Tris-HCl pH-8.8 (Biorad – 1610798) and stacking gel buffer 0.5 M Tris-HCl pH-6.6 (Biorad – 1610799).

Following separation via SDS-PAGE, proteins were transferred to a 0.45 µm nitrocellulose membrane (Biorad – 1620115) by electrophoretic transfer in a wet tank blotting system (Mini Trans-Blot Electrophoretic Transfer Cell, Bio-Rad). The transfer was carried out in transfer buffer (Bio-rad – 1610771 comprising 25 mM Tris, 190 mM glycine and 23% (v/v) ethanol) for ~ 1 hour and 45 minutes at 90 V at room temperature with a block of ice in the transfer tank. The nitrocellulose membrane was then blocked by incubating the membrane for 1 hour at room temperature in 5% (w/v) Bovine Serum Albumin (BSA)(Sigma-Aldrich – A7906) in Tris Buffered Saline Tween 20 (TBST) (20 mM Tris-HCl (Sigma-Aldrich – T1503), 150 mM NaCl (Sigma-Aldrich – 53014), pH 7.4 supplemented with 0.05% (v/v) Tween 20 (Sigma-Aldrich – P2287) and 0.02% (w/v) sodium azide (Sigma-Aldrich – S8032)). The membrane was then incubated overnight at 4°C with the primary antibody under gentle shaking. Primary antibodies were dissolved in blocking solution. The membrane was washed five times in TBST (three washes of 5 minutes followed by two washes of 10 minutes each) followed by 30 min of incubation with the horseradish peroxidase (HRP) conjugated secondary antibody. Following 5 more washes in TBST, the membrane was visualized using the enhanced chemiluminescence (ECL - Perkin Elmer 203-16281). The membrane was exposed to an auto-radiographic CL-XPosure clear blue x-ray films (Thermo-fisher 34093) for times varying from 10 seconds to 1 hour.

***Table 2.2**:* **List of antibodies used for immunoblotting.**

Antibody	Concentration	Catalogue no.	Species	Company
FOXO3	1:1000	SC11351	Rabbit	Santa Cruz
TMEFF2	1:4000	Ab133562	Rabbit	Abcam
FBLN5	1:500	Ab66339	Mouse	Abcam
YPEL4	1:1000	Ab176280	Rabbit	Abcam
NEIL2	1:25000	Ab180576	Rabbit	Abcam
HYAL2	1:5000	Ab68608	Rabbit	Abcam
SMLR1	1:1000	ab204982	Rabbit	Abcam

GAPDH	1:160000	SC32233	Mouse	Santa Cruz
TP53	1:500	SC126	Mouse	Santa Cruz
PUMA	1:500	PC686	Rabbit	Merck Millipore
MDM2	1:200	SC965	Mouse	Santa Cruz
PARP-1	1:1000	9542	Rabbit	Cell signalling
BCL2	1:200	Ab7973-1	Rabbit	Abcam
MCL1	1:200	SC819	Rabbit	Santa Cruz
FOXK2	1:1000	A301-729A	Rabbit	Blyth Laboratories
FOXM1	1:1000	C-20	Rabbit	Santa Cruz
Anti-Mouse	1:65000	P0447	Goat	Dako
Anti-Rabbit	1:65000	A0545	Goat	Sigma-Aldrich

2.11Quantitative Real time PCR (qRT-PCR)

2.11.1 RNA extraction

For data presented in figures 4.2 and 4.4 , cells were harvested by washing once with PBS followed by incubating in Trypsin-EDTA (Sigma-Aldrich – 59418C) for 5 minutes at 37° C. Trypsin was neutralized by adding 9 volumes of DMEM + 10% FBS (Sigma-Aldrich – D6429 and Life Technologies – 10500064). Cell suspension was centrifuged at 1600 RPM at 4°C. Cell pellets were flash frozen on dry ice and stored at -80°C. Total RNA was extracted using RNeasy Mini Kit (Qiagen® – 74104) following the manufacturer's instructions. Cell homogenization was performed using QIAshredder spin column (Qiagen® – 79654) and optional DNA digestion was performed using On-Column RNase-Free DNase set (Qiagen® – 79254).

For the data presented in figures 5.10, 5.12, 5.25, 6.4, 6.5, 6.8 as well as for all microarray analyses, RNA extraction was performed as follows: cells were transfected in 24-well plates. Media was removed and cells were washed once with PBS. Cell lyses was carried out by adding 100 µl Buffer RLT Plus (Qiagen® – 1053393) to each well. Lysates from 12 wells of 24-well plate were pooled and RNA purified using RNeasy Plus Mini Kit (Qiagen® – 74134) following manufacturer's instruction. RNeasy Plus Mini Kit employs gDNA Eliminator Mini Spin Columns. In addition, optional DNA digestion was performed using On-Column RNase-Free DNase set (Qiagen® – 79254).

2.11.2 cDNA Synthesis

The concentrations of RNA preparations were determined by measuring the optical density of the samples at 260 nm (OD260) using a NanoDrop 2000 Spectrophotometer (Thermo Scientific). RNA integrity was confirmed using agarose gel electrophoresis. 1 µg of RNA was used for subsequent cDNA synthesis. cDNA was generated using QuantiTect Reverse Transcription Kit (Qiagen® – 205311). The reaction mix was prepared following the manufacturer's instructions except for the following modifications: before adding the Quantiscript Reverse Transcriptase (Qiagen®), reaction mix was incubated at 65° C for 5 min, then placed on ice for 5 min. Quantiscript Reverse transcriptase was added followed by incubation at 42° C for 15 min and at 50° C for another 15 min followed by incubation at 95° C for 3 min and reaction mix was then moved on ice.

Only for the data presented in figure 4.4 (a) and 4.4 (b), the above procedure was followed except, instead of using the random primer mix provided by the manufacturer, TMEFF2 reverse primer 5' GACCGGGAACAACGTATAGAAC 3' and L19 reverse primer 5' GCAGCCGGCGCAAA 3' were added to the reverse transcription reaction.

2.11.3 qRT-PCR Reaction

Real time PCR was performed using the Power SYBR™ Green PCR MasterMix (Applied Biosystems – 4368577) and the 7900HT Fast Real-Time PCR System (Applied Biosystems). All measurements were performed in a 384-well plate format. The PCR cycling was as follows: 95°C for 10 min and 40 cycles of (95°C for 15 sec followed by 60°C for 1 min), followed by the dissociation curve analysis step. The list of primers used for real time PCR is provided in table 2.3.

2.11.4 Data analysis

2.11.4.1 Relative quantification using $\Delta\Delta C_T$ method

C_T values, average C_T values and C_T values standard deviation were generated by SDS software (Applied Biosystems) integrated with the 7900HT Fast Real-Time PCR System (Applied Biosystems). From the average C_T values provided by the SDS software, fold change was calculated using $\Delta\Delta C_T$ method by first working out ΔC_T using the formula, $\Delta \boldsymbol{C_T} = \boldsymbol{C_{T\,target}} - \boldsymbol{C_{T\,reference}}$ where L19 was used as reference. Subsequently, $\Delta\Delta C_T$ was calculated using the formula, $\Delta\Delta \boldsymbol{C_T} = \Delta \boldsymbol{C_{T\,test\,sample}} -$

$\Delta \boldsymbol{C_{T\ control\ sample}}$. ΔC_T standard deviations (SD) were calculated using the formula $\boldsymbol{SD} = \sqrt{\boldsymbol{SD_{target}}^2 + \boldsymbol{SD_{reference}}^2}$. In order to integrate the standard deviation in the relative fold change, first maximum and minimum $\Delta\Delta C_T$ was calculated as $\Delta\Delta \boldsymbol{C_{T\,Max}} = \Delta\Delta \boldsymbol{C_T} + \boldsymbol{SD}$ and $\Delta\Delta \boldsymbol{C_{T\,Min}} = \Delta\Delta \boldsymbol{C_T} - \boldsymbol{SD}$ respectively. Following this step, the maximum and minimum fold change was calculated using formula $\boldsymbol{Fold\ change_{Max}} = \boldsymbol{2^{-\Delta\Delta C_{T\,Max}}}$ and $\boldsymbol{Fold\ change_{Min}} = \boldsymbol{2^{-\Delta\Delta C_{T\,Min}}}$ respectively. Lastly, $\boldsymbol{Fold\ change_{Max}}$ and $\boldsymbol{Fold\ change_{Min}}$ were averaged and standard deviation calculated using Microsoft Excel. NB: standard curves were generated and comparative reaction efficiencies between the target and reference confirmed before applying $\Delta\Delta C_T$ method. The above-mentioned calculation procedure was adapted from Ref. [82].

Statistical significance (p value) was calculated using the two-sample T-test assuming unequal variance (Welch's T-Test) (two -tailed) using the webpage http://www.statskingdom.com/150MeanT2uneq.html. and confirmed (significance only) through webpage https://www.graphpad.com/quickcalcs/ttest1/?Format=SD. The average fold change and standard deviation along with the number of samples were provided to the above-mentioned webpages/tools. For the latter, the options of 'Welch's unpaired t test' and 'Enter mean, SD and N' were selected.

2.11.4.2 Relative Quantification using Relative Standard Curve Method

Only for the data presented the figures 6.4 and 6.5, the relative standard curve method was used to estimate relative difference in mRNA levels. First a representative solution was prepared by mixing equal volumes of cDNA from all the samples to be analysed. A 4-fold dilution series was prepared with this mixture. A standard curve (C_T value *vs* dilution) was then generated using the SDS software (Applied biosystems). Quantity means were calculated through this standard curve by SDS software along with the quantity standard deviations. The final relative mRNA values were calculated by dividing the quantity mean for test sample by quantity mean for L19 reference gene. Standard deviations for test and reference genes were combined by first calculating the coefficient of variation (CV) for target and reference genes using the formula $\boldsymbol{CV_{target}} = \frac{\boldsymbol{Quantitiy\ Mean_{target}}}{\boldsymbol{Quantity\ Standard\ Deviation_{target}}}$ and $\boldsymbol{CV_{reference}} = \frac{\boldsymbol{Quantitiy\ Mean_{reference}}}{\boldsymbol{Quantity\ Standard\ Deviation_{reference}}}$, CV of quotient was then calculated using the formula $\boldsymbol{CV\ of\ quotient} = \sqrt{\boldsymbol{CV_{target}}^2 + \boldsymbol{CV_{reference}}^2}$. Standard deviation (SD) for the test

sample was then calculated using the formula $\boldsymbol{SD = CV\ of\ quotient\ X\ Quantity\ mean}$. The above mentioned calculation procedure was followed from Ref [82].

Table 2.3: **List of primers used for qRT-PCR.**

Gene	Forward primer 5'-3'	Reverse primer 5'-3'
FOXO3	TCTACGAGTGGATGGTGCGTT	CGACTATGCAGTGACAGGTTGTG
FOXM1	TGCAGGCTAGGGATGTGAATCTTC	GGAGCCCAGTCCATCAGAACT
FOXK2	ACGTCTCCAGTGAAGGCCGTA	CCACCTTGTACCCTGAAGA
HYAL2	ACTCCCAGTCTACGTCTTCA	CTCGCCAATGGTAGAGATGAG
NEIL2	GTCTGCTATACACTGCTGGAC	GAGGCACTCAGGACTGAAC
FBLN5	AAGCCACGACCCGCTAC	GTCATCACCAGGGTGGCA
TMEFF2	CTATCAATATGCAGGAGCCATCT	GACCGGGAACAACGTATAGAAC
SMLR1	TGTGACTGTGTCTGCTCTAAAC	CTCTACCAAGGTTTCTTCACCTC
P53	CCGCAGTCAGATCCTAGCG	AATCATCCATTGCTTGGGACG
TNFAIP3	GGACGAGCAAGTGCAGAAA	TGGGACTGACTTTCCCTGA
TNFSF10	GTGGCATTGCTTGTTTCTTAAA	TCTAACGAGCTGACGGAGTT
HLA-F	ATCGTTGCTGGCCTTGT	CTGCAGCCTGAGAGTAGC
BIRC3	GGAGAGAGAAAGAGCAACTGAG	TGCCATTCTATTCTTCCGGATT
IFI27	GGTGAGGTCAGCTTCACATT	GTTAATTCCGTGGCCTAGAGAG
CXCL10	CTGTACCTGCATCAGCATTAGTA	GACATCTCTTCTCACCCTTCTTT
L19	GCGGAAGGGTACAGCCAAT	GCAGCCGGCGCAAA

2.12 Caspase-Glo® 3/7, caspase activity assay

MCF-7 cells were transfected in 96-well plate (white with clear bottom), 48 hours post transfection, 100 µl of Caspase-Glo® 3/7 Reagent (Promega® – G8091) was added to each well on top of 100 µl of DMEM + 10% FBS already present. The plate was shaken at 450 rpm for 10 sec and incubated for 1 hour at room temperature. The relative luminescence was read using PHERAstar plus (BMG Labtech) plate reader with LUMplus (luminometer) module.

2.13 TMRE + DRAQ7™ Staining- Flow Cytometry- Cell Death Assay

MCF-7 cells were transfected in 24-well plates and 48 hours post-transfection floating and adherent cells were harvested, centrifuged at 1500 rpm for 5 min at 4° C, supernatant discarded, and the cell pellet resuspended in 200 µl Phosphate Buffered Saline (PBS) containing 10 nM Tetramethylrhodamine ethyl ester perchlorate (TMRE) (Sigma-Aldrich – 87917) and 0.5 µl DRAQ7™ (Abcam – ab109202), incubated for 25 min at 37 ° C and another 25 min at room temperature. Cells

were subjected to flow cytometry using LSR II flow cytometer (Becton, Dickinson & Company) with active 405 nm, 488 nm, 561 nm and 633 nm excitation lasers. TMRE staining was detected using 610/20 YG filter and DRAQ7™ staining was detected using 695/40 R filter. In the experiments involving concurrent measurement of GFP fluorescence, GFP was detected using 525/50 B filter. Data was collected using CellQuest™ Pro (Becton, Dickinson & Company) and analysed using FlowJo® (FlowJo LLC).

2.13.1 Caspase inhibition using z-VAD-FMK

MCF-7 cells were transfected in 24-well plates. At 4 hours post-transfection, i.e., time of media change, Opti-MEM™ was replaced with DMEM + 10% FBS containing 50 µM Z-VAD (OMe)-FMK (MP Biomedicals – FK009). Cells were harvested 48 hours post-transfection and assayed for cell death as described above.

2.13.2 TP53 inhibition using Pifithrin-Alpha

MCF-7 cells were transfected in 24-well plates. At 4 hours post transfection, i.e., time of media change, Opti-MEM™ was replaced with DMEM + 10% FBS containing 20 µM PFT-α (Sigma-Aldrich – P4359). Cells were harvested 48 hours post-transfection and assayed for cell death as described above.

2.14 Live cell Imaging

Poly-D-Lysine coated, 35 mm glass bottom dishes (ibidi GmbH) with 4-well silicone insert (ibidi GmbH) were used for live imaging. In each well, 8.5 x 10^3 MCF-7 cells were seeded 24 hours prior to transfection. Transfection was performed using Xfect™ transfection reagent – 96-well plate format was used (table 2.1). At 4 hours post-transfection, Opti-MEM™ was replaced with 100 µl DMEM + 10% FBS and the dish was moved to the stage top incubator (Tokai Hit) on the Zeiss Elyra wide-field microscope which was maintained at 37° C and 5% CO_2 under high relative humidity. Imaging was started at 6 hours post-transfection. Images were captured at 20 min interval till 92 hours post-transfection. Excitation at 488 nm with 10 mW laser power and 50 ms exposure time was used, emission was detected with a 63x oil immersion objective with a high numerical aperture of 1.4 at 495-550 nm for GFP and > 650 nm for phase contrast. Phase contrast images used a mercury lamp. at 8%. Five Z-stacks were acquired for each frame and alternation between GFP and bright field

channels was employed for each individual frame. Image analysis of the Z-stack time series was performed using Fiji – 'batteries-included' image processing package [83].

2.15 Immunofluorescence

Transfections were performed in 8-well glass bottom chamber (ibidi GmbH) coated with Poly-D-Lysine as described previously (table 2.1). Cells were fixed with 4% (v/w) paraformaldehyde (Sigma-Aldrich – 28908) in PBS for 10 min at room temperature, after washing once with PBS, cells were permeabilized in 0.2% (v/v) Triton X-100 (Sigma-Aldrich – T8532) for 10 min at room temperature. Cells were blocked for 1 hour at room temperature in 2% (w/v) bovine serum albumin in PBS. Cells were incubated overnight at 4° C in primary antibody (antibodies were dissolved in blocking solution). Cells were washed five times with PBS for 20 minutes each, incubated in secondary antibody (dissolved in blocking solution) for 1 hour at room temperature and washed again five times for 20 min each with PBS. When co-immunostaining was performed, cells were blocked again for 1 hour at room temperature in 2% (w/v) BSA in PBS, incubated in primary antibody solution overnight at 4° C, washed five times with PBS for 20 minutes each, incubated in a secondary antibody for 1 hour and washed five times with PBS for 20 min each. In either case, after the final wash, cells were covered with VECTASHIELD® antifade-mounting-medium-with-DAPI (Vector laboratories – H-1200). Imaging was performed using Zeiss Elyra wide-field microscope with excitation lasers at 405 nm (Blue), 488 nm (Green), 561 nm (Red) and 642 nm (Far-Red). Emission was detected with a 63x oil immersion objective with a high numerical aperture of 1.4 and band-pass filters at 420-480 nm, 495-550 nm, 570-640 nm and a long-pass filter > 650 nm respectively. Five Z-stacks were acquired for each frame. Image analysis of the Z-stack time series was performed using Fiji – 'batteries-included' image processing package [83]. Primary antibodies used were: Anti-TMEFF2 (Abcam – Ab133562 | 1:200 dilution), Anti-FBLN5 (Abcam – ab66339 | 1:300 dilution), Anti-ATPB (Abcam – ab14730 | 1:200), Anti-α-Tubulin (Abcam – ab7291 | 1:1400 dilution). Secondary antibodies used were: Donkey Anti-mouse Alexa Fluor 647 (Molecular Probes – A-31571 | 1:1000 dilution), Donkey Anti-Rabbit Alexa Fluor 647 (Molecular Probes – A-31573 | 1:1000 dilution) and Donkey Anti-mouse Alexa Fluor 488 (Molecular Probes – R37114 | 1:1000 dilution).

2.16 Microarray Analyses

2.16.1 Microarray Processing

Affymetrix® GeneChip™ Clariom™ D human microarrays were purchased from Thermo Fisher Scientific (catalogue number 902922). RNA extraction and purification procedures for the microarray experiments have already been elaborated. Quantity and integrity of RNA was measured using the RNA Pico Chip (Agilent Technologies). For each sample, 10 ng of RNA was converted into labelled cDNA using the NuGEN Ovation® Pico WTA System V2 (NuGEN Technologies Inc.), followed by biotinylating step using the NuGEN Encore® biotin Module (NuGEN Technologies Inc.). Seven mg of labelled cDNA was hybridised to Affymetrix GeneChip™ Clariom™ D human microarrays for 20 hours at 45° C The microarrays were washed, stained (GeneChip Fluidics Station 450) and scanned (GeneChip Scanner 3000 7G) according to the manufacturer's instructions (NuGEN Technologies Inc and Affymetrix ®).

2.16.2 Data analysis

2.16.2.1 Differential Gene Expression Analysis

The data processing for microarray (quality control checks and subsequent statistical analysis) was conducted using Transcriptome Analysis Console v 4.0.1 (TAC) (Applied biosystems) following the instructions from the developer [84]. According to the developer [84], '*TAC serves as a friendly graphical interface to a set of R and Bioconductor modules Including Limma, EventPointer, Dbscan Rtsne and Apcluster*' (p.14).

The CEL files were imported into Transcriptome Analysis Console. For the differential gene expression profiling of MCF-7 cells transfected with control *vs* MCF-7 cells transfected with TMEFF2, the analysis parameters were as follows:

- Array type: Clariom_D_Human
- Analysis type: expression (Gene + Exon)
- Analysis version: version 2
- Summarization method: Gene + Exon - SST-RMA
- Pos *vs* Neg AUC Threshold: 0.7
- Genome version: hg38 (Homo sapiens)
- Annotation: Clariom_D_Human.r1.na36.hg38.a1.transcript.csv

- Map file: Clariom_D_Human.r1.map
- Condition (comparison): MCF7 control; MCF7 TMEFF2
- Undergoing apoptosis (normalization group): NO, YES

Expression Analysis Settings:

• Gene-level fold change < -2 or > 2

• Gene-level p-value < 0.05

• Splicing index < -2 or > 2

• Exon-level p-value < 0.05

• Anova method: ebayes

• A Probeset (Gene/Exon) is considered expressed if ≥ 50% samples have DABG values below DABG threshold.

• DABG < 0.05

• Event algorithm method: both

• Event pointer p-value < 0.1

• Event score > 0.2

For the differential gene expression profiling of CV-1 and COS-7, wild type and transfected with FBLN5, the analysis parameters were as follows:

- Array type: Clariom_D_Human
- Analysis type: expression (Gene + Exon)
- Analysis version: version 2
- Summarization method: Gene + Exon - SST-RMA
- Pos *vs* Neg AUC threshold: 0.7
- Genome version: hg38 (Homo sapiens)
- Annotation: Clariom_D_Human.r1.na36.hg38.a1.transcript.csv
- Map file: Clariom_D_Human.r1.map
- Condition (comparison): COS7 Control; COS7 FBLN5; CV1 Control; CV1 FBLN5
- Transfection (normalization group): transfected, untransfected

Expression Analysis Settings:

• Gene-level fold change < -2 or > 2

• Gene-level p-value < 0.05

• Splicing index < -2 or > 2

• Exon-level p-value < 0.05

• Anova method: ebayes

• A Probeset (Gene/Exon) is considered expressed if ≥ 50% samples have DABG values below DABG threshold.

• DABG < 0.05

• Event algorithm method: both

• Event pointer p-value < 0.1

• Event score > 0.2

2.16.2.2 Gene Ontology Analyses / Pathways analyses

Two separate bioinformatic tools were employed for pathways analyses. Firstly, the Transcriptome Analysis Console has integrated WikiPathways tool which employs Fisher's Exact Test to overlap the relevant molecules in the differential gene expression analysis with the canonical pathways in the WikiPathways database [85].

Secondly, the pathway analysis, networks, functional analyses and comparison analyses were conducted through the use of Ingenuity Pathway Analysis - IPA (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis). The details about the functioning of IPA can be found in Ref [86].

For the pathway analysis of differentially expressed genes upon TMEFF2 overexpression in MCF-7 cells, the list of differentially expressed genes was imported into IPA and gene identities were indexed using Affymetrix probe set ID as first preference, gene symbols as second preference and public gene IDs as third preference (more than 90% of transcripts were identified through Affymetrix probe set IDs). Following the recommendation of the software provider, first a list of differentially expressed genes with slightly loose filter (fold change ≥ 1.5 to -1.5≤, $p < 0.05$ and FDR < 0.25) was provided to IPA. Moreover, the transcripts type provided to IPA excluded 'non-coding' transcripts. 'Multiple_complex', 'Precusor_MicroRNA', 'Small_RNA', 'Pseudogenes', 'Ribosomal', 'tRNA' and 'Unassigned' transcripts were included in the list provided to IPA along with the 'Coding' transcripts. The analysis parameters were as follows:

- Analysis type: core analysis
- Reference set: Clariom D Human transcript array
- Relationships to consider: direct and indirect relationships
- Interaction networks: yes
- Include endogenous: molecules per network = 70, networks per analysis = 25
- Node types: all
- Data sources: all
- Confidence: experimentally observed only
- Species: all
- Tissue & cell lines: all
- Mutations: all
- Expression fold change cut-off: -2.0 down and 2.0 up
- p-value cut-off: 0.005

- FDR p-value cut-off: 0.05

For the pathway analysis of differentially expressed genes between wild type CV-1 cells *vs* CV-1 cells transfected with FBLN5, wild type COS-7 cells *vs* COS-7 cells transfected with FBLN5 and wild type COS-7 cells *vs* wild type CV-1 cells, the import and indexing method used was the same as described above for the TMEFF2 overexpressing MCF-7 cells. For these analyses, the differentially expressed genes with slightly loose filter (fold change ≥ 1.5 to -1.5 ≤, $p < 0.05$ and FDR < 0.25) were provided to IPA. The transcripts type provided to IPA excluded 'non-coding' transcripts. 'Multiple_complex', 'Precusor_MicroRNA',' Small_RNA', 'Pseudogenes', 'Ribosomal', 'tRNA' and 'Unassigned' transcripts were included in the list provided to IPA along with the 'Coding' transcripts. The following parameters were applied for the three individual pathways analyses:

- Analysis type: core analysis
- Reference set: Clariom D Human transcript array
- Relationships to consider: direct and indirect relationships
- Interaction networks: yes
- Include endogenous: molecules per network = 70, networks per analysis = 10
- Node types: all
- Data sources: all
- Confidence: experimentally observed only
- Species: all
- Tissue & cell lines: all
- Mutations: all
- Expression fold change cut-off: -2.0 down and 2.0 up
- p-value cut-off: 0.05
- FDR p-value cut-off: 0.05

Following the completion of three individual core analyses, a 'comparison analysis' was performed between these three core analyses.

2.17 Phenotype Inspection cell death assay

CV-1 and COS-7 cell were co-transfected with GFP and either FBLN5, TMEFF2, luciferase or tBID in 1:4 ratio (1 GFP: 4 test plasmid) in 96-well plates using the Xfect™ transfection reagent as described previously (table 2.1). At 48 hours post-transfection, green fluorescent cells were imaged using the IN Cell Analyzer 2000 (GE Healthcare) with 20x objective and FITC wavelength filter. 36 fields in each

well of a 96-well plate were imaged. ~ 3000 COS-7 cells and ~ 2500 CV-1 cells were counted manually for each condition. Three independently transfected populations from three different wells of a 96-well plate were counted for each condition. Rounded, clustered and/or cells with disfigured morphology were considered as dead.

2.18 Automated Cell Counting Assay

CV1 and COS-7 cell were co-transfected with GFP and either FBLN5, TMEFF2, luciferase or tBID in 1:4 ratio (1 GFP: 4 test plasmid) in 96-well plates. At 24 and 48 hours post-transfection, media was removed and cells were fixed with 4 % paraformaldehyde. After washing twice with PBS, green fluorescent cells were imaged using the IN Cell Analyzer 2000 (GE Healthcare) with 10x objective and FITC wavelength filter. Nine fields in each well of a 96-well plate were imaged. Green fluorescent cells were counted through automation using Fiji – 'batteries-included' image processing package by first adjusting brightness and contrast (manually) followed by converting the image to mask and subsequently using the 'Analyse Particles' function [83].

2.19 Gene Sub-Cloning

FBLN5 and TMEFF2 were cloned into pcDNA3.1 vector (Invitrogen®). For subcloning of FBLN5, the gene was excised from the pME18SFL3 vector through restriction digest using *Eco*RI (FastDigest - Thermo Scientific – FD0274) and *Xba*I (FastDigest - Thermo Scientific – FD0684) following manufacturer's instructions. pcDNA3.1 empty vector was also subjected to same restriction digest. The DNA fragments were separated through DNA electrophoresis (1.0% w/v agarose Sigma-Aldrich – A9539, dissolved in TAE buffer Fermentas – B49, with 1:10000 Gel Red™ Biotium-UK – 41002), excised from the agarose gel and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega® – A9282) according to manufacturer's instructions. The sticky-end ligation was performed using T4 DNA ligase HC (Thermo Scientific – EL0013) following the manufacturer's instructions. Ligated DNA was transformed into NEB-5α chemically competent *E. coli* (New England BioLabs – C2987H) using the standard protocol provided by the manufacturer. Following the growth of colonies on Lysogeny (LB) – agar in the presence of 100 µg/mL ampicillin (Sigma-Aldrich – A9518), 10 clones were screened for the presence of FBLN5 insert using standard molecular biology techniques (miniprep, restriction digest and separation through DNA electrophoresis). The clone containing FBLN5 insert was grown in Terrific broth (Sigma-Aldrich – T0918) supplemented with 100 µg/mL ampicillin and plasmid DNA

purified using PureLink™HiPure Midiprep or Maxiprep plasmid purification kits (Invitrogen – K210004, K210006) according to manufacturer's protocol. The sequence of cloned gene was confirmed through Sanger sequencing using CMV-Forward primer 5' CGCAAATGGGCGGTAGGCGTG 3'.

For subcloning of TMEFF2 into pcDNA 3.1, first the gene was amplified through PCR using the forward primer 5' CATCATCTCGAGGCCGCCACCATGGTGCTGTGG 3' and reverse primer 5' CTGGCTCCGTACTTCTAGAGTGAAACATGCTCCCT 3' using TMEFF2 in pME18SFL3 as template. PCRBIO HiFi Polymerase (PCR Biosystems Ltd – PB10.41-02) was used following the manufacturer's instructions. The PCR product was purified using Wizard® SV Gel and PCR Clean-Up System (Promega® – A9282) according to manufacturer's instructions. Both the purified PCR product and pcDNA 3.1 empty vector were then subjected to restriction digest using *Xho*I (FastDigest - Thermo Scientific – FD0695) and *Xba*I (FastDigest - Thermo Scientific – FD0684) by following manufacturer's instructions. The subsequent procedures of DNA separation, purification, ligation, bacterial transformation, screening, plasmid purification and Sanger sequencing were the same as described above for FBLN5 cloning.

2.20 Crystal Violet staining

Media was removed and the adherent cells were washed once with PBS. Cells were fixed with 4% (w/v) paraformaldehyde (Sigma-Aldrich – 28908) in PBS for 15 min at room temperature. After washing once with deionized water, cells were incubated with 0.2% (w/v) Crystal Violet (Sigma-Aldrich C3886) for at least 2 hours at room temperature. Plates were immersed in running tap water for 15 min. Plates were dried overnight and imaged using a conventional table-top scanner.

3 Identification and Selection

3.1 Background

As stated in the Introduction, 22 novel anticancer genes were isolated through phased systematic screening by Qize Ding [59]. In the first phase, several hundreds of cell-death inducing genes in HEK293T cells were isolated through robotic single cDNA investigation (RISCI) [57]. In the second phase, these lethal genes were transfected into CV-1 cells and a subset of the ones, which did NOT cause cell death, were identified as anticancer genes. In the final phase, for the confirmation of their anticancer activity, these anticancer genes were transfected into CV-1 cells transformed by oncogenes and mutated tumour-suppressor genes (table 3.1). Nineteen out of 22 anticancer genes caused cell death in transformed CV-1 cells.

The human cDNA library used for genetic screening was obtained from the Japanese National Institute of Technology and Evaluation, NITE. The NITE library consists of approximately 30,021 individual cDNA clones [60]. 68% of NITE library cDNA clones came directly from the FLJ library collection. FLJ is short for 'Full Length Japan', a project that resulted in a human cDNA library that was generated from full length mRNAs. mRNAs were isolated from 61 tissues, 21 primary cell cultures and 16 cell lines [61]. The NITE cDNA library contains both protein coding and non-coding sequences, cloned predominantly in pME18SFL3 shuttle vector. A unique GenBank [87] accession number (AK number) identifies each cDNA clone. Moreover, as the clones were provided in a 96-well format, their respective 96-well coordinates were also used to identify them. For example, if the clone AK074989 were located in the well F10 on plate number 40, it would be identified as 40F10. The NITE library database provided by the supplier identified clones with both their AK numbers and the 96-well coordinates. Over the years and after various rounds of screening in Stefan Grimm's lab, cDNA clones were redistributed into new 96-well plates. For example, the cDNA clone AK074989 was initially screened out for its ability to induce apoptosis. This clone was subsequently moved to a new 96-well plate along with other apoptosis inducers and catalogued separately.

The list of the 22 anticancer genes screened out by Qize Ding is provided in table 3.1. Each gene was identified through its AK number and the relative plate coordinates. Table 3.1 also provides information regarding the transformation of CV-1 for each gene. As the cDNA clones were redistributed over time, the correct AK numbers of these clones needed to be confirmed through

sequencing. Before the sequencing results were obtained, I started by designating these genes as ACG1 to ACG22, ACG for anticancer genes, followed by the chronological number in which these were identified.

Table 3.1: **List of anticancer genes used in this study.** The genes were isolated through the systematic functional genetic screen by Bevan Lin and Qize Ding

96-well plate Coordinates	GenBank Accession number	Annotated ACG	CV-1 cells transformation method
1B01	AK074989	1	cMYC overexpression, HRAS overexpression, E1A + E1B55 overexpression, RB knockdown.
4G04	AK026893	2	cMYC overexpression, HRAS overexpression, E1A + E1B55 overexpression, WT1 knockdown.
3C09	AK091900	3	cMYC overexpression, HRAS overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown, microRNA-17-92 overexpression.
4G03	AK024890	4	cMYC overexpression, HRAS overexpression, E1A + E1B55 overexpression, microRNA-17-92 overexpression.
3B03	AK127225	5	cMYC overexpression, HRAS overexpression, E1B55 overexpression, E1A + E1B55 overexpression, WT1 knockdown, microRNA-17-92 overexpression.
1A03	AK074677	6	cMYC overexpression, HRAS overexpression, E1B55 overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown.
2B08	AK055260	7	cMYC overexpression, E1A + E1B55 overexpression, RB knockdown, microRNA-17-92 stable.
3B06	AK127945	8	cMYC overexpression, E1A + E1B55 overexpression, WT1 knockdown.
2A05	AK056206	9	cMYC overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown.
1B08	AK075147	10	cMYC overexpression, E1A + E1B55 overexpression.
2H03	AK124577	11	cMYC overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown, microRNA-17-92 overexpression.
4E10	AK024986	12	E1A overexpression
4E06	AK026890	13	cMYC overexpression, E1A + E1B55 overexpression, RB knockdown.

3D05	AK092480	14	cMYC overexpression, HRAS overexpression, E1B55 overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown, microRNA-17-92 overexpression.
4E11	AK024844	15	Induced cell death in HEK293T but not in wild type CV1. Does not cause cell death in any of transformed cells either.
3A03	AK125051	16	Causes cell death in HEK293T but not in wild type CV1. Does not cause cell death in any of transformed cells either.
4F06	AK024692	17	cMYC overexpression, HRAS overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown, microRNA-17-92 overexpression.
4F04	AK024913	18	cMYC overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown.
2A07	AK056273	19	E1B55 overexpression, E1A + E1B55 overexpression, RB knockdown.
4E08	AK024691	20	Induced cell death in HEK293T but not in wild type CV-1. Does not cause cell death in any of transformed cells either.
4F12	AK024955	21	cMYC overexpression, HRAS overexpression, E1A + E1B55 overexpression, WT1 knockdown, RB knockdown.
4D09	AK024690	22	E1A overexpression.

Abbreviations E1A (early viral 1A), E1B55 (early viral B 55 kDa), RB (retinoblastoma-associated protein), WT1 (Wilms tumour 1).

3.2 Results

3.2.1 Sequence Verification

The plasmid DNA preps of the 22 screened anticancer genes were subjected to Sanger sequencing [88]. Short representative peak views for ACG1 to ACG22 are shown in figure 3.1. Further, I checked if the resulting sequences obtained matched with the respective sequences provided by the NITE library [61] (for the corresponding 96-well plate coordinates). Alignment was performed using The National Center for Biotechnology Information (NCBI) nucleotide blast two-sequence alignment tool, Basic Local Alignment Search Tool-N (BLASTN) v 2.2.32+ [67, 68]. The results of the alignments are enlisted in table 3.2 and the representative sequence matching on NCBI BLASTN 2.2.32+[68] is depicted in the supplementary figure S.1. The sequences of ACG1, ACG2, ACG3, ACG4, ACG5, ACG7,

ACG8, ACG9, ACG10 ACG11, ACG14, ACG16, ACG17, ACG18, ACG19 and ACG22 aligned successfully with their corresponding NITE library sequences.

On the other hand, the sequences for ACG6, ACG12, ACG13, ACG15, ACG20 and ACG21 did not match with the corresponding library sequences. The sequences for these genes were then subjected to the NCBI nucleotide BLAST using the NCBI database [68]. The results with the highest sequence alignments corresponded to the cDNA sequences submitted by NITE to the NCBI for the same FLJ (Full Length Japan-library supplier) cDNA library that was used for screening. These results indicated that there was cross-contamination within the library during various stages of screening. The correct NCBI database accession numbers for the corresponding sequences are enlisted in table 3.3.

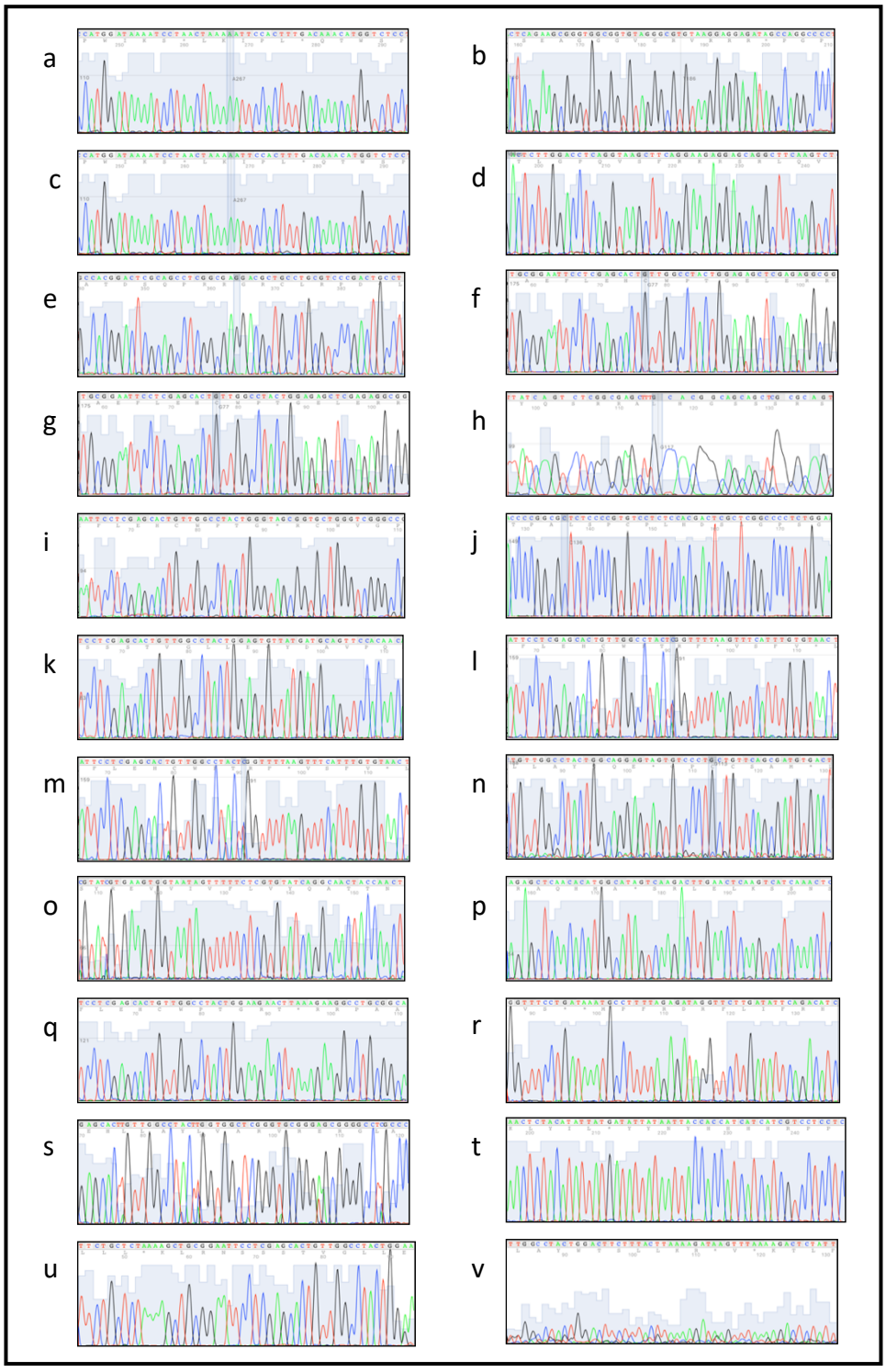


Figure 3.1: **Short representative peak views for the Sanger sequencing results.** (a) ACG1, (b) ACG2, (c) ACG3, (d) ACG4, (e) ACG5, (f) ACG6, (g) ACG7, (h) ACG8, (i) ACG9, (j) ACG10, (k) ACG11, (l) ACG12, (m) ACG13, (n) ACG14, (o) ACG15, (p) ACG16, (q) ACG17, (r) ACG18, (s) ACG19, (t)ACG20, (u) ACG21, (v) ACG22.

Table 3.2: **Results for NCBI nucleotide blast two-sequence alignment tool**, BLASTN 2.2.32+ for ACG1 to ACG22.

96-well plate Coordinates in library database	Annotated ACG	GenBank accession number	Did the sequence match with library sequence?
1B01	1	AK074989	YES
4G04	2	AK026893	YES
3C09	3	AK091900	YES
4G03	4	AK024890	YES
3B03	5	AK127225	YES
1A03	6	AK074677	NO
2B08	7	AK055260	YES
3B06	8	AK127945	YES
2A05	9	AK056206	YES
1B08	10	AK075147	YES
2H03	11	AK124577	YES
4E10	12	AK024986	NO
4E06	13	AK026890	NO
3D05	14	AK092480	YES
4E11	15	AK024844	NO
3A03	16	AK125051	YES
4F06	17	AK024692	YES
4F04	18	AK024913	YES
2A07	19	AK056273	YES
4E08	20	AK024691	NO
4F12	21	AK024955	NO
4D09	22	AK024690	YES

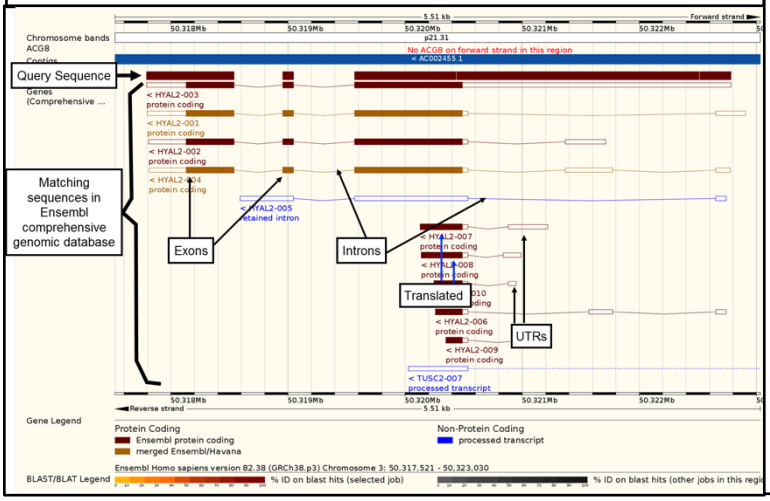
Table 3.3. **Correct sequence accession numbers** for the genes which did not align with their corresponding NITE library reference sequences.

96-well plate Coordinates in library database	**Annotated ACG**	**Correct GenBank accession number according to NCBI BLAST**
1A03	**6**	AK074632.1
4E10	**12**	AK024844.1
4E06	**13**	AK024927.1
4E11	**15**	AK024868.1
4E08	**20**	AK024973
4F12	**21**	AK024689.1

3.2.2 Gene Identities

Although the NCBI database and their alignment algorithm (BLASTn) are excellent at matching the query sequences against the previously known (compressive) database, when it comes to the gene annotation, NCBI BLASTn is hardly straightforward. The results of the sequence alignments can give a false impression about the annotations of the queried sequences. For example, if a query sequence overlaps only a small portion of a previously annotated gene, the alignment results would still provide a very high alignment score. To avoid such discrepancies, I subjected the NITE library reference cDNA sequences of all 22 genes to the BLAST-like Alignment Tool (BLAT) [69] using the Ensembl genomic interpretation system [70]. The results are shown in figures 3.2 through to 3.23.

Ensembl Genome Browser Legend



Box 3.1: A brief guide to analysing the results from Ensembl genomic interpretation system. Boxes represent exons and the connecting lines indicate the introns. Filled boxes represent coding sequences whereas empty boxes represent untranslated regions. Image adapted from: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

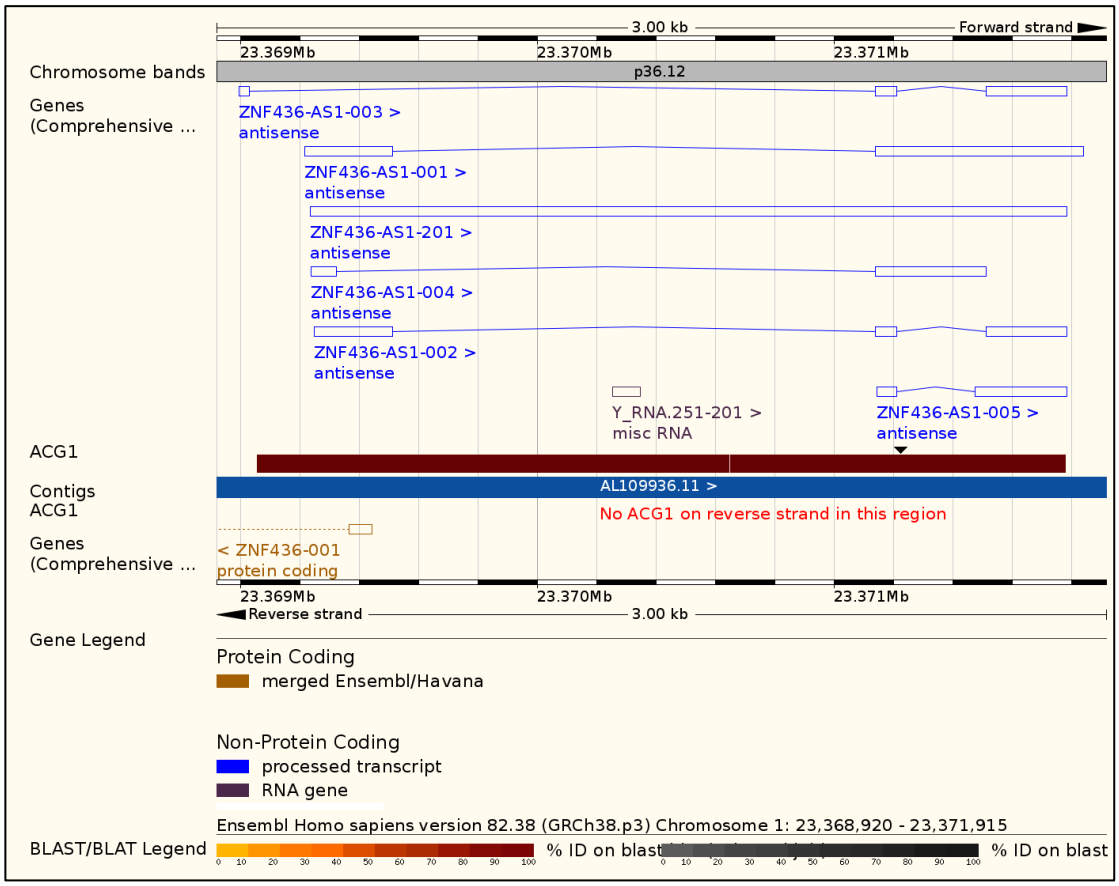


Figure 3.2: **Ensembl BLAT results for ACG1**. The query ACG1 cDNA sequence overlapped introns and exons of *ZNF436-AS1*. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

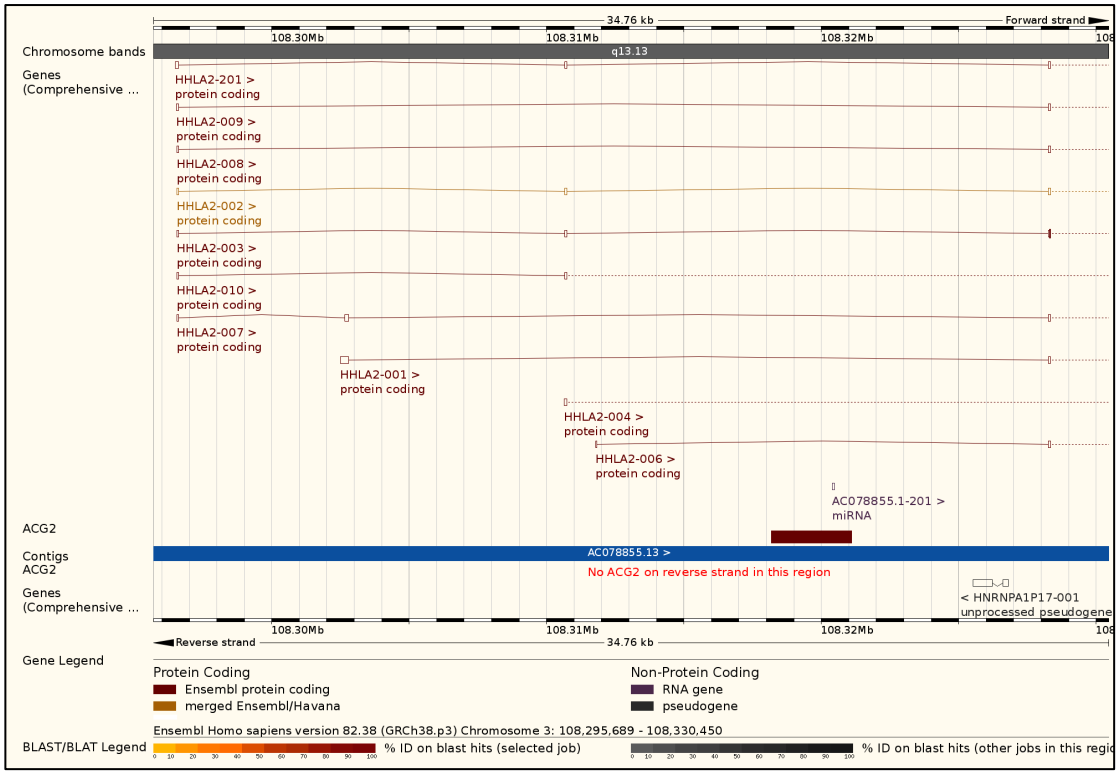


Figure 3.3: **Ensembl BLAT results for ACG2**. The query ACG2 cDNA sequence partially overlapped the *HHLA2* intron and AC0788551.1-201 miRNA sequence. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html..

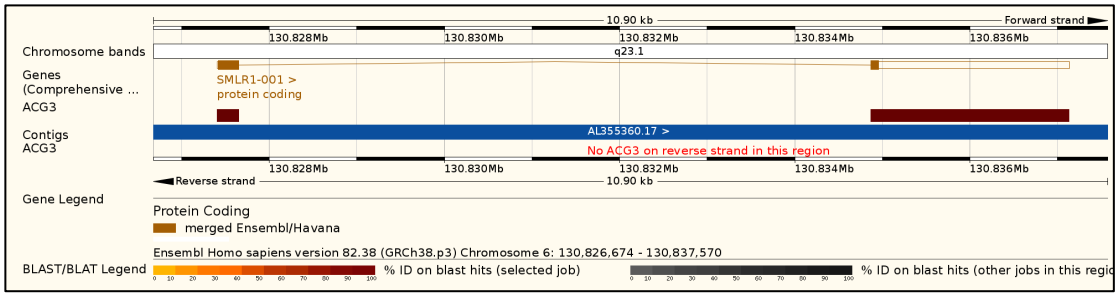


Figure 3.4: **Ensembl BLAT results for ACG3.** The query ACG3 cDNA sequence matched *SMLR1* splice variant 001. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

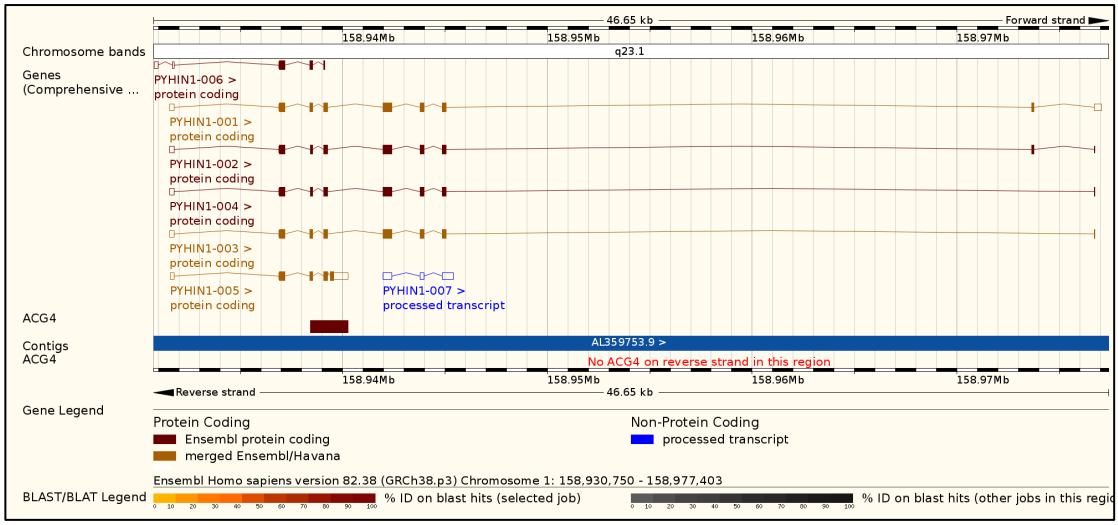


Figure 3.5: Ensembl BLAT results for ACG4. The query ACG4 cDNA sequence partially overlapped *PHYN1*. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html..

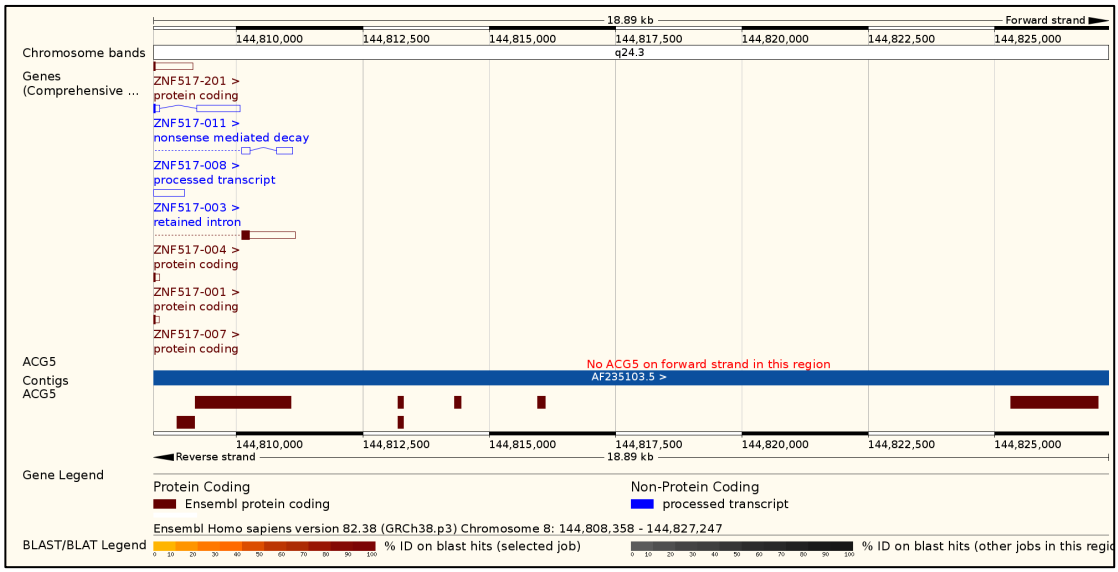


Figure 3.6: **Ensembl BLAT results for ACG5**. The query ACG5 cDNA sequence exhibited typical exon structure but did not overlap any annotated gene on Ensembl database. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

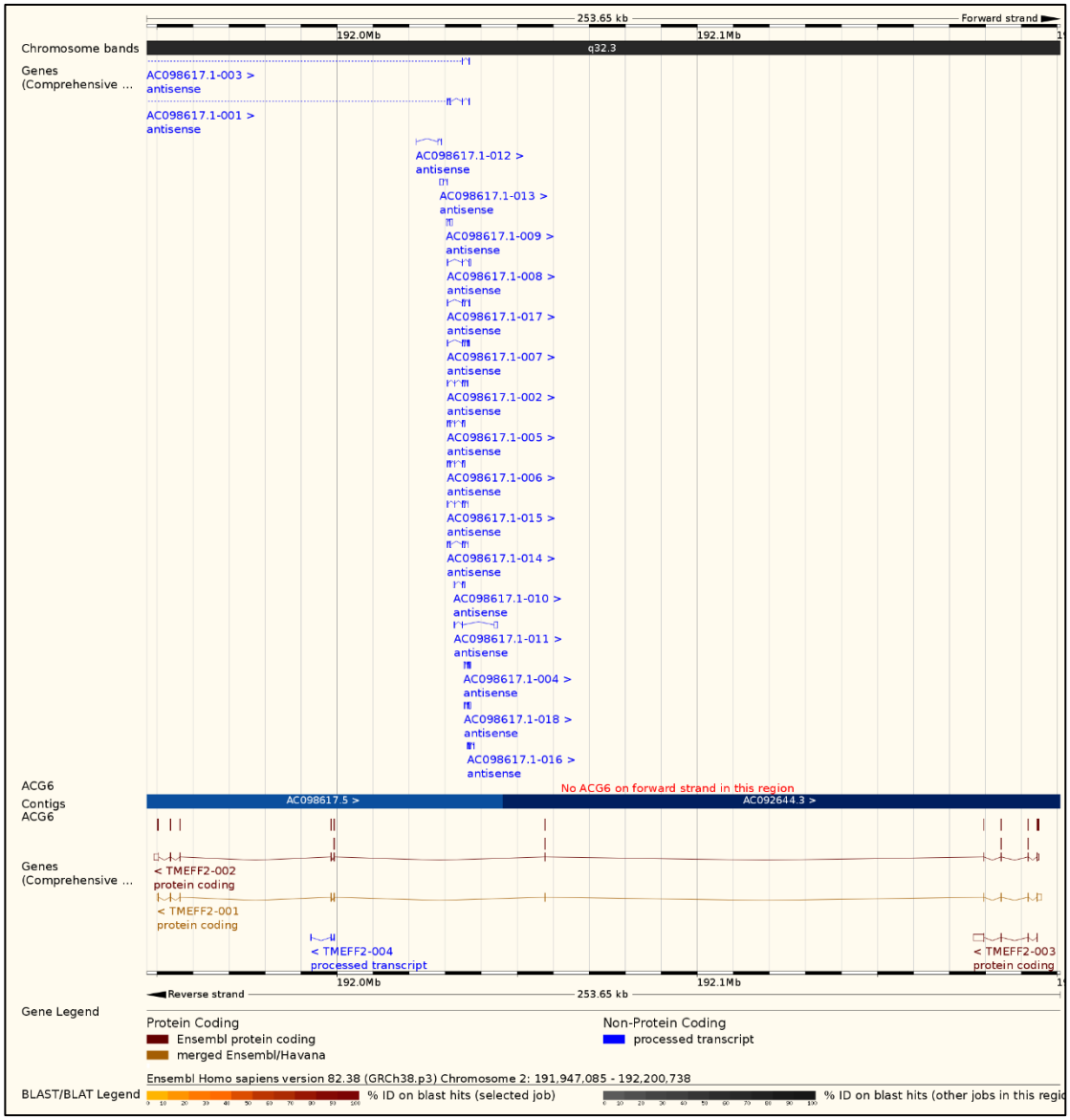


Figure 3.7: **Ensembl BLAT results for ACG6**. The query ACG6 cDNA sequence matched TMEFF2, the 5' sequence matched splice variant 001 while the 3' end matched the sequence of TMEFF2 splice variant 002. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

.

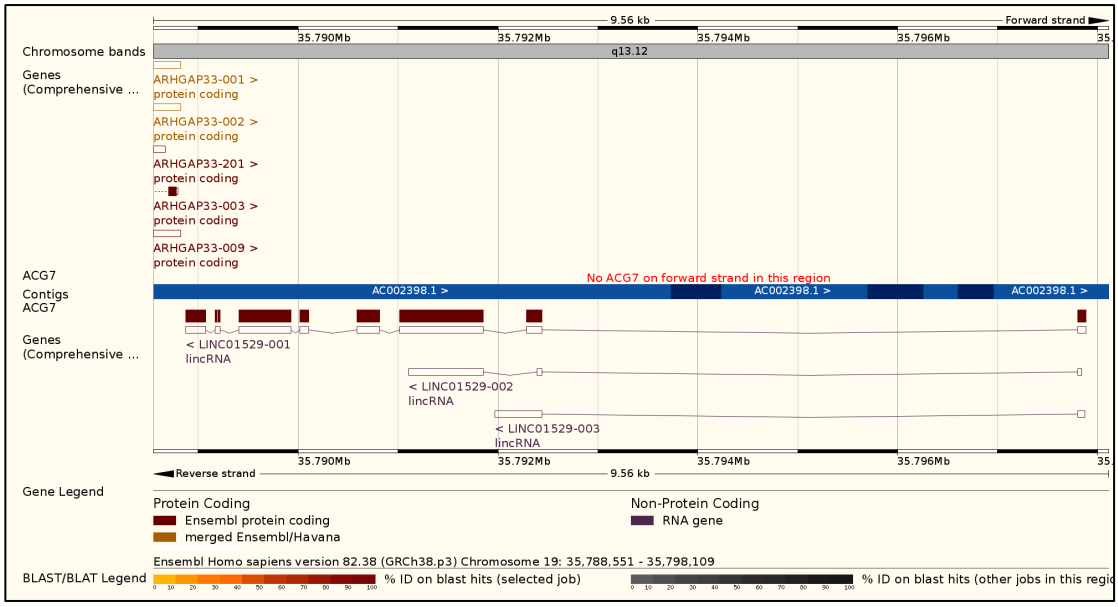


Figure 3.8: **Ensembl BLAT results for ACG7.** The query ACG7 cDNA sequence matched LINC01529lincRNA splice variant 001. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

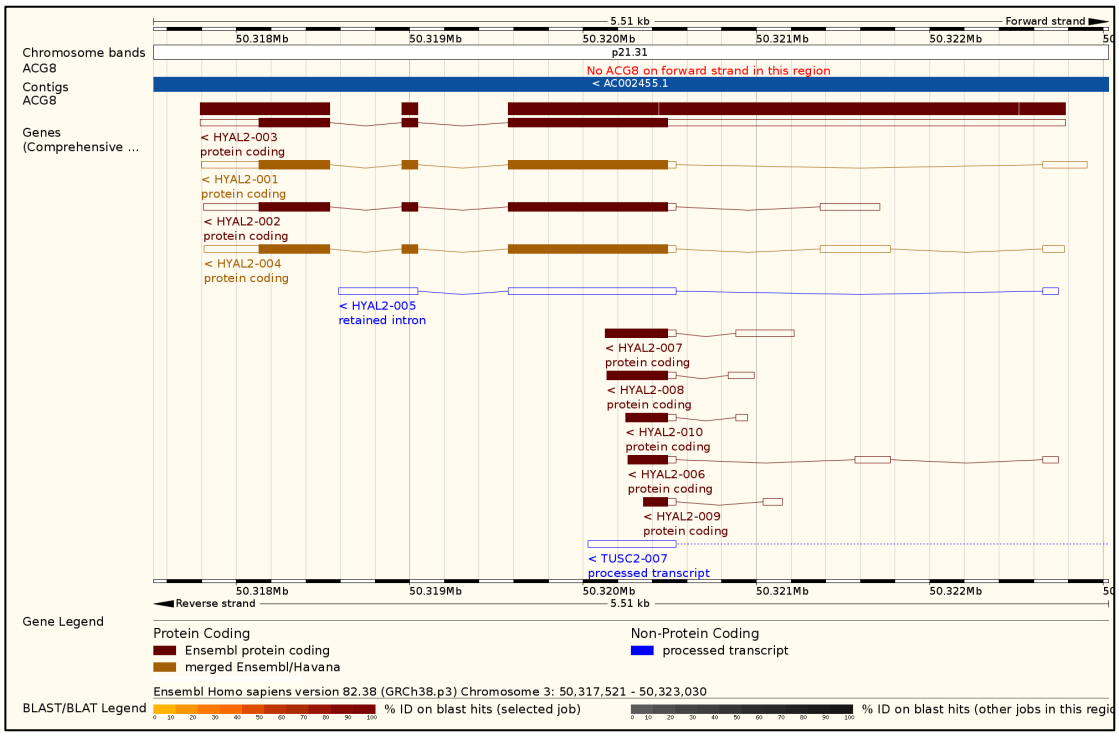


Figure 3.9: **Ensembl BLAT results for ACG8**. The query ACG8 cDNA sequence matched *HYAL2* splice variant 003. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

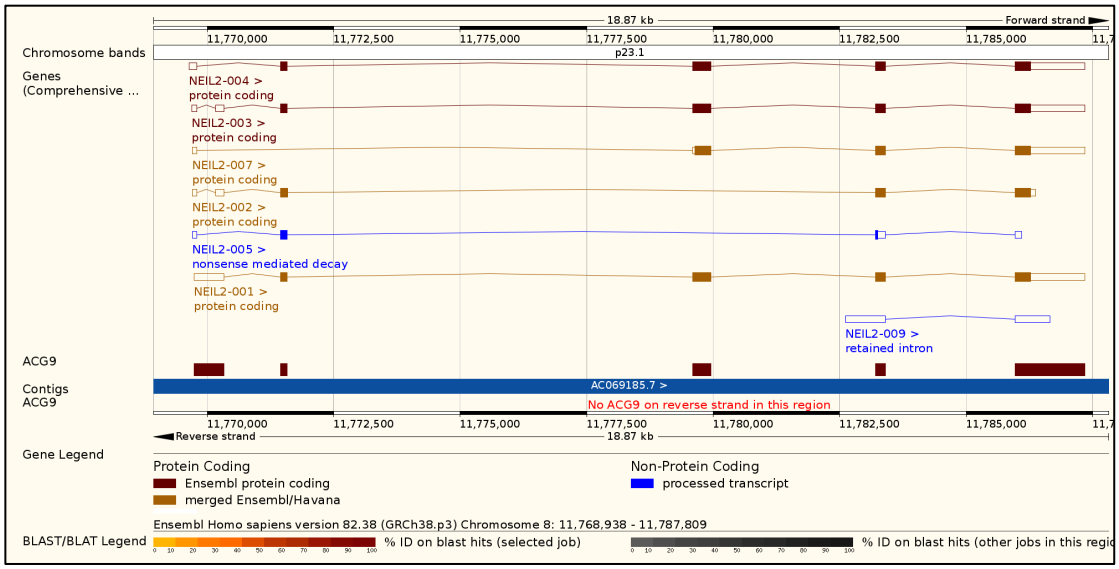


Figure 3.10: **Ensembl BLAT results for ACG9.** The query ACG9 cDNA sequence matched NEIL2 splice variant 001. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

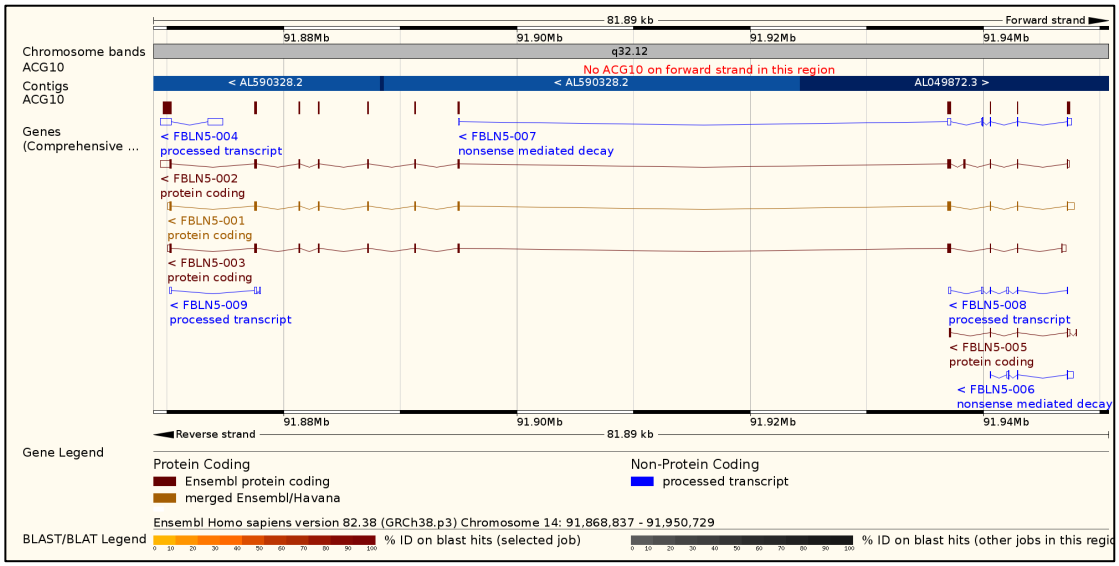


Figure 3.11: **Ensembl BLAT results for ACG10**. The query ACG10 cDNA sequence matched *FBLN5* splice variant 002. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

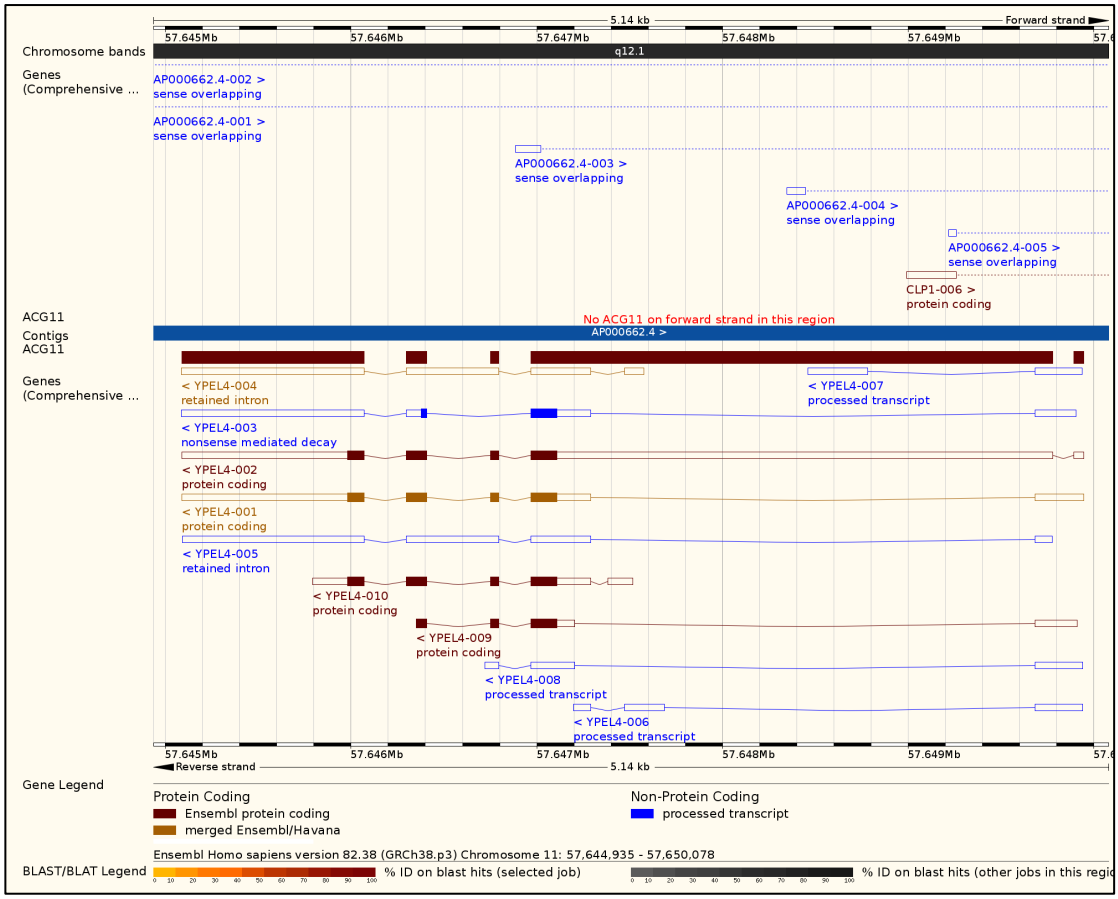


Figure 3.12: **Ensembl BLAT results for ACG11**. The query ACG11 cDNA sequence matched *YPEL4* splice variant 002. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

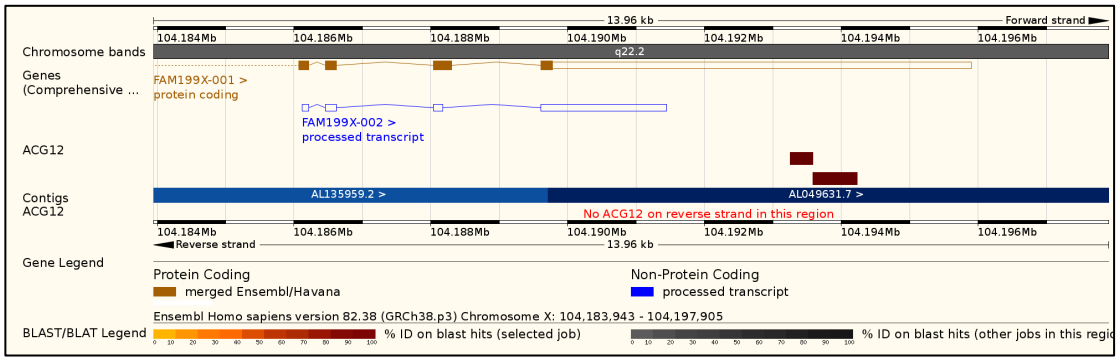


Figure 3.13: **Ensembl BLAT results for ACG12**. The query ACG12 cDNA sequence partially overlapped an exon of *FAM199X*. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

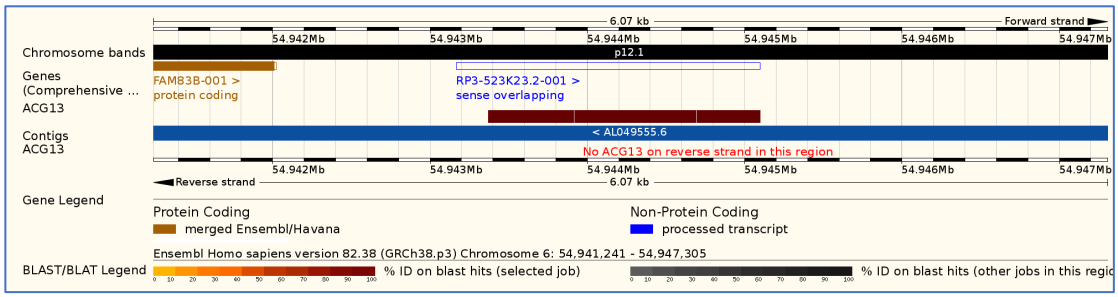


Figure 3.14: **Ensembl BLAT results for ACG13.** Query ACG13 cDNA sequence closely matched the RP3-523K23.2 processed transcript. A processed transcript is a non-coding transcript that does not contain an open reading frame. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

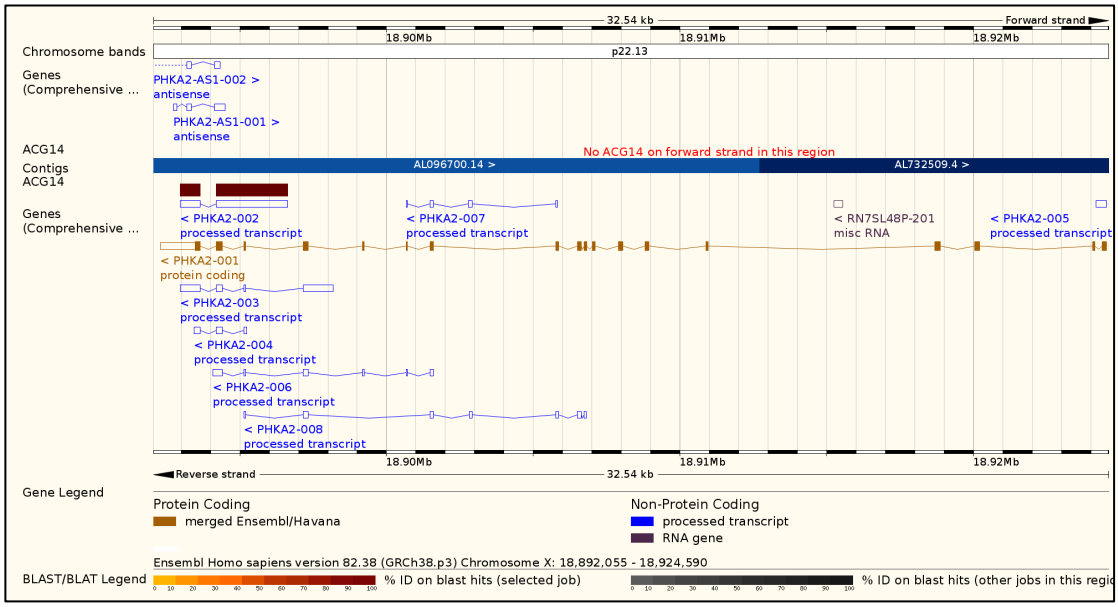


Figure 3.15: **Ensembl BLAT results for ACG14.** The Query ACG14 cDNA sequence matched the *PHKA2-002* processed transcript. A processed transcript is a non-coding transcript that does not contain an open reading frame. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

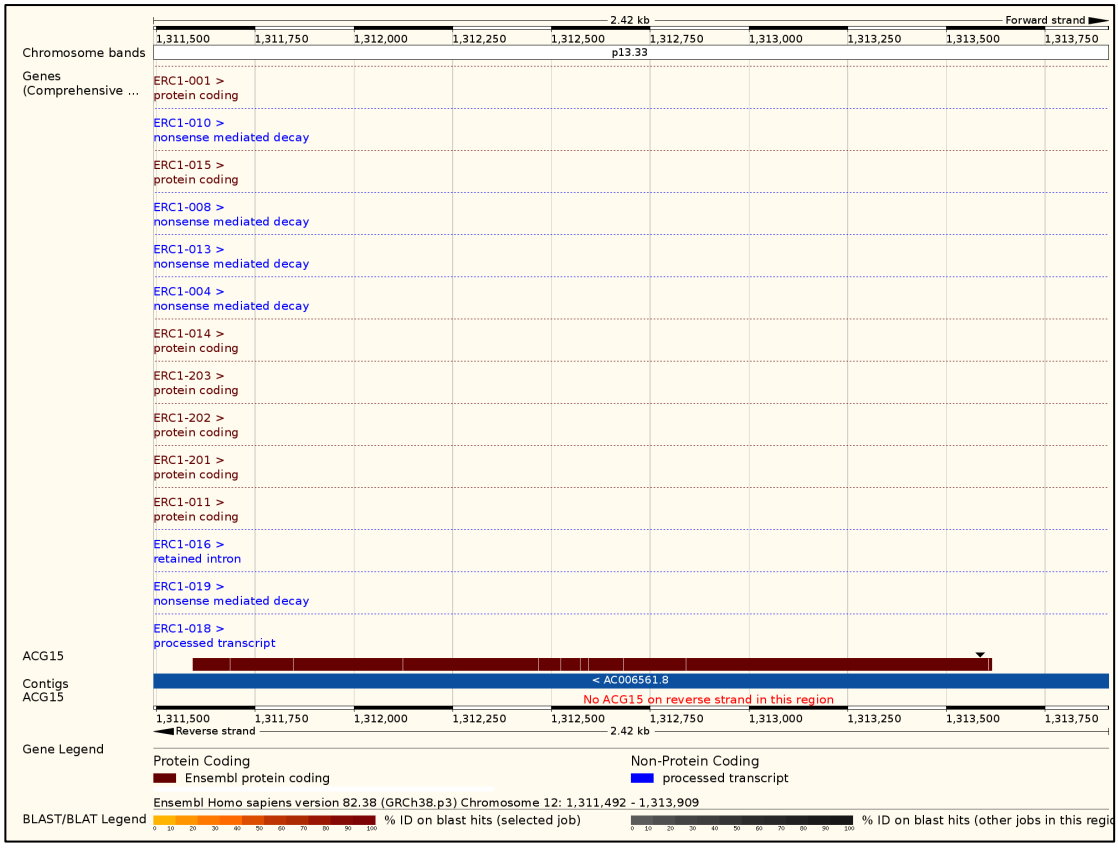


Figure 3.16: **Ensembl BLAT results for ACG15.** The query ACG15 cDNA sequence partially overlapped an intron of *ERC1*. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

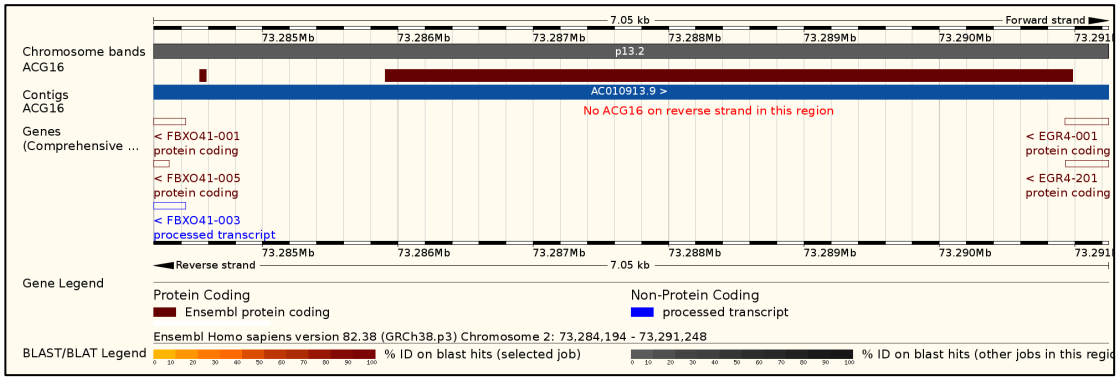


Figure 3.17: **Ensembl BLAT results for ACG16.** The query ACG16 cDNA sequence showed a typical intergenic pattern but did not overlap any gene on Ensembl comprehensive genomic database. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

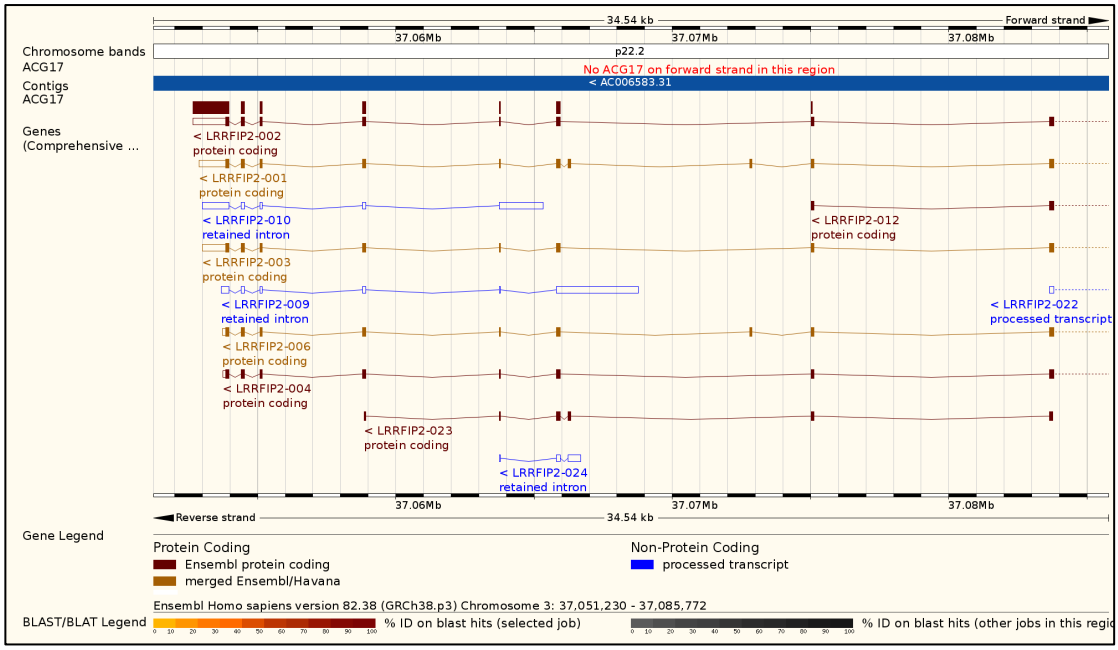


Figure 3.18: **Ensembl BLAT results for ACG17**. The query ACG17 cDNA sequence partially matched *LRRFIP2* splice variant 002. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

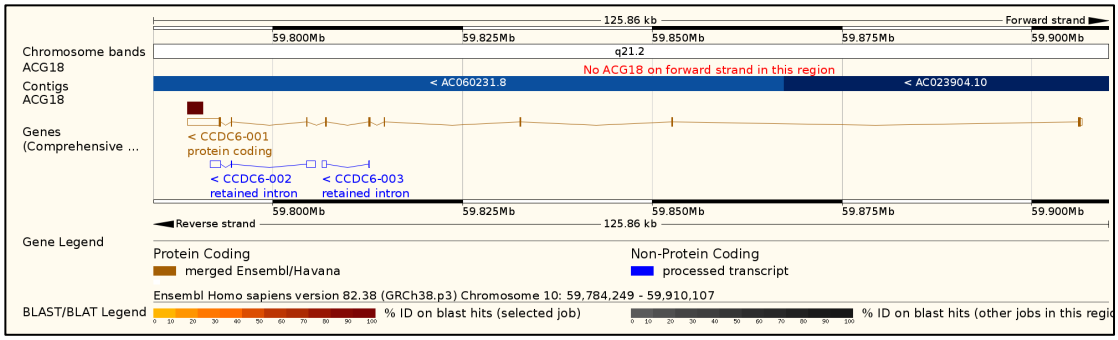


Figure 3.19: **Ensembl BLAT Results alignment for ACG18.** The query ACG18 cDNA sequence partially overlapped an exon of *CCDC6*. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

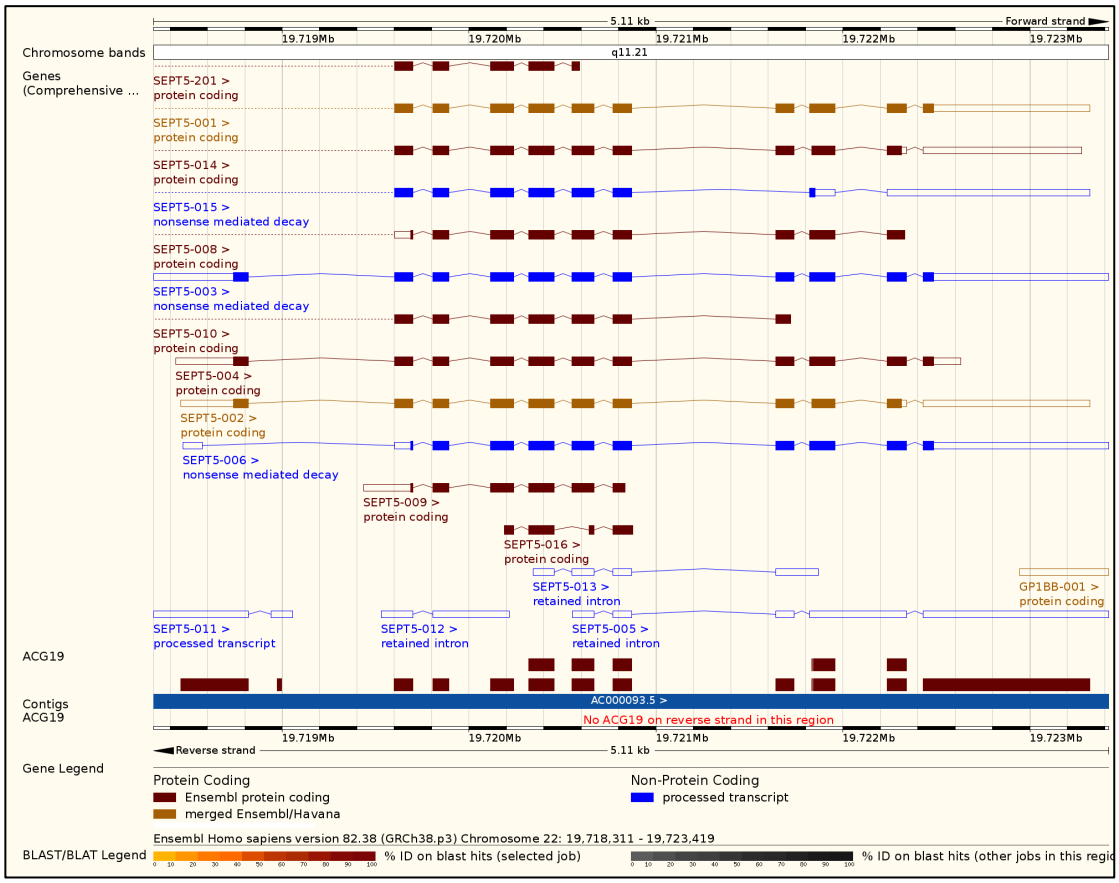


Figure 3.20: **Ensembl BLAT results for ACG19.** The query ACG19 cDNA sequence closely matched *SEPT5* splice variant 002. ACG19 appeared to contain an additional exon. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

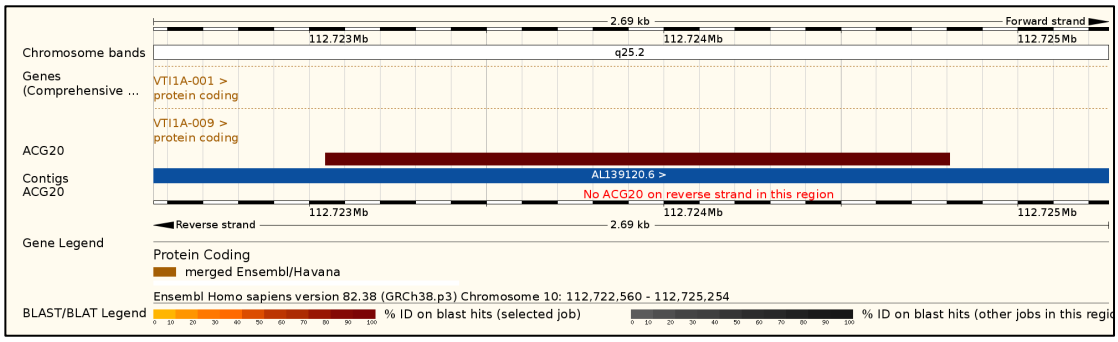


Figure 3.21: **Ensembl BLAT results for ACG20**. The query ACG20 cDNA sequence partially overlapped an intron of *VTI1A.* Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

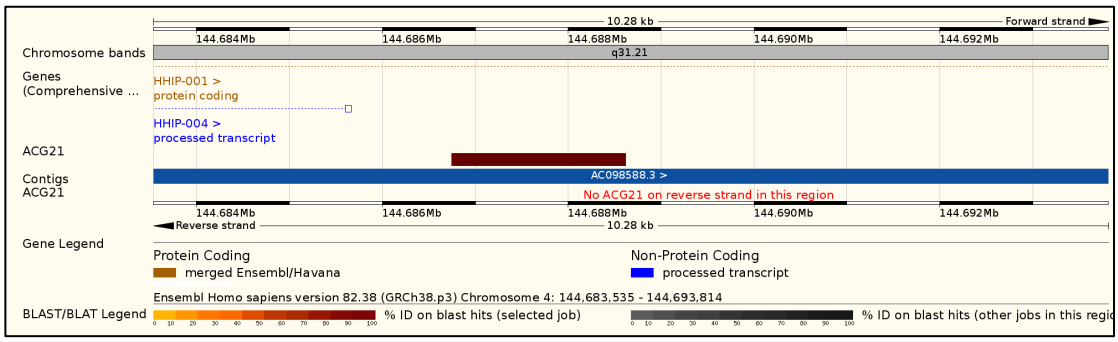


Figure 3.22: **Ensembl BLAT results for ACG21**. The query ACG21 cDNA sequence partially overlapped an intronic region of *HHIP*. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

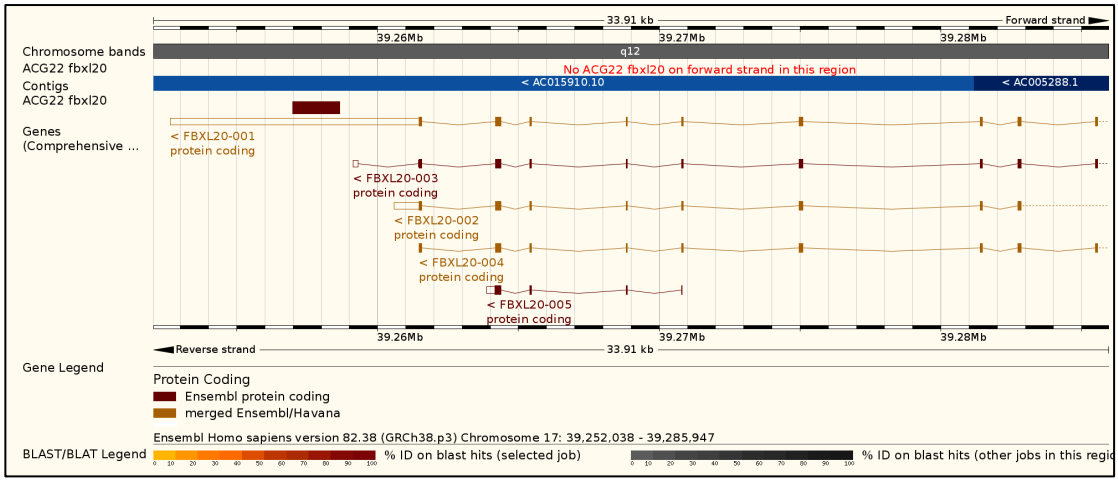


Figure 3.23: **Ensembl BLAT results for ACG22**. The query ACG22 cDNA sequence partially overlapped an exon of *FBXL20*. Source: Ensembl genome browser 77, http://oct2014.archive.ensembl.org/index.html.

Summarizing these results, ACG1 cDNA sequence overlapped all the exons and introns ZNF436 antisense RNA, *ZNF436-AS1* indicating that the cDNA generated for this sequence was from a full-length RNA. The cDNA sequences for ACG3 matched only the exons of *SMLR1*, which implied that the mRNA used to generate the cDNA was not full length but spliced (processed). Similarly, the cDNA sequences for ACG6, ACG7, ACG8, ACG9, ACG10, ACG11, ACG14 and ACG19 overlapped the exons of *TMEFF2, LINC01529, HYLA2, NEIL2, FBLN5, YPEL4, PHKA2* and *LRRFIP2* respectively. Looking closely, two of these transcripts were non-coding genes: in the case of LINK01529, it was a long intergenic non-coding RNA while for ACG14, the cDNA sequence matched the processed transcript of *PHKA2*. In the case of ACG6, the 5' query cDNA sequence matched splice variant 001 while the 3' end matched the sequence of splice variant 002 of *TMEFF2.* For ACG17, the query cDNA sequence overlapped with only the last 7 exons of *LRRFIP* splice variant 002, which could be a novel splice variant for the gene. ACG19 cDNA sequence closely matched the exons of *SEPT5* splice variant 002 but ACG19 appeared to have an additional exon which was not previously reported in any splice variants of *SEPT5.* Just as ACG17 appeared to be a novel splice variant of *LRRFIP2*, so does ACG19 seem to be a novel splice variant of *SEPT5*.

Nine out of the 22 cDNA sequences did not overlap the intergenic structure for any previously annotated genes on the Ensembl genomic database. Two among these 9 cDNA sequences, ACG5 and ACG16, appeared to be having intergenic sequences which overlapped specific chromosomal regions

(figures 3.6 and 3.17), thus exhibiting a peculiar structure which appears typical of exons. Preliminarily, it seems as if these two novel genes are yet to be characterized (at least by the Ensembl project). ACG2, ACG4, ACG12, ACG15, ACG20, ACG21 and ACG22 only overlapped small regions of the exons or the introns of some previously identified genes. One could speculate several possibilities regarding these transcripts ranging from novel coding splice variants, novel non-coding transcripts to the transcribed regulatory elements. The gene identities deciphered for the 22 anticancer genes are listed in table 3.4. For the purpose of representation in this thesis, I will keep referring to the genes which did not have a previously defined ID as ACGs. Although it appears that novel transcripts and splice variants have been identified, gene characterization is beyond the scope of my research aims.

Table 3.4: **Identities of anticancer genes.**

96-well plate Coordinates	Annotated ACG	GenBank accession number	Gene Symbol	Gene ID	Protein Coding?
1B01	1	AK074989	**ZNF436-AS1**	ZNF436 antisense RNA	No
4G04	2	AK026893	nd[1]	Novel-ACG2	nd
3C09	3	AK091900	**SMLR1**	Small leucine-rich protein 1	Yes
4G03	4	AK024890	nd	Novel-ACG4	nd
3B03	5	AK127225	nd	Novel-ACG5	nd
1A03	6	AK074632.1	**TMEFF2**	Transmembrane protein with EGF-like and two follistatin-like domains 2	Yes
2B08	7	AK055260	**LINC01529**	Long intergenic non-protein coding RNA 1529	No
3B06	8	AK127945	**HYAL2**	Hyaluronoglucosaminidase 2	Yes
2A05	9	AK056206	**NEIL2**	Nei-like DNA glycosylase 2	Yes
1B08	10	AK075147	**FBLN5**	Fibulin 5	Yes

2H03	11	AK124577	**YPEL4**	Yippee-like 4	Yes
4E10	12	AK124577	nd	Novel-ACG12	nd
4E06	13	AK024927.1	nd	Novel-ACG13	nd
3D05	14	AK092480	**PHKA2**	Non-coding, processed transcript	No
4E11	15	AK024868.1	nd	Novel-ACG15	nd
3A03	16	AK125051	nd	Novel-ACG16	nd
4F06	17	AK024692	**LRRFIP2**	Leucine rich repeat (in FLII) interacting protein 2.- Partial overlap.	nd
4F04	18	AK024913	nd	Novel-ACG18	nd
2A07	19	AK056273	**SEPT5**	Septin 5 – novel splice variant	nd
4E08	20	AK024973	nd	Novel-ACG20	nd
4F12	21	AK024955	nd	Novel-ACG21	nd
4D09	22	AK024690	nd	Novel-ACG21	nd

[1]nd, not determined

3.2.3 Prognostic Value

The concept of anticancer genes dictates that these genes do not necessarily have a prognostic value in cancer, i.e., an anticancer gene does not necessarily lose functionality during neoplastic transformation and progression. However, in order to expand the knowledge regarding the novel anticancer genes, I checked for their prognostic value in breast, ovarian, gastric and lung cancers. For this purpose, I employed the online database Kaplan-Meier Plotter http://kmplot.com/analysis/ [71]. For the detailed functionality of this webpage and related databases see Refs [72, 73, 74, 75]. I excluded SEPT5 and LRRFIP2 from these analyses because the information regarding their identities was not certain – they could be novel splice variants with properties different from those of the previously defined splice variants. Kaplan-Meier plots [89] for ZNF436-AS1, SMLR1, TMEFF2, HYAL2, NEIL2, FBLN5 and YEPL4 are provided in figures 3.24, 3.25, 3.26, 3.27, 3.28, 3.29 and 3.30, respectively. One striking commonality observed in almost all the genes tested was that their mRNA expression levels showed a positive correlation with the survival of breast cancer patients (SMLR1 being the only exception). Similar results were obtained for NEIL2 and FBLN5 in lung cancer patients. In addition, HYAL2 also exhibited a positive correlation between its mRNA levels and the ovarian cancer patient survival. Interestingly, ZNF436-AS1, FBLN5, YPEL4 and HYAL2 showed a negative correlation between their mRNA expression levels and the survival probability of gastric cancer patients while TMEFF2 exhibited a similar trend.

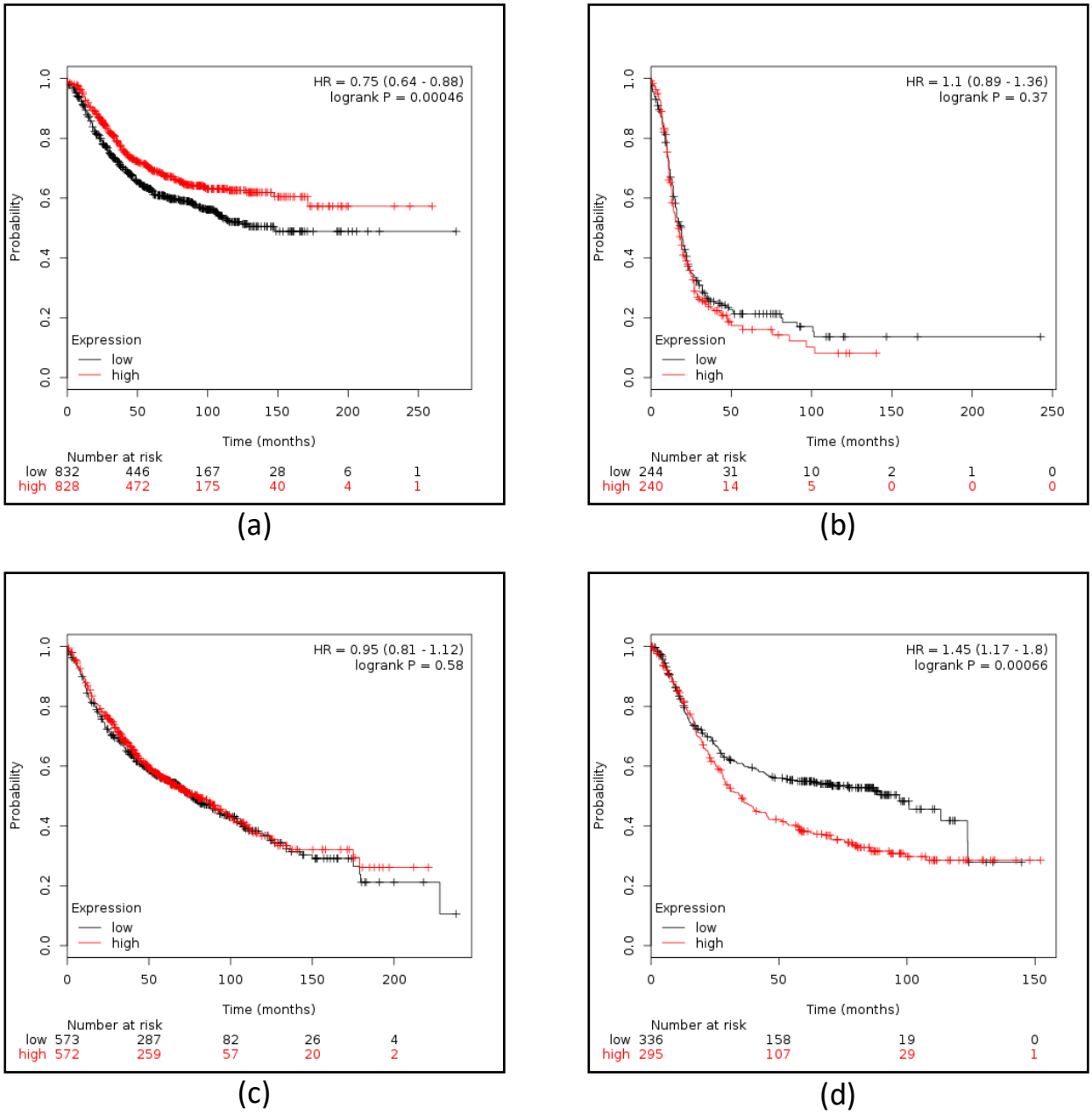


***Figure 3.24*: Kaplan-Meier plots showing the correlations between *ZNF436_AS1* expression and patient survival:** (a) breast cancer, (b) ovarian cancer, (c) lung cancer, (d) gastric cancer. Source: Kaplan-Meier Plotter (http://kmplot.com). Expression levels were measured through DNA microarrays, Affymetrix ® probe set ID: 229086_at.

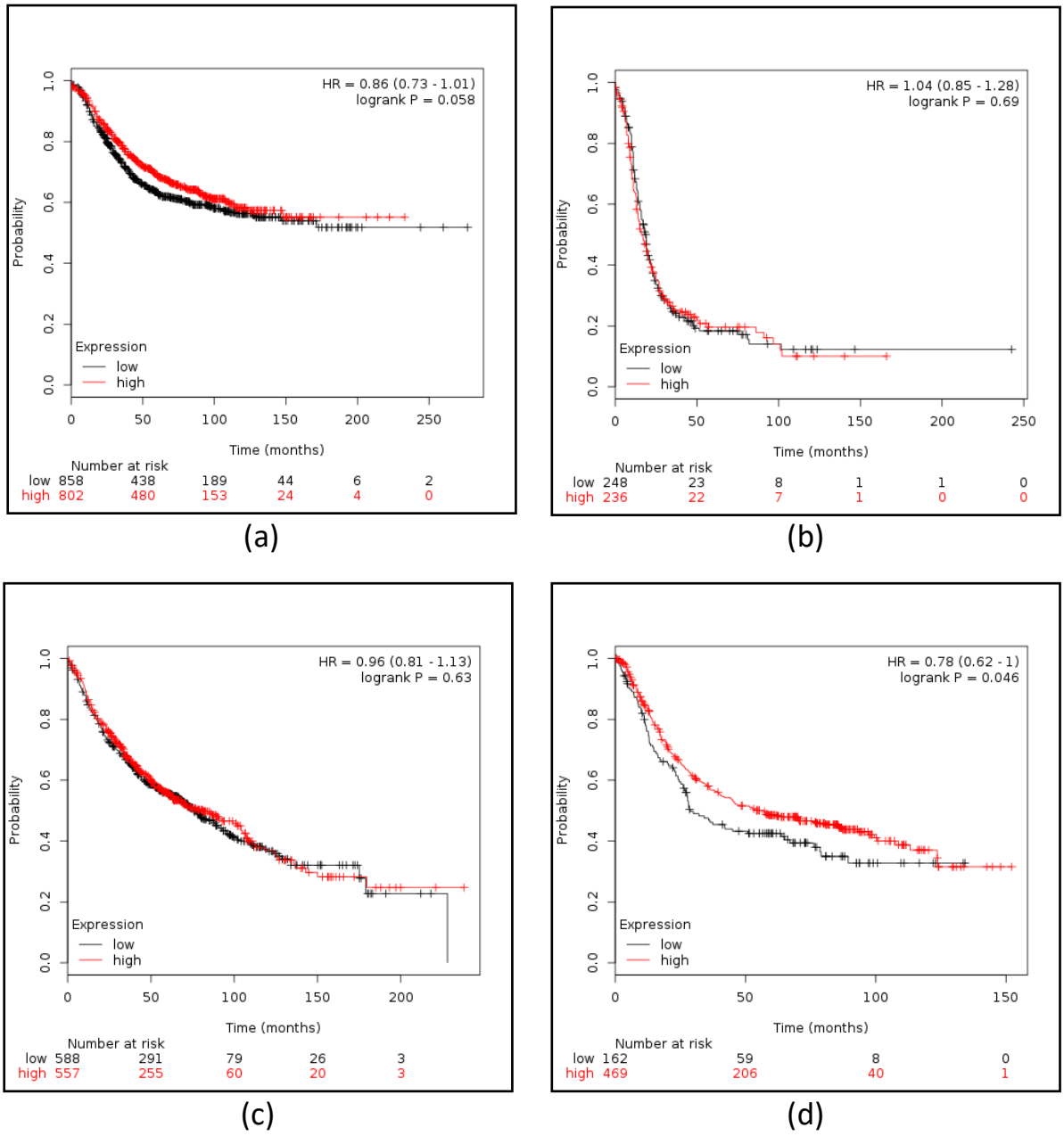


Figure 3.25: **Kaplan-Meier plots showing the correlations between *SMLR1* expression and patient survival:** (a) breast cancer, (b) ovarian cancer, (c) lung cancer, (d) gastric cancer. Source: Kaplan-Meier Plotter (http://kmplot.com). Expression levels were measured through DNA microarrays, Affymetrix ® probe set ID: 244151_at.

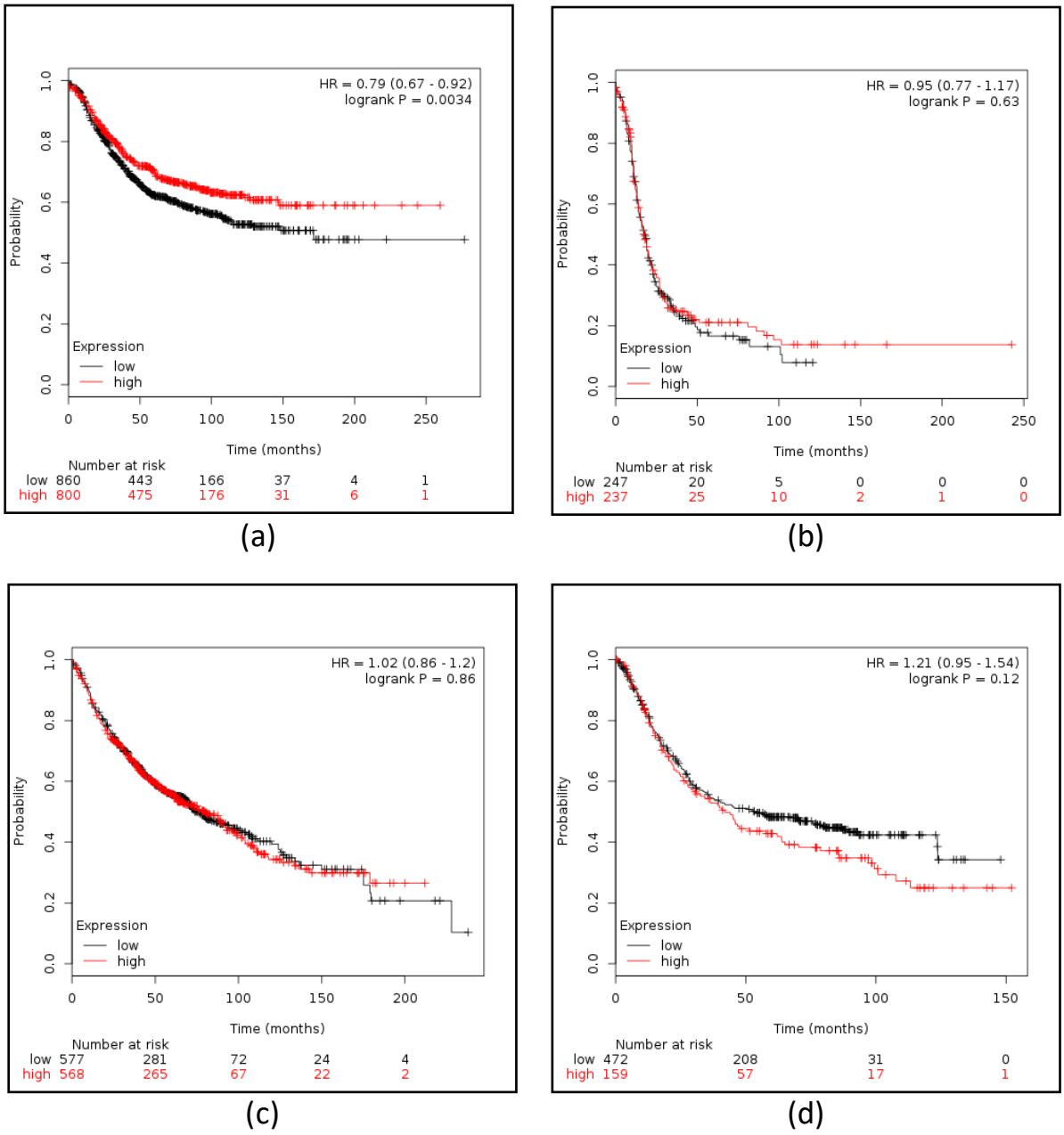


Figure 3.26: **Kaplan-Meier plots showing the correlations between *TMEFF2* expression and patient survival:** (a) breast cancer, (b) ovarian cancer, (c) lung cancer, (d) gastric cancer. Source: Kaplan-Meier Plotter (http://kmplot.com). Expression levels were measured through DNA microarrays, Affymetrix ® probe set ID: 233910_at.

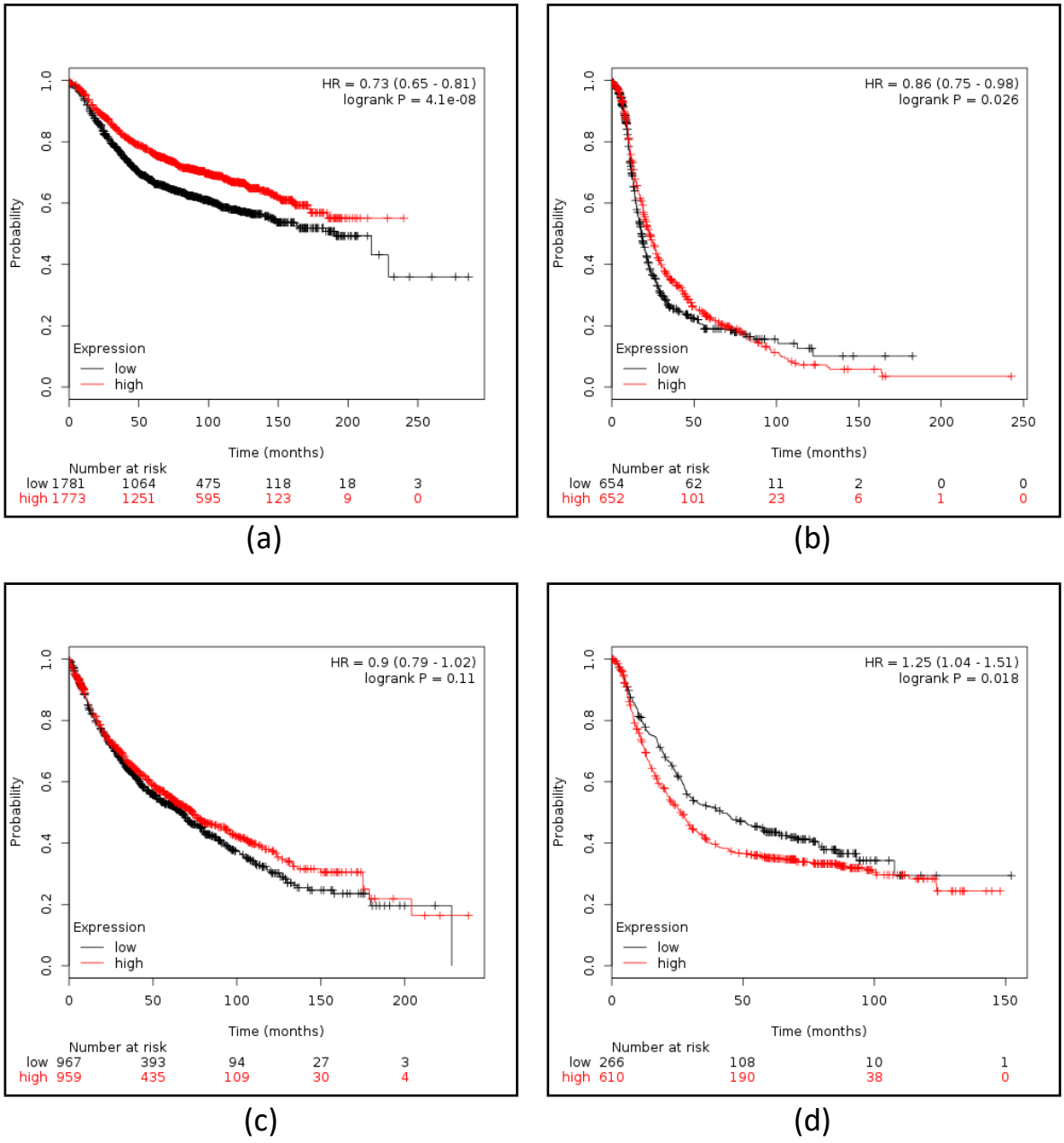


Figure 3.27: **Kaplan-Meier plots showing the correlations between *HYAL2* expression and patient survival:** (a) breast cancer, (b) ovarian cancer, (c) lung cancer, (d) gastric cancer. Source: Kaplan-Meier Plotter (http://kmplot.com). Expression levels were measured through DNA microarrays, Affymetrix ® probe set ID: 206855_s_at.

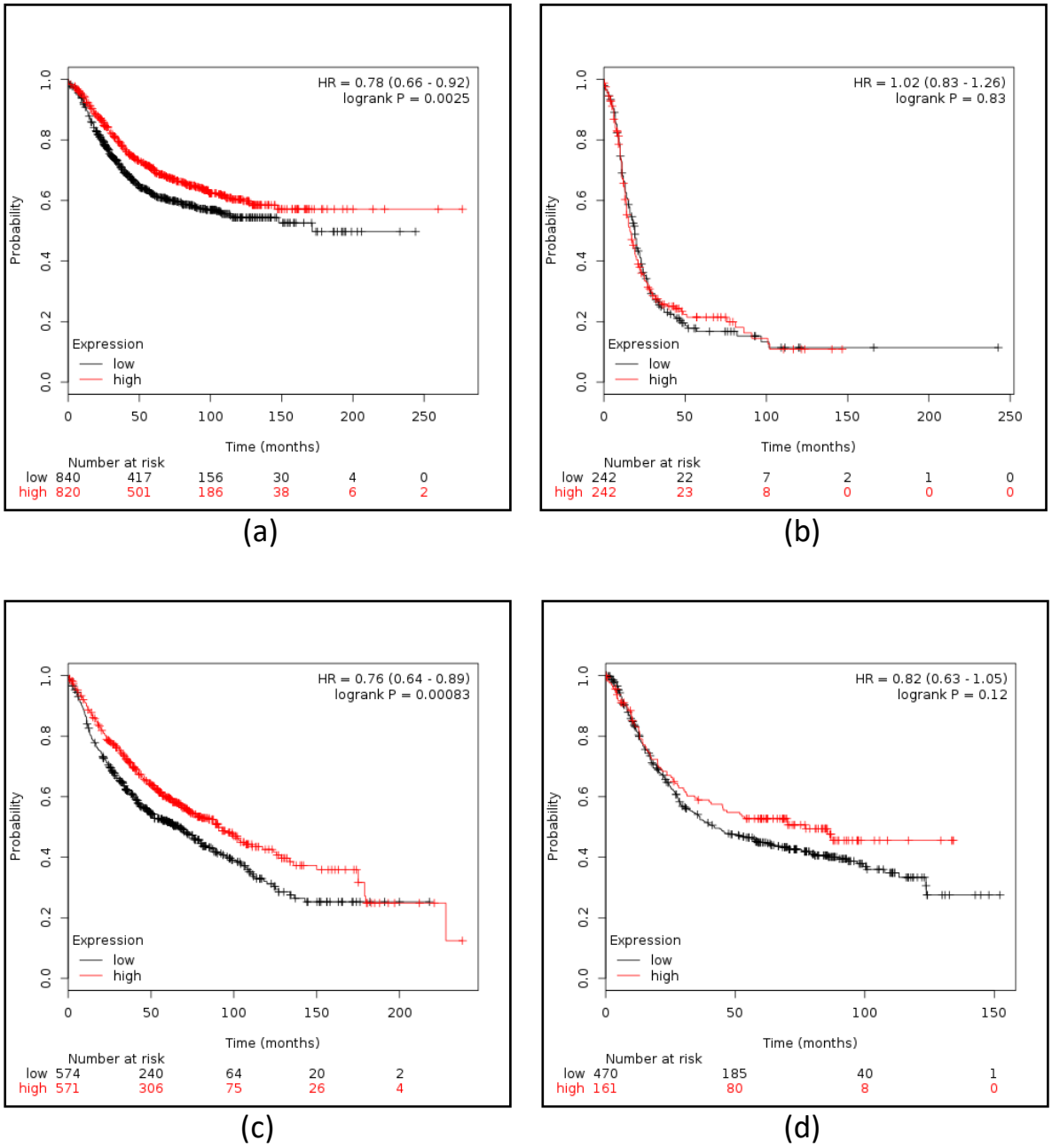


Figure 3.28: **Kaplan-Meier plots showing the correlations between *NEIL2* expression and patient survival:** (a) breast cancer, (b) ovarian cancer, (c) lung cancer, (d) gastric cancer. Source: Kaplan-Meier Plotter (http://kmplot.com). Expression levels were measured through DNA microarrays, Affymetrix ® probe set ID: 226585_at.

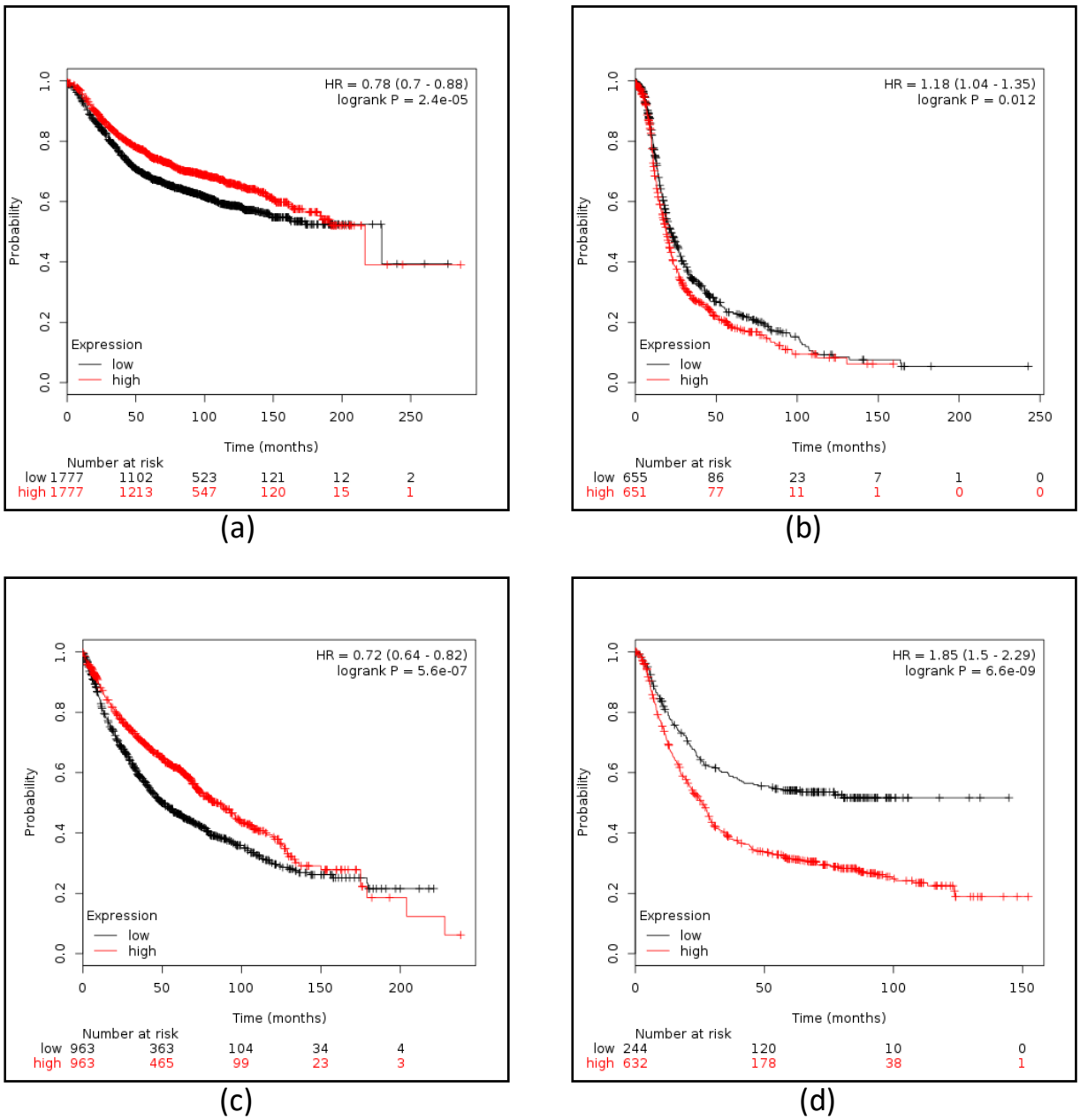


Figure 3.29: **Kaplan-Meier plots showing the correlations between *FBLN5* expression and patient survival:** (a) breast cancer, (b) ovarian cancer, (c) lung cancer, (d) gastric cancer. Source: Kaplan-Meier Plotter (http://kmplot.com). Expression levels were measured through DNA microarrays, Affymetrix ® probe set ID: 203088_at.

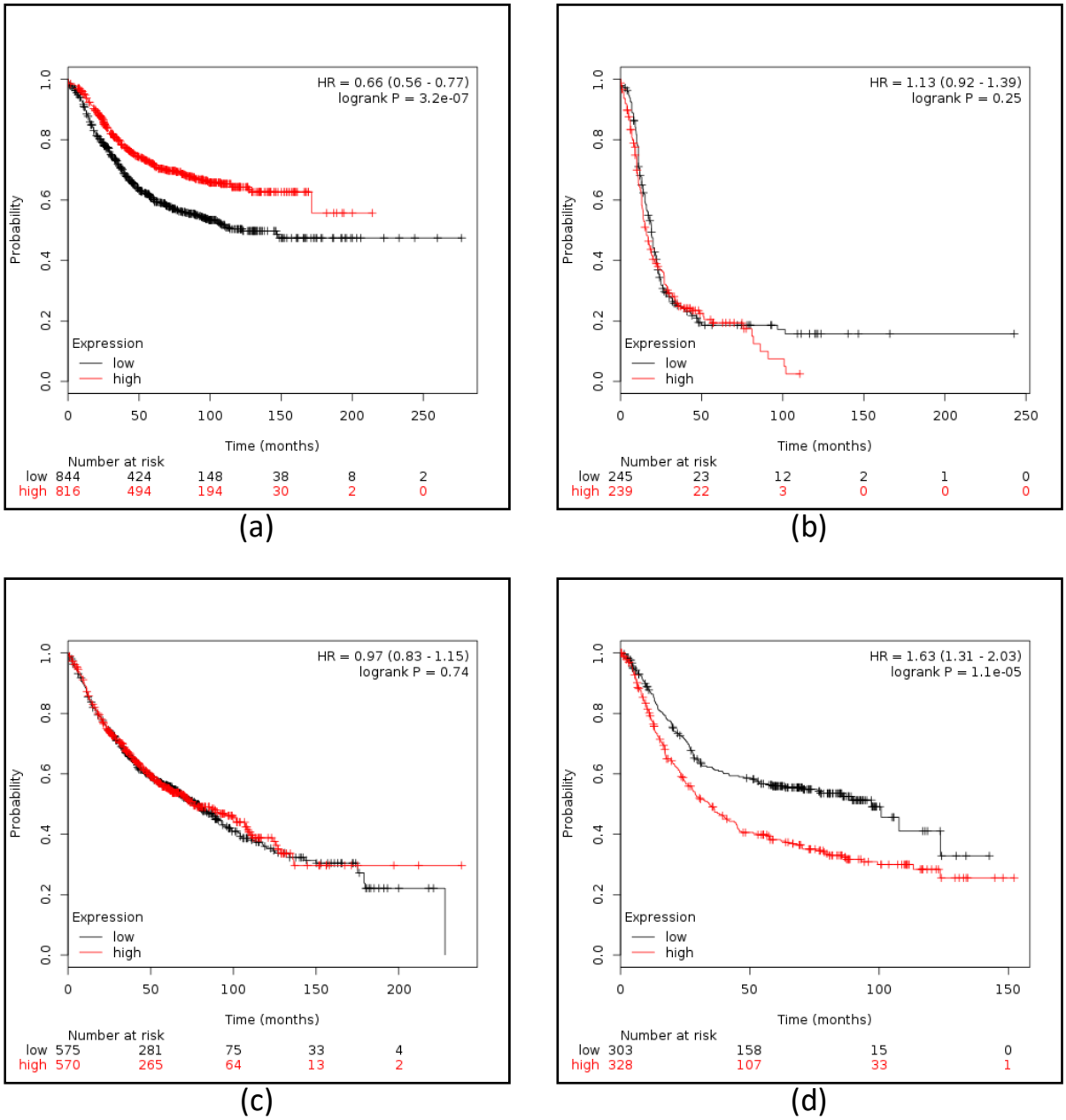


Figure 3.30: **Kaplan-Meier plots showing the correlations between *YPEL4* expression and patient survival:** (a) breast cancer, (b) ovarian cancer, (c) lung cancer, (d) gastric cancer. Source: Kaplan-Meier Plotter (http://kmplot.com). Expression levels were measured through DNA microarrays, Affymetrix ® probe set ID: 235916_at.

3.2.4 Anticancer genes in MCF-7 cells

As the 22 genes had gone through the same screening procedures, some commonalities in their regulation and functioning were not unexpected. The most vivid commonality observed in the Kaplan-Meier plots was the positive correlation between the mRNA levels of almost all of the tested anticancer genes and the survival probability of breast cancer patients. Based on this observation, I decided to probe further into the role of these novel anticancer genes in breast cancer. The oestrogen receptor (ER)-positive and epidermal growth factor receptor 2 (HER2)-negative breast cancer subtype has the highest incidence rate and thus constitutes the 'bulk' of clinical breast cancers [90]. MCF-7 cells are ER-positive and HER2-negative, a well-established breast cancer cell line derived from the ductal adenocarcinoma (reviewed elaborately in Ref [91]). In addition, MCF-7 cells present the advantages of relatively non-fastidious tissue culture requirements, rapid division rate and a relatively trouble-free transfection. I therefore decided to use MCF-7 cells as the breast cancer cell model for the further investigation of the novel anticancer genes.

I checked for the effects of the overexpression of anticancer genes on the cell viability and cell death in MCF-7 cells. For the latter, I employed $Dioc_6$ plus Propidium iodide (PI) combined staining which was detected using the flow cytometry [92, 93] – 3,3'-Dihexyloxacarbocyanine Iodide ($Dioc_6$) is a cell-permeant, green-fluorescent, lipophilic dye which selectively stains the mitochondria of live cells. Propidium iodide (PI) is a non-permeant red-fluorescent nuclear counterstain. Being non-permeant, PI only stains dead cells. The loss of $Dioc_6$ staining and/or gain of PI staining was used as a marker for cell death. Results are presented in figure 3.31. The experimental overexpression of ZNF436-AS1, ACG2, SMLR1, ACG5, TMEFF2, LINC01529, HYAL2, NEIL2, FBLN5, YEPL4 and SEPT5 induced substantial cell death in MCF-7 cells. On the other hand, ACG12, ACG13, ACG15, ACG17, ACG18, ACG20 and ACG22 caused little or no cell death.

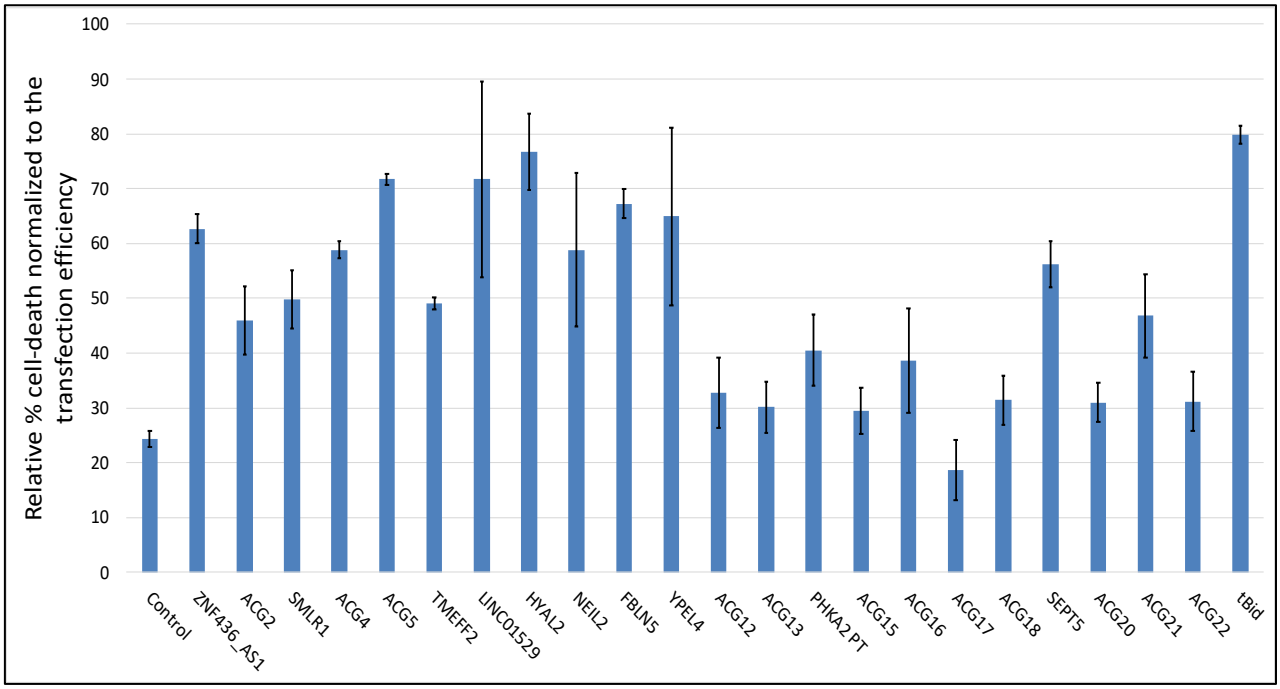
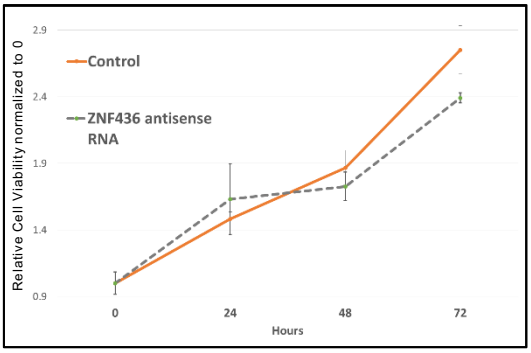


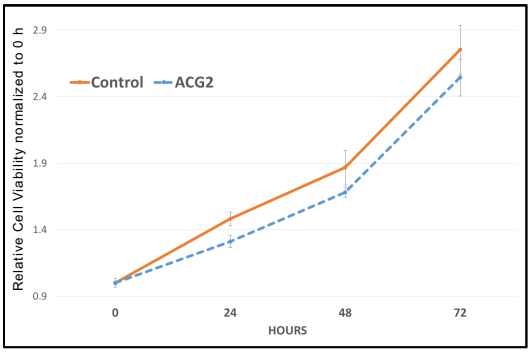
Figure 3.31: **Cell death in MCF-7 cells upon experimental overexpression of novel anticancer genes**. MCF-7 cells were transfected with the specified anticancer genes. 48 hours post-transfection, cells were harvested, stained with $Dioc_6$ + PI and subjected to flow cytometry. MCF-7 cells were transfected with GFP in parallel in order to monitor transfection efficiency. Results were adjusted to untransfected cells and normalized to the transfection efficiency. Luciferase was used as a negative 'control' and tBID as positive control. Bars represent average from two wells of a 24-well plate each with an independently transfected population of cells. Error-bars represent the +/- standard deviation. NB: this experiment was performed in collaboration with Qize Ding, data for this experiment is therefore shared between this and Qize Ding's PhD thesis.

In order to investigate the effects of the 22 anticancer genes on the short-term viability of MCF-7 cells, I performed the MTT colorimetric assay [94]. The results are shown in figure 3.32. The highest reductions in cell viability were observed in the cells transfected with ACG5, NEIL2, HYAL2 and YPEL4. A moderate reduction in cell viability over time was observed in cells transfected with ZNF436AS1, SMLR1, TMEFF2, FBLN5, ACG16, SEPT5 and ACG20, whereas the transfection with ACG2, ACG12, ACG13, ACG14, ACG15, LRRFIP2 processed transcript, ACG18 and ACG22 did not cause any reduction in the cell viability of MCF-7 cells over time. The results of the MTT colorimetric assay were by and large in line with those of cell death ($Dioc_6$ + PI) assay in terms of the trends. The only exception in this regard was ACG20, which did not cause cell death, but its overexpression resulted in a substantial reduction in cell viability

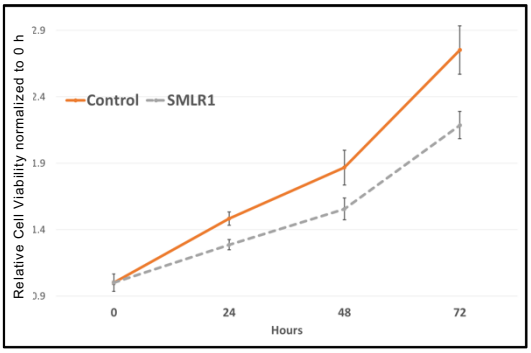
a


Control
ZNF436 antisense RNA
Relative Cell Viability normalized to 0
Hours

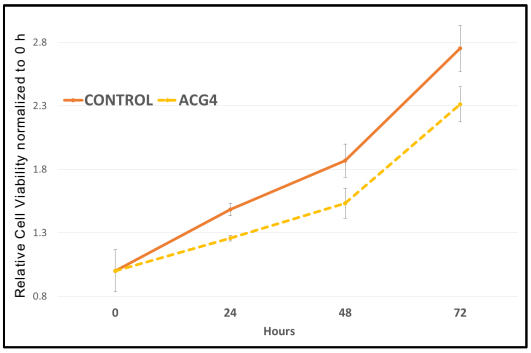
b


Control
ACG2
Relative Cell Viability normalized to 0 h
HOURS

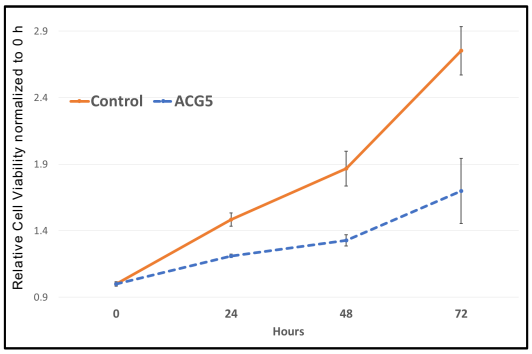
c


Control
SMLR1
Relative Cell Viability normalized to 0 h
Hours

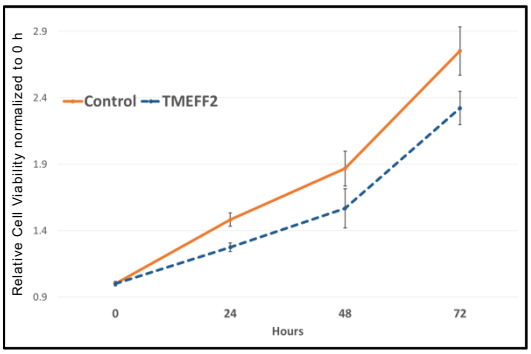
d


CONTROL
ACG4
Relative Cell Viability normalized to 0 h
Hours

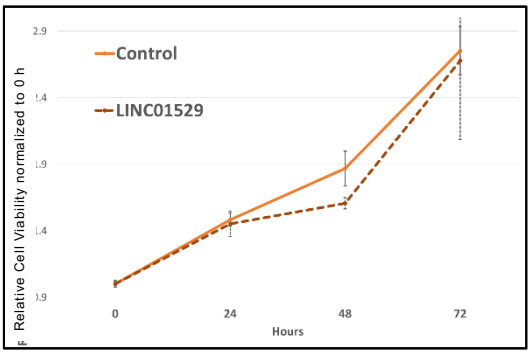
e


Control
ACG5
Relative Cell Viability normalized to 0 h
Hours

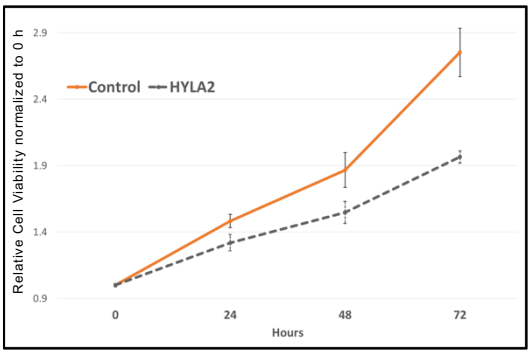
f


Control
TMEFF2
Relative Cell Viability normalized to 0 h
Hours

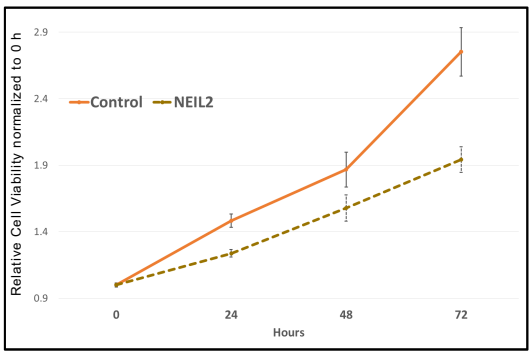
g


Control
LINC01529
Relative Cell Viability normalized to 0 h
Hours

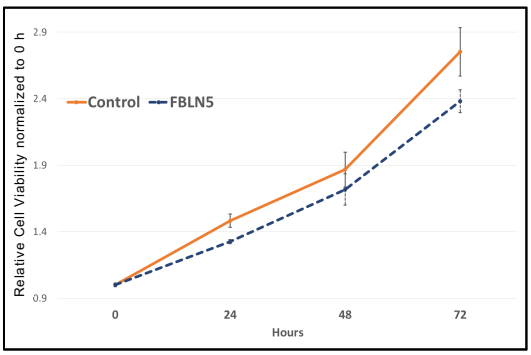
h


Control
HYLA2
Relative Cell Viability normalized to 0 h
Hours

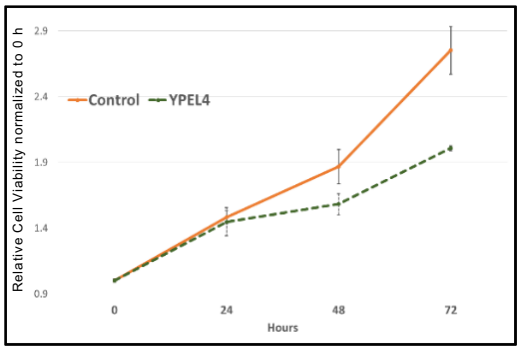
i


Control
NEIL2
Relative Cell Viability normalized to 0 h
Hours

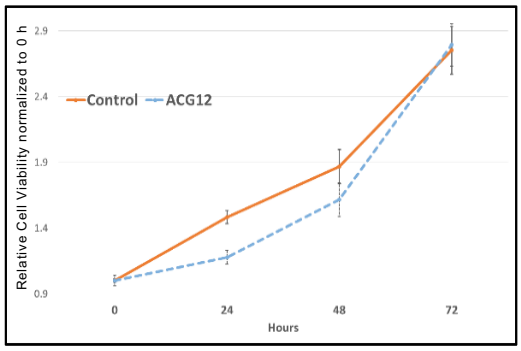
j


Control
FBLN5
Relative Cell Viability normalized to 0 h
Hours

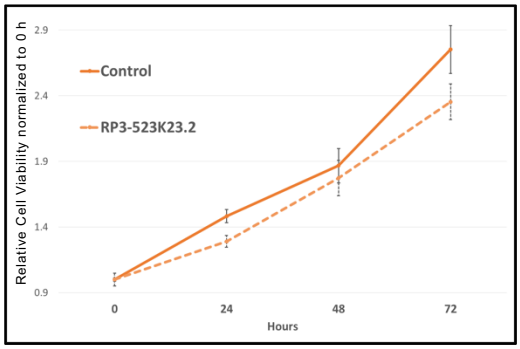
k


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
YPEL4
0
24
48
72
Hours

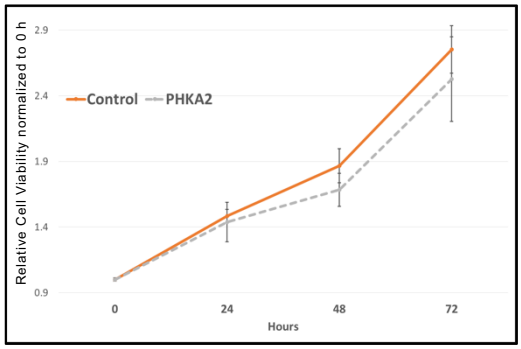
l


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
ACG12
0
24
48
72
Hours

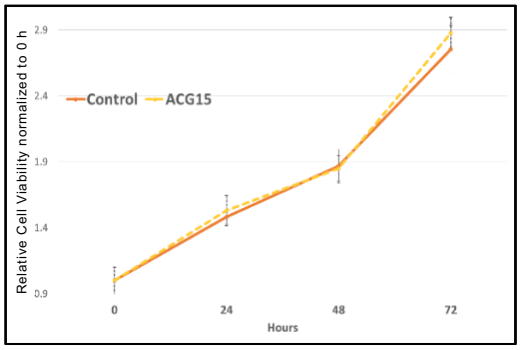
m


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
RP3-523K23.2
0
24
48
72
Hours

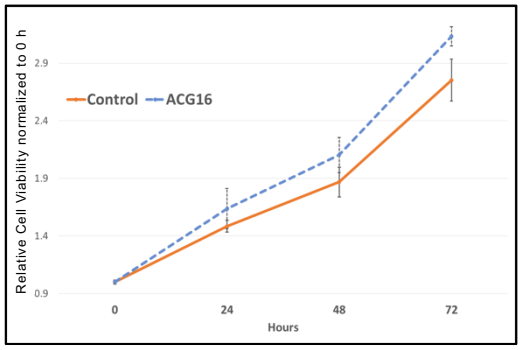
n


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
PHKA2
0
24
48
72
Hours

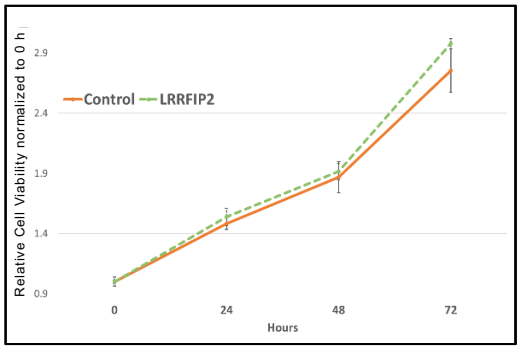
o


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
ACG15
0
24
48
72
Hours

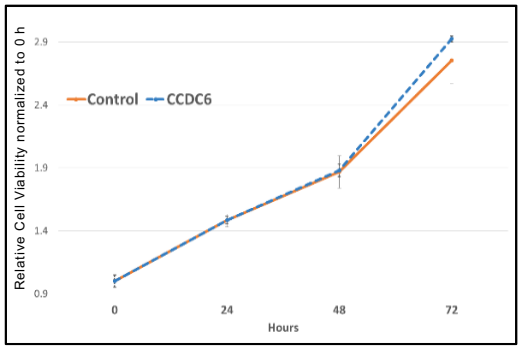
p


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
ACG16
0
24
48
72
Hours

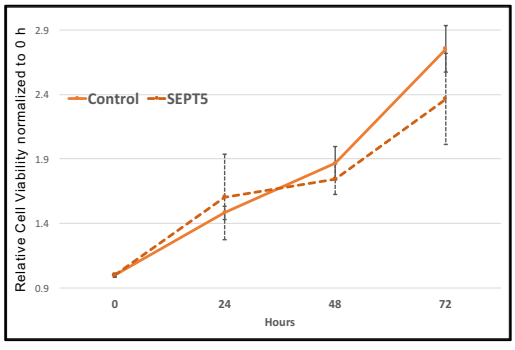
q


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
LRRFIP2
0
24
48
72
Hours

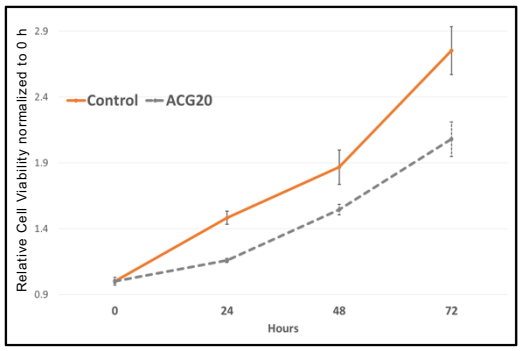
r


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
CCDC6
0
24
48
72
Hours

s


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
SEPT5
0
24
48
72
Hours

t


Relative Cell Viability normalized to 0 h
2.9
2.4
1.9
1.4
0.9
Control
ACG20
0
24
48
72
Hours

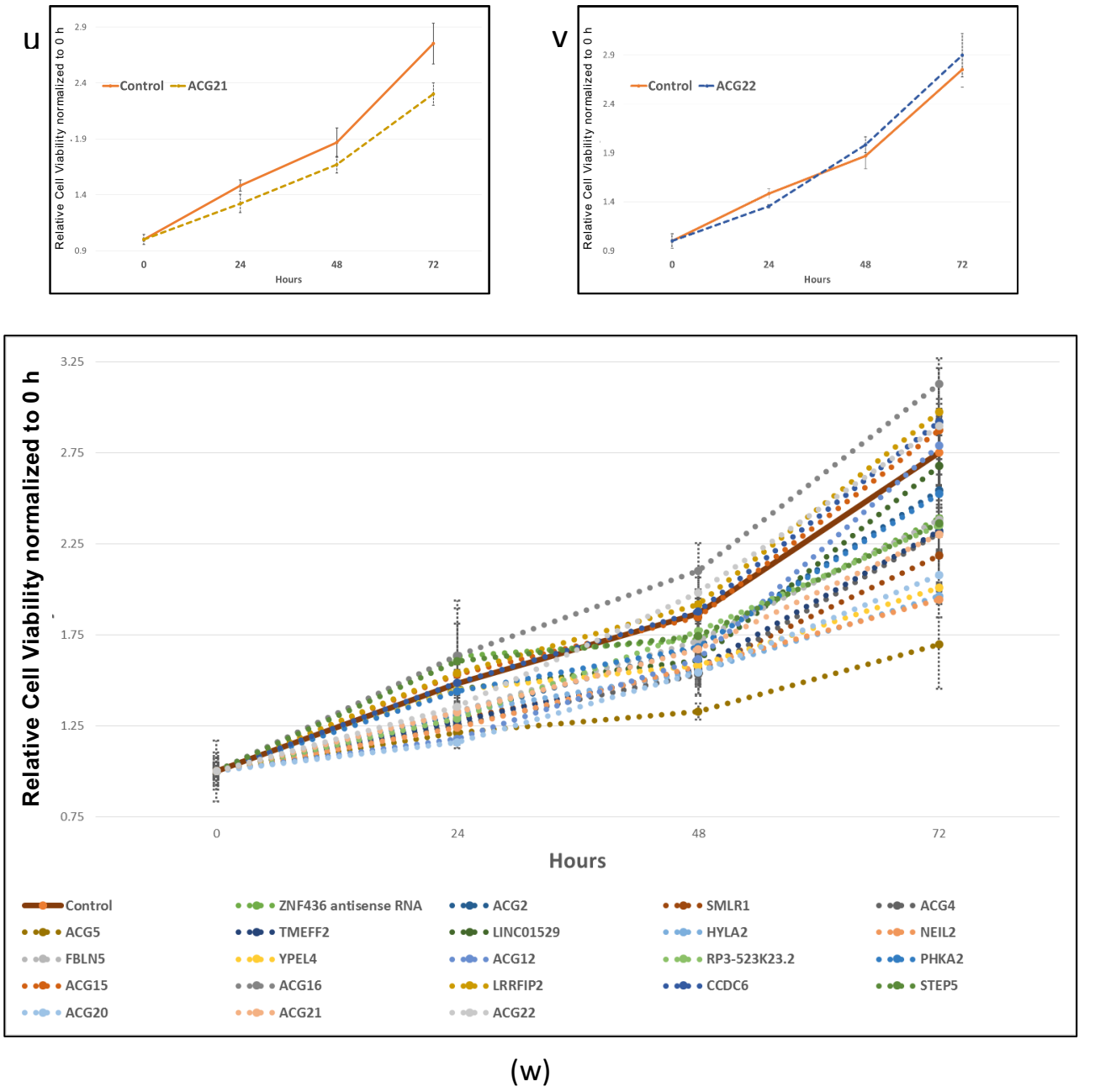


***Figure 3.32*: Relative viability of MCF-7 cells upon overexpression of novel anticancer genes – MTT colorimetric assay**. (a) ZNF436 antisense RNA, (b) ACG2, (c) SMLR1, (d) ACG4, (e) ACG5, (f) TMEF2, (g) LINC01529, (h) HYLA2, (i) NEIL2, (j) FBLN5, (k) YPEL4, (l) ACG12, (m) ACG13 (n) PHKA2 processed transcript, (o) ACG15, (p) ACG16, (q) ACG17, (r) ACG 18, (s) SEPT5, (t) ACG20, (u) ACG21, (v) ACG22, (w) compilation graph. β-Galactosidase was used as a negative 'control'. MCF-7 cells were transfected in 96-well plates. Each point represents the average reading of three wells of a 96-well plate, each with an independently transfected population of cells. Error bars represent +/- standard deviation. NB: this experiment was performed in collaboration with Qize Ding, and data for this experiment are therefore shared between this and Qize Ding's PhD thesis.

3.2.5 Selection of a subset for further investigation

Following these observations, I needed to choose a subset of genes for further investigation. It was deemed preferable to work with genes which were (i) protein coding, (ii) had some knowledge regarding their properties and functionality available, (iii) caused cell death and loss of cell-viability in MCF-7 cells, (iv) showed positive correlation between their expression levels and survival probability of breast cancer patients and (iv) had functioning antibodies available for them. The non-protein coding genes ZNF436_AS1, ACG2, ACG4, ACG5, LINC01529, ACG12, ACG13, PHKA2-proccessed transcript, ACG15, ACG16, ACG18, ACG20, ACG21 and ACG22 were thus deselected. In the case of LRRFIP2 and SEPT5, the sequence overlaps presumably identified novel transcript variants of these two genes. As the properties of these novel splice variants might have been different from those of the previously known splice variant, LRRFIP2 and SEPT5 were also deselected for further investigation.

3.2.6 Validation of overexpression

I overexpressed SMLR1, TMEFF2, HYAL2, NEIL2, FBLN5 and YPEL4 in MCF-7 cells and subjected the cell lysates to immunoblotting. Results are shown in figure 3.33. The highest degree of overexpression was observed in the case of TMEFF2 and FBLN5. HYAL2 and NEIL2 exhibited moderate overexpression whereas no overexpression was observed for SMLR1 and YPEL4. Moreover, the lysates of cells transfected with SMLR1 and YPEL4 exhibited numerous non-specific bands. After several failed attempts at getting the immunoblots of acceptable quality, I reasoned that the antibodies used to detect these two proteins were not appropriate enough as to proceed studying SMLR1 and YPEL4. These two genes were therefore deselected for further investigation.

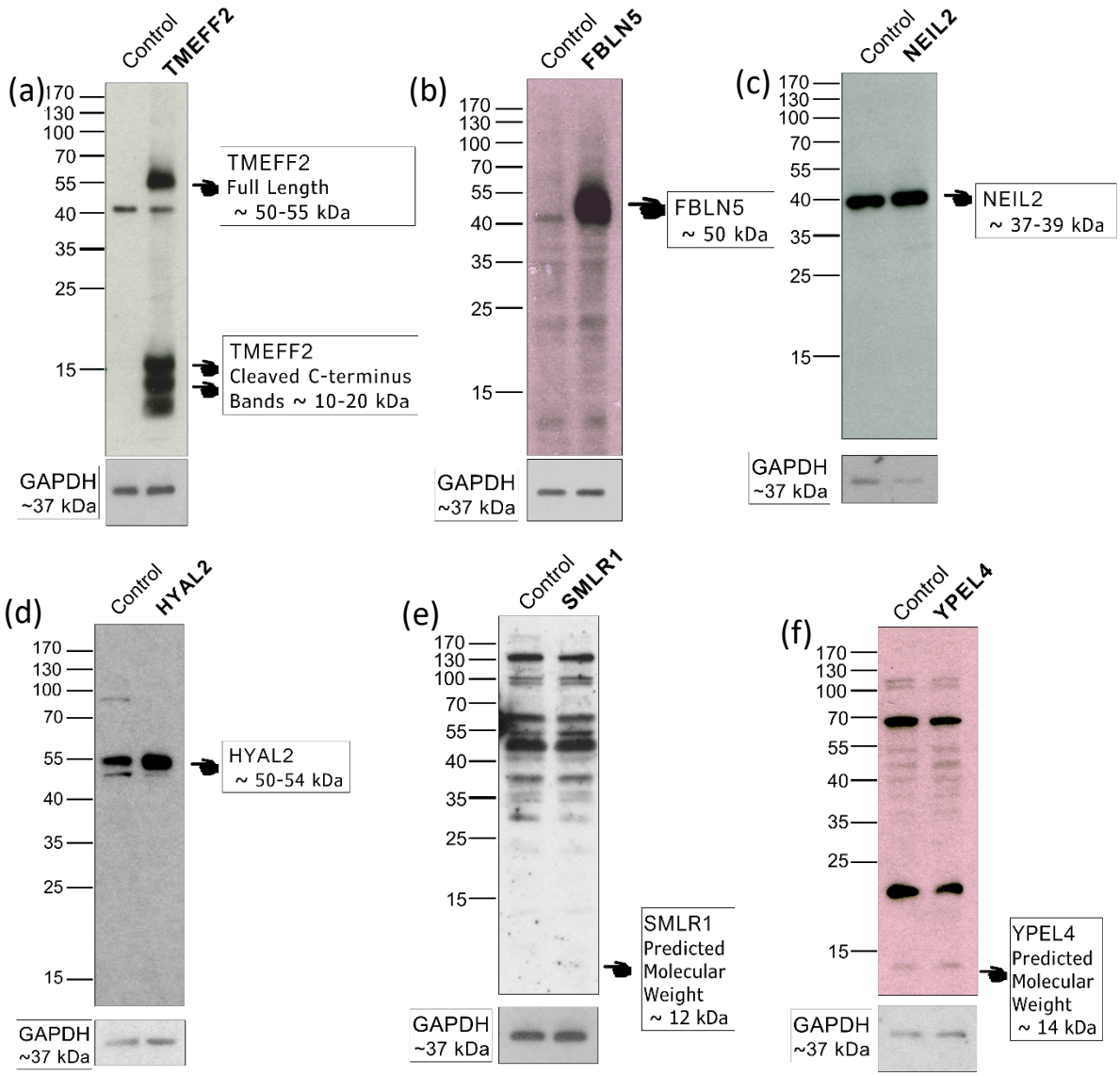


***Figure 3.33*: Experimental overexpression of anticancer genes in MCF-7 cells.** MCF-7 cells were transfected with (a) TMEFF2, (b) FBLN5, (c) NEIL2, (d) HYAL2, (e) SMLR1 and (f) YPEL4. Transient transfection with GFP in parallel was used as negative 'control'. Transfection was performed in 10 cm dishes and cells were harvested 30 hours post-transfection. Proteins were separated through SDS-PAGE under reducing conditions and blotted onto nitrocellulose membranes. Immunoblotting procedure and antibodies used are described in Materials and Methods. Immunoblotting (for the provided cell pellets) was performed by Master student Miss Ottilie Swinyard.

3.3 Discussion

The Ensembl genome browser served as a valuable tool in deciphering the identities of the anticancer genes. NITE [61] employed primer walking for sequencing the cDNA library transcripts. I subjected the anticancer genes plasmid preps to Sanger sequencing which provided relatively short sequences (mostly ~ 800 bp). To avoid the laborious and time-consuming primer walking, I simply checked if the results from Sanger sequencing matched with the ones provided by NITE. Once the right sequence was confirmed, I used the sequences provided by NITE to run the BLAST like alignment tool (BLAT) [69] through the Ensembl genome browser [70].

It was interesting to notice that 4 out of 22 anticancer genes aligned with the non-coding RNAs, i.e., ZNF436-AS1, LINC01529lincRNA, RP3-523K23.2-001 sense-overlapping-processed-transcript and PHKA2-002 processed-transcript. Other genes which either did not overlap any previously annotated transcripts or overlapped the introns of some previously annotated genes could also possibly be non-protein-coding. According to MacFarlane and Murphy [95], '*miRNA precursors are commonly found in clusters through many different regions of the genome, most frequently within intergenic regions and introns of protein-coding genes*' (p.538). ACG2, ACG15, ACG20 and ACG21 lie within the intronic regions of *HHLA2, ERC1, VTI1A,* and *HHIP* respectively. These four ACGs could be the precursor sequences for novel microRNAs. Non-coding RNAs have attracted major attention over the past two decades and have been implicated in diverse processes ranging from apoptosis and differentiation to metastasis and epigenetic regulations [96]. To the best of my knowledge, it is the first time that non-coding transcripts have been identified as anticancer genes. I believe the possibility of non-coding RNAs as anticancer genes holds some exciting prospects. The sequences of ACG5 and ACG16 overlapped with the genomic sequence in a manner similar to exons. These could be novel genes. Subjecting their sequences to the bioinformatic nucleotide-to-protein translation tools could yield interesting results (these two could also be non-coding RNAs). In addition, it appears that novel splice variants for LRRFIP2 and SEPT5 have been identified, which is also worth further investigation. I did not conduct these further studies because the objectives of my project where to select a subset of genes and initiate investigation regarding their regulation and functioning.

Regarding the protein-coding anticancer genes, these belonged to a wide range of functional spectrum. Hyaluronidase 2 (HYAL2), for example, is a ubiquitous extracellular matrix (ECM) hyaluronan-degrading enzyme with some novel functions, such as CD44 regulation [97]. Recently,

mutations in *HYAL2* have been implicated in orofacial clefting [98]. NEIL2 (Nei-Like DNA Glycosylase 2) is involved in DNA base excision repair [99]. Fibulin-5 (FBLN5) is a secreted ECM glycoprotein which is involved in the development and stabilization of ECM fibres (especially elastic fibres) as well as integrin signalling [100]. The functions of SMLR1, YPEL4 and TMEFF2 are largely unknown. The small leucine-rich protein 1 (SMLR1) is the most nascent among these three as I could not find any study regarding its function, but the protein appears to be part of the extracellular matrix (ECM), small leucine-rich proteoglycan family [101]. Yippee-like-4 (YPEL4), on the other hand, has been reported to be involved in cell proliferation [102]. The transmembrane protein with an EGF-like and two follistatin-like domains 2 (TMEFF2) is involved in a wide range of processes including providing neuroprotection in Alzheimer's disease and triggering the JAK-STAT signalling pathway in colorectal cancer (please refer to chapter 5 for details). The functions profile of these genes fit nicely with the concept of anticancer genes, i.e., they can belong to any strata of biological functions. For example, MDA-7/IL-24 is a secreted cytokine, Orctl3 is an orphan receptor, E4orf4 is an adenoviral early-expression gene which is involved in activating mTOR pathway during viral infection, and HAMLET is, in fact, a misfolded protein. It appears that in the heavily distorted signalling circuitry of transformed and cancer cells, the prospects of synthetic lethality know no bounds.

On the other hand, almost all of the tested genes exhibited a positive correlation between their mRNA expression levels and the survival of breast cancer patients. In addition, NEIL2 and FBLN5 also showed similar trends in lung cancer. The question that arises here is if anticancer genes fundamentally differ from tumour suppressor genes in the sense the latter are responsible for inhibition of the neoplastic transformation and/or metastasis at the physiological expression levels, and the former only cause selective cell death in cancer and transformed cells upon ectopic overexpression on the basis of synthetic lethality, why was a positive correlation between mRNA expression levels of these novel anticancer genes and cancer patient survival observed? Among the explanations are: (i) as I explained in the introduction, tumour suppressor genes can also have anticancer gene activity; the prime example is that of PAR4, and (ii) the anticancer genes could be playing a part in the prevention of neoplastic transformation and/or spread through synthetic lethality rather than the canonical pathways involved in the functioning of tumour suppressor genes. That is, higher levels of MDA-7/IL-24 would cause ER stress and ROS production which would be lethal to cancer cells and not the normal cells. In comparison, p53 would act by inducing apoptosis. In this regard, the ectopic overexpression of anticancer genes would provide the advantage of selective cell death only in cancer

cells, while the overexpression of tumour suppressor genes would cause cell death indiscriminately. Therefore, the possibilities of anticancer genes having a positive or no correlation between their mRNA levels and cancer patient survival probability are equal. However, observing the negative correlation between the mRNA levels of the tested anticancer genes and survival of gastric cancer patients was surprising. The probable explanations for this phenomenon are: (i) the anticancer genes tested could indeed have an oncogenic role in gastric cancer; for example, the expression levels of FBLN5 have been reported to decrease in cancers of breast, ovary, bladder, liver, lung and colon, but other reports have demonstrated the upregulation of FBLN5 expression in gastric and nasopharyngeal cancers (please refer to chapter 6 for details and references). (ii) On the other hand, TMEFF2 downregulation has been reported in gastric cancer; moreover, the report also demonstrated a direct correlation between TMEFF2 promoter methylation and worsening patient survival [103]. This implies that mRNA levels cannot always be the true depiction of the functional capacity of a gene. For example, p53 is primarily silenced through mutation rather than deletion/silencing [1]. Concerning how mRNA levels could depict the true state of affairs in breast cancer and the opposite in gastric cancer, the results observed in one system cannot necessarily be transferable to other systems. It would not be surprising for a gene and/or its protein product to act differently in tissue, cell and microenvironmental contexts. Nevertheless, platforms like kmplot.com provide rapid and reasonably accurate preliminary results. Taken together, these qualities were useful in the context of my study where I had to select a subset of genes for further investigation.

As the genes went through the same genetic screening process, some commonalities in their functioning were not surprising: most of the tested genes showed a positive correlation with the patient survival in breast cancer. Assuming that this commonality could be extrapolated to the rest of 15 anticancer genes, for which I could not get the patient survivability data, I tested for the effects of overexpression of all 22 anticancer genes in MCF-7 cells. I have already explained the reasons behind choosing the MCF-7 cells as the cell culture model for breast cancer in the results section. All seven anticancer genes which showed a positive correlation with breast cancer patient survival also exhibited induction of cell death in MCF-7 cells upon their overexpression. The results were confirmed through checking the short-term cell viability of MCF-7 cells upon overexpression of anticancer genes using the MTT assay [94]. Following these experiments, I shortlisted the anticancer genes which were to be investigated further. As mentioned previously, SMLR1, TMEFF2, HYAL2, NEIL2, FBLN5 and YPEL4 were chosen because, firstly, these were protein coding genes with some previous

knowledge available about them and secondly, these caused cell death in MCF-7 cells and showed a positive correlation between their mRNA levels and breast cancer patient survival. Moreover, I had also decided upon the cancer cell model in which to investigate the regulation and function of these genes, i.e., MCF-7 cells.

Lastly, I employed immunoblotting to determine the overexpression of six anticancer genes. The reasons behind using immunoblotting and not qRT-PCR (quantitative real-time polymerase chain reaction) were: (i) although the cDNA sequences for these genes provided by NITE [61] overlapped with those of previously coded genes, the Sanger sequencing through primer walking could still be prone to error and the transcripts for these genes could still be non-coding. (ii) If I were to investigate the functioning of these genes, it was desirable to have working antibodies for them. (iii) Higher mRNA levels would not necessarily translate into protein – in case the selective cell death was being caused by mRNA only, altogether different strategies would have to be employed for further investigation regarding the functioning of these genes. The highest protein overexpression was observed for TMEFF2 and FBLN5, some degree of overexpression was seen in NEIL2 and HYAL2, while no upregulated protein band was observed for either SMLR1 or YPEL4. The latter also exhibited large scale non-specific binding and I concluded that the antibodies used for these two proteins were not functional.

4 Anticancer Genes and FOX proteins

4.1 Background

4.1.1 Forkhead box (FOX) proteins

The forkhead box proteins comprise a family of transcription factors having the 'winged-helix' DNA-binding domain. The name forkhead comes from the appearance of an ectopic head structure in *D. melanogaster* embryos upon mutation in this homeotic gene [104]. Since the discovery of the first FOX protein by Weigel *et al.* [105] in 1989, a great deal of information has become available implicating the role of FOX proteins in regulating diverse cellular processes, ranging from apoptosis and senescence to cell cycle progression, differentiation, DNA damage repair and metabolism [104]. It appears that the FOX transcription factors have evolved primarily to coordinate temporal and spatial gene expression during embryonic development [106]. However, owing to their ability to impart pleotropic changes over diverse cellular processes, the involvement of FOX proteins in cancer initiation, progression and survival is not surprising. In this regard, three members of the FOX family – FOXM1, FOXO3 and FOXK2 – are important in particular.

4.1.1.1 FOXM1

Forkhead box M1 (FOXM1), also known as Trident, HFH-11, HNF-3, MPP2 and Win [104], is one of the key transcription factors that regulate the complex spatiotemporal gene expression throughout embryonic and foetal development. On the other hand, FOXM1 expression is hardly detectable in the differentiated adult tissue except for in the actively proliferating cells [107]. FOXM1 aids cell cycle progression through G1 to S phase and also through G2 to M phase transition [104]. FOXM1 drives progression through G1 to S phase by modulating the gene expression of the key components of the Skp/Cullin/F-box containing complex (SCF complex). The SCF complex in turn is responsible for the degradation of cyclin-dependent kinase inhibitor (CDKI) proteins during the G1-S transition [106]. For G2 to M phase progression, FOXM1 controls several critical regulators of G2 to M phase transition; these include cyclin A2, cyclin B, cell division cycle 25B (CDC25B), polo-like kinase 1 (PLK1), survivin, aurora A kinase, aurora B kinase and centromere protein A (CENPA), CENPB and CENPF [106]. As FOXM1 is a pivotal regulator of cell proliferation and cell cycle progression, the deregulated FOXM1 leads to the loss of critical cell cycle checkpoints resulting in neoplasia. It is not surprising that FOXM1

upregulation is found in a variety of carcinomas [108]. The underlying molecular mechanisms whereby FOXM1 drives tumorigenesis are not clear yet. So far it is clear, though, that FOXM1 interacts with several key pathways involved in neoplastic transformation and progression, such as PI3K/AKT, the oestrogen receptor (ER), epidermal growth factor (EGFR), ataxia telangiectasia mutated (ATM) and the retinoblastoma protein (RB)-E2F pathways [109]. The role of FOXM1 does not end at tumour formation but FOXM1 also helps the tumours with angiogenesis [110] and metastasis [106]. In addition, FOXM1 is also involved in the epithelial-to-mesenchymal transition (EMT) and in imparting undifferentiated stem cell-like properties at the same time [106]. Furthermore, FOXM1 has emerged as a major player in the genotoxic agent resistance [111].

4.1.1.2 FOXO3

As opposed to FOXM1, the members of subclass 'O' of forkhead transcription factors (FOXO) are *bona fide* tumour suppressors [104]. FOXOs in general, and FOXO3 (also known as FKHRL1) in particular, are suppressors of proliferation and survival. Interestingly, FOXM1 and FOXO3 share the same forkhead response element (FHRE) and compete for its binding and, subsequently, the opposite downstream effects [106]. FOXO3 suppresses the activity of FOXM1 by competing for the same DNA binding site. One of the examples of this phenomenon is the FHRE in proximal region of the VEGF promoter in the breast cancer cells where FOXM1 promotes VEGF transcription while FOXO3 represses it [110]. FOXM1, in fact, is an active transcriptional repressor of FOXO3 [104]. One of the ways in which the FOXOs suppress cell proliferation is through stalling cell cycle progression by inducing the expression of p15 (INK4B), p19 (INK4D) [112], p21 (CIP1) [113], p27 (KIP1) [114, 115] and GADD45A [116]. In addition, FOXOs repress cyclin G2 [117], cyclin D1 and cyclin D2 [63, 104]. The tumour suppressor role of FOXO3 is not limited to its ability to halt cell cycle, but under certain external stimuli (discussed later), FOXO3 and other members of the FOXO family can also induce regulated cell death. FOXO3 induces apoptosis by expression of pro-apoptotic genes FASLG [62] and BCL2L11 [118]. Other key pro-apoptotic genes including TNFSF10, BAD and CASP9 are under the direct transcription control of FOXOs [104]. FOXO3, along with other members of the FOXO family, is regulated mainly under the PI3K/AKT pathway. Other oncogenic pathways like NFκB, IKKβ and ERK also regulate FOXO3 [104]. Moreover, the primary mode of FOXO3 regulation is its subcellular localization, which in turn is governed by its post-translational modification state [119]. For example, under the regulation of insulin and insulin-like growth factors, AKT and SGK (serum and glucocorticoid-induced kinase) phosphorylate FOXO3 at T32, S253 and S315, which causes both the

nuclear export and cytoplasmic sequestration of FOXO3, thus preventing FOXO3 from aiding the expression of anti-proliferative and pro-apoptotic genes [104]. On the other hand, phosphorylation of FOXO3 at S207 by mammalian Ste20-like kinase1 (MST1) re-localizes it from cytoplasm to nucleus [104]. Phosphorylation of FOXO3 on S07 by p38 in response to doxorubicin also causes its nuclear re-localization [120]. Acetylation of FOXO3 interferes with its DNA binding ability and the ubiquitination of FOXO3 by the inhibitor of nuclear factor kappa B kinase subunit beta (IKKβ) at S644 marks it for degradation [104]. FOXO3 also indirectly inhibits metastasis by inducing the expression of the CREB-binding protein/p300-interacting transactivator with Asp/Glu-rich C-terminal domain 2 (CITED2). CITED2 prevents p300 from binding HIF1α; this in turn prevents the hypoxia-inducible factor 1-α (HIF1α) from inducing the expression of metastatic genes [104]. As mentioned previously, FOXO3 also suppresses VEGF in breast cancer cells, thus preventing vascularization and angiogenesis.

4.1.1.2.1 FOXO3-A3, the Always Active Mutant

Brunet *et al.*, [62] developed a mutant of FOXO3 where the AKT phosphorylation sites T32, S253, and S315 were replaced by the non-phosphorylatable alanine residues. As mentioned earlier, this inability to phosphorylate at these three residues leaves FOXO3 incapable of being exported from the nucleus and sequestered into the cytoplasm. This mutant form of FOXO3 is rereferred to as FOXO3-A3-always active mutant.

4.1.1.3 FOXK2

Of all the family members of the forkhead box family, the least investigated so far is FOXK2. A recent study has shown that the activation of FOXO3 by the conventional chemotherapeutic agents epirubicin and paclitaxel is mediated by FOXK2 [65]. Moreover, FOXK2 imparts anti-proliferative effects in ERα-positive breast cancer cells where FOXK2 inhibits the transcriptional activity of ERα by acting as a scaffold protein between ERα and its E3 ubiquitin ligase, the BRCA1-BARD1 heterodimer [121]. FOXK2 also acts as a sequence specific nuclear localization aid for the tumour suppressor BAP1 [122]. In addition, FOXK2 suppresses starvation-induced autophagy (regulated by mTOR) through inhibiting the expression of key autophagy related genes [123].

4.1.2 FOX proteins and Anticancer genes?

FOXM1 and FOXO3 play an important role in both the targeted and conventional anticancer therapies. Targeted therapy with tyrosine kinase inhibitors, such as imatinib, indirectly activates FOXO3 through

inhibiting the PI3K/AKT pathway. The epidermal growth factor receptor (EGFR) inhibitor, gefitinib, on the other hand, is suggested to activate FOXO3 through mechanisms independent of PI3K/AKT signalling [124]. As mentioned before, the conventional chemotherapeutic agent doxorubicin at least in part confers its cytotoxicity by inducing nuclear localization of FOXO3 [120]. It is also suggested that other conventional chemotherapeutic agents like epirubicin and paclitaxel activate the anti-proliferative and pro-apoptotic functions of FOXO3 [104]. As FOXO3 competes with FOXM1 for DNA binding and FOXM1 transcriptionally represses FOXO3, any effects of anticancer therapy imparted through FOXO3 would also involve FOXM1 and *vice versa*. Moreover, FOXM1 is found to be overexpressed and FOXO3 activity is suppressed in a variety of drug-resistant cancer cells [104]. Therefore, both the conventional and non-conventional chemotherapeutic agents, in one way or the other, influence the FOXM1-FOXO3 axis (figure 4.1). Hence, in the context of anticancer genes, the FOXM1-FOXO3 axis presents an attractive point of investigation. I also included FOXK2 in the study because of its anti-proliferative role and correlation with FOXO3.

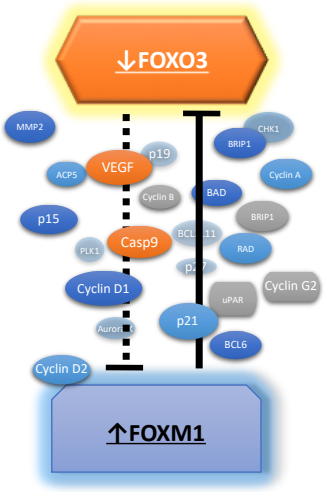


***Figure 4.1*: The FOXM1-FOXO3 axis:** the oncogenic transcription factor FOXM1 actively represses the transcription of the tumour suppressor gene, FOXO3. The transcription factor FOXO3 indirectly represses FOXM1 by competing for the same DNA binding site.

4.2 Results

4.2.1 Regulation of Anticancer genes by FOX proteins

I had selected SMLR1, TMEFF2, HAYL2, NEIL2, FBLN5 and YPEL4 for further investigation. As forkhead box (FOX) proteins are pleotropic transcription factors, I checked if the selected anticancer genes were

being regulated by FOX proteins. The reasons behind choosing FOXM1, FOXO3 and FOXK2 for this purpose are explained already in the background section and the reasons for using MCF-7 cells are explained in the previous chapter. In order to test if the selected anticancer genes were being regulated by the chosen FOX proteins in MCF-7 cells, I overexpressed and knocked-down FOXM1, FOXO3 and FOXK2 in MCF-7 cells and tested for the changes in expression levels of the anticancer genes using qRT-PCR and immunoblotting. The challenge in this regard was that the native expression levels of the chosen anticancer genes were very low in MCF-7 cells. Consequently, getting an efficient PCR reaction was very difficult. I spent considerable time and effort in undertaking various optimization procedures to obtain efficient PCR reactions. These included redesigning primers, varying concentrations of primers and template in the reactions, changing the reaction conditions in terms of temperatures and timings and employing gene specific primers for reverse transcription. A PCR reaction with efficiency comparable to that of the reference housekeeping gene (L19) was successfully obtained for SMLR1 through two-step qRT-PCR by redesigning of primers. For TMEFF2, the gene-specific primer was used in the reverse transcription step of the two-step qRT-PCR in order to obtain the PCR reaction with comparable efficiency (to that of housekeeping gene, L19). For FBLN5, HYAL2 and NEIL2, an additional 15 minutes of incubation at 50°C in the reverse transcription reaction worked successfully (canonically 15 min at 42°C was in practice). Despite all the efforts, a reaction efficiency comparable to L19 housekeeping gene could not be achieved for YPEL4. In the interest of time, I am not presenting the results of the lengthy optimization procedures undertaken.

Another challenge in this experiment was that in MCF-7 cells, the PI3K/AKT pathway is not only constitutively active but this active state is crucial for the survival of these cells [125]. Consequently, it was highly probable that even the overexpressed FOXO3 would be phosphorylated and inactivated (by constitutively active AKT). As a solution, in addition to transfection with FOXO3, I also transfected the MCF-7 cells with FOXO3-A3 mutant (explained earlier).

4.2.1.1 Confirmation of FOX overexpression and knockdown

For the overexpression, I transfected MCF-7 cells with plasmids harbouring FOXM1, FOXO3, FOXO3-A3-always active mutant and FOXK2 in 10 cm tissue culture dishes. The procedure was repeated twice, i.e., overexpression was performed in triplicate. After harvesting the cells at 30 hours post-transfection, I divided the cell pellets into two groups, one for qRT-PCR and the other one for immunoblotting. A similar approach was adopted for the knockdowns, except, in this case, the siRNA

transfection was performed only in one batch of MCF-7 cells. The qRT-PCR results for the confirmation of overexpression of FOXM1, FOXO3 and FOXO3-A3-always active mutant and FOXK2 in all three batches (separately) of MCF-7 cells are depicted in figures 4.2(a), 4.2(c) and 4.2(e) respectively. The qRT-PCR results for confirmation of FOXM1, FOXO3 and FOXK2 knockdowns are presented in figures 4.2(b), 4.2(d) and 4.3(f) respectively. Moreover, the immunoblot showing the overexpression of FOXM1, FOXO3 and FOXO3-A3-always active mutant and FOXK2 from one of the three batches, and the knockdown from the only batch of MCF-7 cells is shown in figure 4.3. Considerable variation in the degree of overexpression of all three FOX genes was observed between the three batches of MCF-7 cells. The highest degree of this inter-batch variation was observed in the case of FOXK2 overexpression, where the third batch showed a very low degree of overexpression. Consequently, I assayed for the mRNA expression levels of the selected anticancer genes in all three batches of MCF-7 cells with FOXM1, FOXO3 and FOXO3-A3-always active mutant overexpression while I used only the first two batches of the cells with FOXK2 overexpression.

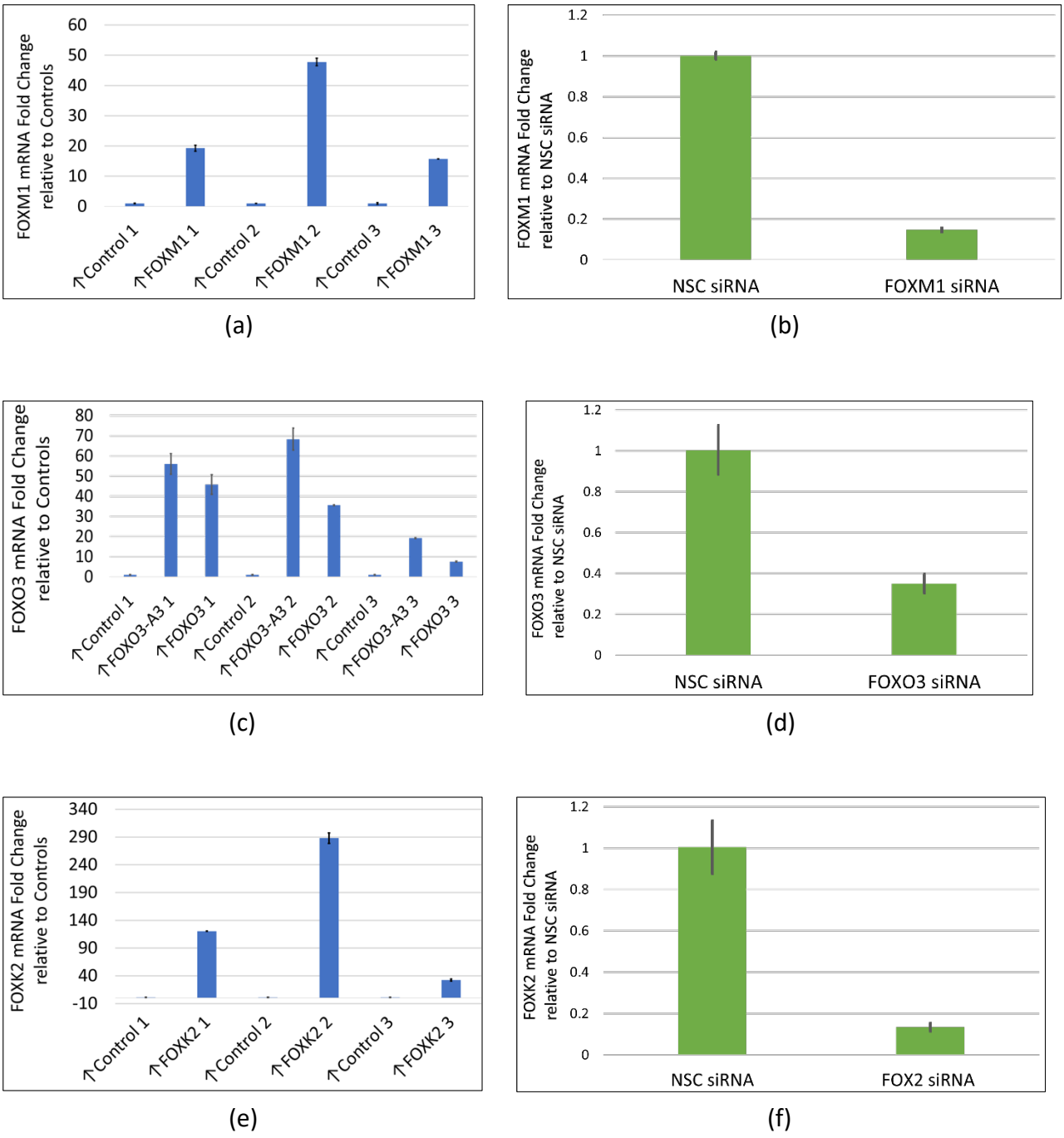


***Figure 4.2*: qRT-PCR results for the validation of overexpression and knockdown of FOXM1, FOXO3, FOXK2 and FOXO3-A3-always active mutant (overexpression only) in MCF-7 cells.** Relative mRNA fold change was calculated according to ΔΔCt method using L19 as housekeeping gene. For overexpression, transient transfection with GFP was used as negative 'control(s)' and cells were harvested 30 hours post-transfection. For knockdown, nonspecific siRNA (NSC) was used as negative control and cells were harvested 48 hours post-transfection. (a) Fold increase in FOXM1 mRNA levels upon transient transfection with FOXM1 in three batches of MCF-7 cells. (b) Fold decrease in FOXM1 mRNA level upon transient knockdown of FOXM1 using siRNA. (c) Fold increase in FOXO3-A3-always active mutant and FOXO3 mRNA levels upon transient transfection with FOXO3 in three batches of MCF-7 cells. (d) Fold decrease in FOXO3 mRNA level upon

transient knockdown of FOXO3 using siRNA. (e) Fold increase in FOXK2 mRNA levels upon transient transfection with FOXK2 in three batches of MCF-7 cells. (f) Fold decrease in FOXK2 mRNA level upon transient knockdown of FOXK2 using siRNA. Bars represent the average form 3 PCR-reactions. Error bars represent +/- standard deviation.

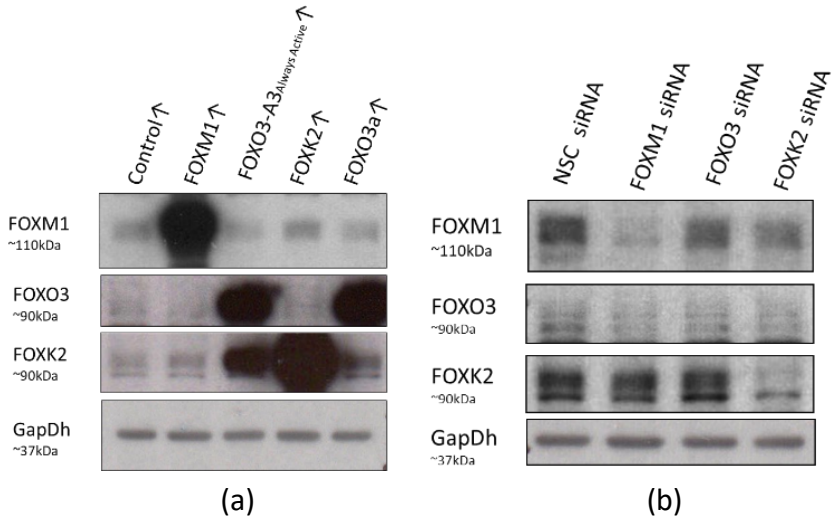
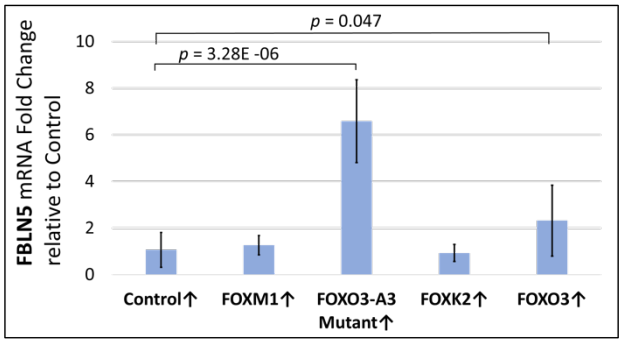


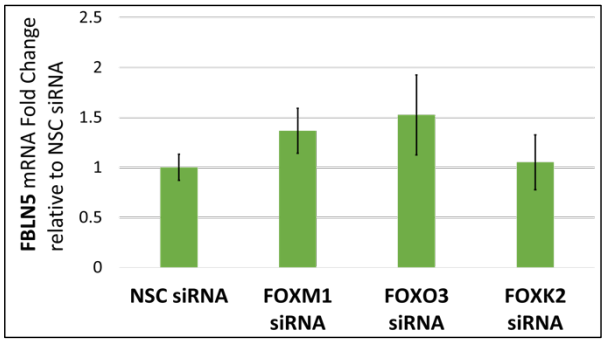
***Figure 4.3*: Immunoblots for the confirmation of (a) overexpression and (b) knockdown of FOXM1, FOXO3, FOXK2 and FOXO3-A3-always active mutant (overexpression only) in MCF-7 cells.** For overexpression, transient transfection with GFP was used as negative 'control' and cells were harvested 30 hours post-transfection. Immunoblot is showing the results for one out of three independently transfected batches of MCF-7 cells. For knockdown, non-specific (NSC) siRNA was used as negative control and cells were harvested 48 hours post-transfection.

4.2.1.2 Effects of FOX overexpression and knockdown on the mRNA levels of selected anticancer genes.

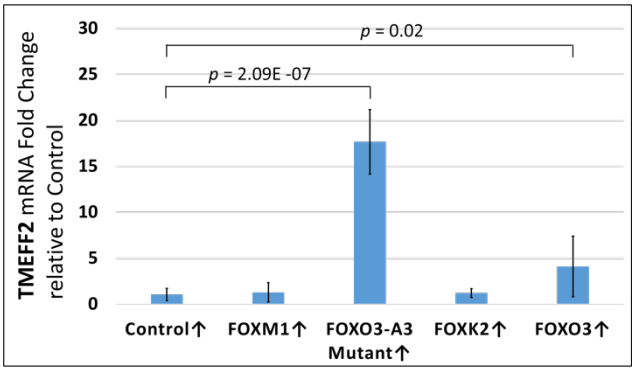
The changes in mRNA levels of FBLN5, TMEFF2, SMLR1, HYAL2 and NEIL2 were assayed using qRT-PCR upon transient overexpression and knockdown of FOXM1, FOXO3, FOXK2 and FOXO3-A3-always active mutant (overexpression only) in MCF-7 cells. Results are presented in figure 4.4. FBLN5 and TMEFF2 showed a dramatic ~7-fold and ~18-fold increase, respectively, in their relative mRNA levels upon overexpression of FOXO3-A3-always active mutant. The mRNA levels of both also increased marginally upon overexpression of wild type FOXO3. An opposite trend, however, was not observed in the case of FOXO3 knockdown. Moreover, no noticeable change in mRNA levels of FBLN5 or TMEFF2 was observed upon overexpression, or knockdown, of FOXM1 or FOXK2. A small increase in mRNA levels of SMLR1 and a small decrease in mRNA levels of HYAL2 were observed upon overexpression of FOXO3-A3-always active mutant.


p = 0.047
p = 3.28E -06
FBLN5 mRNA Fold Change relative to Control
10
8
6
4
2
0
Control↑
FOXM1↑
FOXO3-A3 Mutant↑
FOXK2↑
FOXO3↑

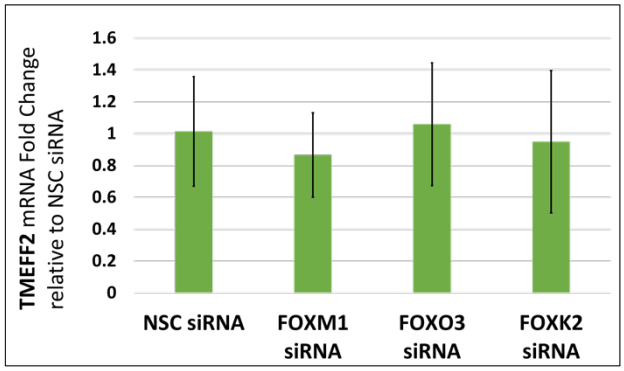
(a)


FBLN5 mRNA Fold Change relative to NSC siRNA
2.5
2
1.5
1
0.5
0
NSC siRNA
FOXM1 siRNA
FOXO3 siRNA
FOXK2 siRNA

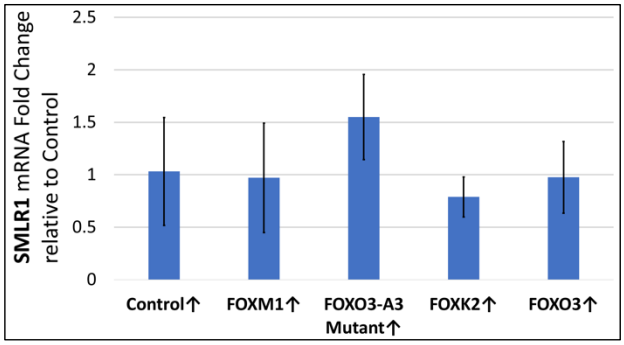
(b)


p = 0.02
p = 2.09E -07
TMEFF2 mRNA Fold Change relative to Control
30
25
20
15
10
5
0
Control↑
FOXM1↑
FOXO3-A3 Mutant↑
FOXK2↑
FOXO3↑

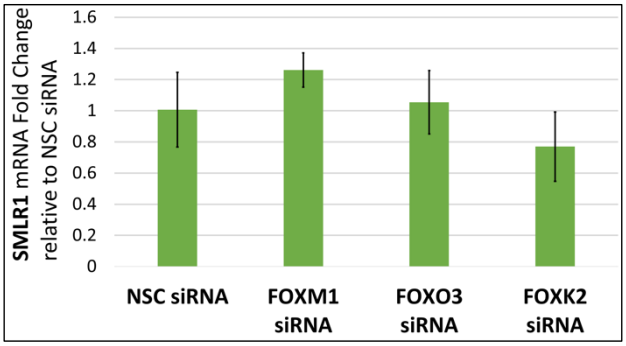
(c)


TMEFF2 mRNA Fold Change relative to NSC siRNA
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
NSC siRNA
FOXM1 siRNA
FOXO3 siRNA
FOXK2 siRNA

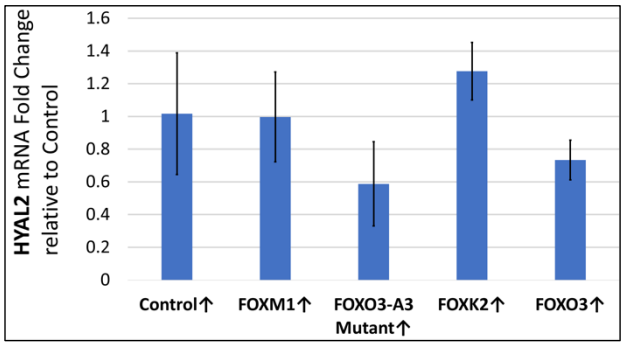
(d)


SMLR1 mRNA Fold Change relative to Control
2.5
2
1.5
1
0.5
0
Control↑
FOXM1↑
FOXO3-A3 Mutant↑
FOXK2↑
FOXO3↑

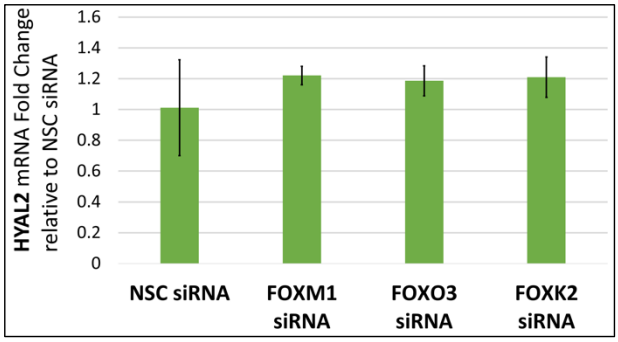
(e)


SMLR1 mRNA Fold Change relative to NSC siRNA
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
NSC siRNA
FOXM1 siRNA
FOXO3 siRNA
FOXK2 siRNA

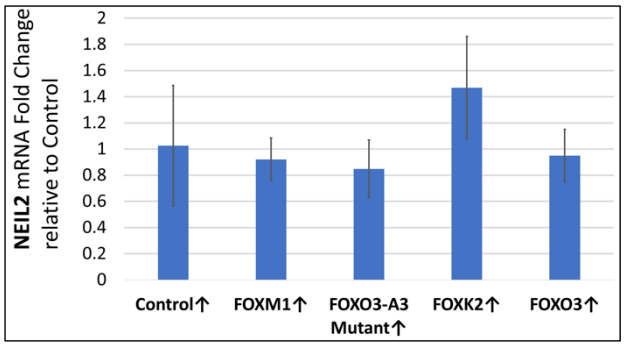
(f)


HYAL2 mRNA Fold Change relative to Control
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Control↑
FOXM1↑
FOXO3-A3 Mutant↑
FOXK2↑
FOXO3↑

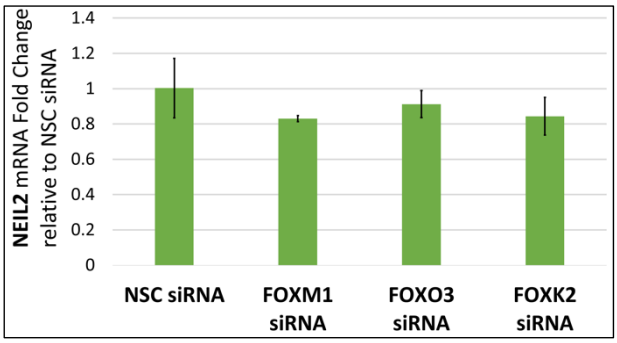
(g)


HYAL2 mRNA Fold Change relative to NSC siRNA
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
NSC siRNA
FOXM1 siRNA
FOXO3 siRNA
FOXK2 siRNA

(h)



(i)



(j)

***Figure 4.4*: Effects of overexpression and knockdowns of FOXM1, FOXO3, FOXK2 and FOXO3-A3-always active mutant (overexpression only) on the mRNA levels of anticancer genes in MCF-7 cells.** The relative mRNA fold change was calculated according to ΔΔCt method using L19 as housekeeping gene. For overexpression, cells were harvested 30 hours post-transfection and transient transfection with GFP was used as negative 'control'; qRT-PCR readings represented are the average from three independently transfected batches of MCF-7 cells, except for FOXK2 overexpression where qRT-PCR readings are the average of two independently transfected batches of MCF-7 cells. For knockdown, cells were harvested 48 hours post-transfection and nonspecific (NSC) siRNA was used as negative control; qRT-PCR reactions were performed on cDNA produced from a single batch of transiently transfected MCF-7 cells and bars represent average of 3 reactions. (a) Fold change in mRNA levels of FBLN5 upon transient overexpression of FOXM1, FOXO3-A3-always active mutant, FOXO3 (wild type) and FOXK2 relative to transiently transfected GFP as control. p value was calculated using two-sample t-test assuming unequal variances. (b) Fold change in mRNA levels of FBLN5 upon transient knockdown of FOXM1, FOXO3 and FOXK2 through siRNAs relative to NSC siRNA. (c) Fold change in mRNA levels of TMEFF2 upon transient overexpression of FOXM1, FOXO3-A3-always active mutant, FOXO3 (wild type) and FOXK2 relative to transiently transfected GFP as control. p value was calculated using two-sample t-test assuming unequal variances. (d) Fold change in mRNA level of TMEFF2 upon transient knockdown of FOXM1, FOXO3 and FOXK2 through siRNAs relative to NSC siRNA. (e) Fold change in mRNA levels of SMLR1 upon transient overexpression of FOXM1, FOXO3-A3-always active mutant, FOXO3 (wild type) and FOXK2 relative to transiently transfected GFP as control. (f) Fold change in mRNA levels of SMLR1 upon transient knockdown of FOXM1, FOXO3 and FOXK2 through siRNAs relative to NSC siRNA. (g) Fold change in mRNA levels of HYAL2 upon transient overexpression of FOXM1, FOXO3-A3-always active mutant, FOXO3 (wild type) and FOXK2 relative to transiently transfected GFP as control. (h) Fold change in mRNA levels of HYAL2 upon transient knockdown of FOXM1, FOXO3 and FOXK2 through siRNAs relative to NSC siRNA. (i) Fold change in mRNA levels of NEIL2 upon transient overexpression of FOXM1, FOXO3-A3-always active mutant, FOXO3 (wild type) and FOXK2 relative to transiently transfected GFP as control. (j) Fold change in mRNA levels of NEIL2 upon transient knockdown of FOXM1, FOXO3 and FOXK2 through siRNAs relative to NSC siRNA.

An increase of 0.5-fold was observed in the relative mRNA levels of NEIL2 upon FOXK2 overexpression, but the opposite was not observed in the case of FOXK2 knockdown. No significant change in the mRNA levels of any of the tested anticancer genes was observed upon knockdown of FOXM1, FOXO3 or FOXK2.

Lastly, I also attempted several experiments in trying to decipher the effects of the overexpression of TMEFF2 and FBLN5 on FOX proteins in MCF-7 cells. In short, results were inconsistent (data not shown) and this experimental avenue was not pursued any further.

4.3 Discussion

FBLN5 and TMEFF2 showed a dramatic increase in their relative mRNA expression levels upon overexpression of FOXO3-A3-always active mutant in MCF-7 cells. The mRNA levels of both were also upregulated proportionately upon overexpression of wild type FOXO3. Marina *et al.* [125] demonstrated that the PI3K/AKT pathway is constitutionally active in MCF-7 cells. In this study, when the PI3K/AKT pathway was inhibited, it led to cell death in MCF-7 cells but not in the MDA-MB-231 breast cancer cells. This indicates that the PI3K/AKT pathway is not only constitutionally active, but crucial for the survival of these cells. The constitutionally activated PI3K/AKT pathway would be expected to downregulate the functioning of FOXO3 (by phosphorylation at T32, S253 and S315). Moreover, FOXOs in general are largely regulated through post-translational modification rather than transcriptional modulation [119]. I included the transfection with FOXO3-A3-always active mutant in this experiment due to the concern that even the overexpressed wild type FOXO3 would be inactivated through phosphorylation by the constituently active PI3K/AKT pathway in MCF-7 cells. The results in figure 4.4(a) and 4.4(c) prove that these concerns were legitimate as both FBLN5 and TMEFF2 showed only a marginal increase upon transfection with wild type FOXO3, whereas when FOXO3-A3-always active mutant was overexpressed, the fold increase in mRNA levels of both genes was substantially large. These results are clearly suggestive of a direct or indirect regulation of TMEFF2 and FBLN5 in MCF-7 cells by FOXO3. One could argue that if these two genes are being positively regulated by FOXO3, a decrease in their mRNA expression levels should be expected upon knockdown of FOXO3 and/or overexpression of FOXM1. The absence of the former can be explained by the fact that, as FOXO3 is constitutionally inactive in MCF-7 cells, a further knockdown of an already inactive transcription factor cannot be expected to make a noticeable difference. It would only be upon the activation of a constituently inactive protein that a strong contrast in the

downstream effects would be observed. This (at least in part) also explains the inability of overexpression or knockdown of FOXM1 to impart any noticeable difference in the mRNA levels of FBLN5 and TMEFF2. In case these two genes are being regulated by FOXO3, the overexpression of FOXM1 would be expected to decrease their mRNA levels by repressing the activity of FOXO3. Again, the repression of an already inactive factor would not yield noticeable results in the downstream changes. When it comes to knockdown of FOXM1, it certainly should relieve the repression on FOXO3, leading to the upregulation of the putative downstream *TMEFF2* and *FBLN5* genes. However, the biological systems cannot be considered in isolation and FOXM1 is not the sole repressor of FOXO3 activity. Another factor that gives legitimacy to the suggested positive regulation of FBLN5 and TMEFF2 by FOXO3 is the biological reproducibility of the experiment. The relative fold increase in TMEFF2 and FBLN5 mRNA levels upon the overexpression of FOXO3-A3-always active mutant was both substantial and significant across all three independently transfected populations of MCF-7 cells. Similar was the case with the wild type FOXO3 and upon its overexpression, the increase in mRNA levels of both TMEFF2 and FBLN5 was observed across all the biological and technical replicates.

Considering the dramatic increase in the mRNA levels of TMEFF2 and FBLN5 upon the overexpression of FOXO3-A3-always active mutant, I attempted to verify this upregulation using the immunoblot. However, I was unable to detect TMEFF2 bands in either control or test conditions. For FBLN5, bands of negligible intensity were observed in both control and test conditions, but no change was detectable. As I explained earlier, the native expression levels of both of these genes in MCF-7 cells are very low and several unconventional steps were taken in order to get the PCR reactions with acceptable efficiency. As the sensitivity of the qRT-PCR assay is higher than immunoblotting by several orders of magnitude, it was not surprising not to be able to observe any change in the protein expressions of TMEFF2 or FBLN5, i.e., immunoblotting is probably not the appropriate assay (not sensitive enough) for detection of changes in the protein levels of TMEFF2 or FBLN5 in MCF-7 cells. Therefore, I am not showing the result of these experiments and reporting the facts only.

Overall, these results are indicative of FBLN5 and TMEFF2 being positively regulated by FOXO3. As FOXO3 is a bona fide tumour suppressor gene [104], the positive regulation of FBLN5 and TMEFF2 by a tumour suppressor transcription factor in MCF-7 cells is suggestive of these two genes having onco-suppressive activities in breast cancer.

5 TMEFF2 in Breast Cancer

5.1 Background

After deciphering the identities of anticancer genes and establishing the link between FOX proteins and two anticancer genes, I aimed to narrow down the further research to one anticancer gene and investigate the possible mechanisms behind its tumour specific activities. I decided to choose between TMEFF2 and FBLN5 because (i) both were protein coding genes with some knowledge available regarding their functions, (ii) the protein overexpression for both had been confirmed and working antibodies were available, (iii) both TMEFF2 and FBLN5 showed a significant correlation with the survival of breast cancer patients, thus giving an indication of an onco-suppressive activity, and (iv) both showed clear upregulation in response to FOXO3 overexpression.

As I have already mentioned, Qize Ding had reported that TMEFF2 and FBLN5 induced cell death in HEK293T and the CV-1 cells transformed through overexpression of oncogenes or mutated tumour suppressor genes (table 3.1). Among the oncogenes used for transformation of CV-1 cells, the one on which Qize Ding concentrated the most was cMYC. Ding had reported changes in cell morphology and growth characteristics in CV-1 cells upon stable transfection with cMYC [59]. There was no report, however, regarding the anchorage independent growth of these cMYC overexpressing CV-1 cells. As anchorage independent growth is a hallmark and a necessary trait of neoplastic transformation [1], I did not use these cMYC overexpressing CV-1 cells as a model for the transformed cells. Among other options, a relatively straightforward one was using the HEK293T cells as the model for the transformed cells, but apparently the large genetic differences between an adult African green monkey (*C. aethiops*) kidney cell line (CV-1) and the human embryonic kidney fibroblasts transformed using adenovirus DNA (HEK293T) [126, 127] would pose significant challenges while trying to decipher the differences in function(s) of the anticancer gene between the two systems. Providentially, a transformed version of CV-1 cells was already available in the form of the well-established cell line COS-7 cells. COS-7 cells were developed by Yakov Gluzman in Cold Spring Harbour Laboratory during the 1980s by transforming the CV-1 cells using the SV40 origin defective DNA [128]. Moreover, the ability of COS-7 cells to grow in soft agar had already been established in at least one study [129]. I therefore decided to use CV-1 cells as a model for the normal cells, COS-7 cells as a

model for the transformed cells and MCF-7 cells as a cancer cell model. To the best of my knowledge, the functions of neither TMEFF2 nor FBLN5 have been studied in COS-7 or MCF-7 cells previously.

The protein overexpression of both FBLN5 and TMEFF2 in MCF-7 cells had already been confirmed (figure 3.34). I shall report the functioning of TMEFF2 and FBLN5 in CV-1 and COS-7 cells in detail in the next chapter. For the purpose of background information for the results reported in this chapter, the reader is referred to figures 6.3 and 6.4 from chapter 6. I performed immunoblotting and qRT-PCR after transfecting FBLN5 and TMEFF2 in COS-7 and CV-1 cells at various time points. In comparison to COS-7 cells, the extent of TMEFF2 overexpression in CV-1 cells was very low. There could be two possibilities, (i) Qize Ding had identified TMEFF2 as an anticancer gene because of the inability of its plasmid to transfect successfully in CV-1 cells, or (ii) the biochemical and genetic peculiarities in CV-1 cells actively destabilize the mRNA of the overexpressed TMEFF2. The former is less probable because, firstly, FBLN5 was in the same plasmid as TMEFF2 (and even larger in size) but still showed proportional levels of overexpression in both CV-1 and COS-7 cells (figure 6.6). Secondly, Qize Ding had successfully optimized the transfection efficiency in CV-1 cells (as reported in his thesis) [59]. As I have mentioned in the introduction, the suppression of a cell-death inducing gene in normal cells and the alleviation of this suppression in the transformed cancer cells should still fall under the umbrella of synthetic lethality. Nonetheless, owing to the aims my of research, it was preferable to study the differences in the downstream effects of anticancer genes between the normal, transformed and cancer cells rather than investigating the possible phenomenon of active gene suppression by normal cells versus the opposite in the transformed and cancer cells. I therefore decided to choose FBLN5 for further studies.

However, the degree of cell death induced by FBLN5 in MCF-7 cells varied from one experiment to another (supplementary figure S.2). Consequently, I chose to study the functioning of TMEFF2 in MCF-7 cells and that of FBLN5 in CV-1 and COS-7 cells. This chapter focuses on the investigation of TMEFF2 function(s) in MCF-7 cells.

5.1.1 TMEFF2, an Overview

TMEFF2 (transmembrane protein with an EGF-like and two Follistatin-like domains 2) is an 1,814-bp gene encoding, 374-residue long single polypeptide, type-I transmembrane proteoglycan. The HUGO Gene Nomenclature Committee [130, 131] designates this gene with the name TMEFF2 and mentions

TENB2, HPP1, TR, TPEF and CT120.2 as alias symbols and 'transmembrane protein TENB2', 'tomoregulin' and the 'cancer/testis antigen family 120, member 2' as the alias names. The N-terminal of TMEFF2 harbours a signal peptide marking it for secretion; it is followed by two follistatin-like domains, an EGF (epidermal growth factor)-like domain, a transmembrane portion and a short intracellular domain (figure 5.1). The EGF-like domain of TMEFF2 appears to be functionally ineffective because of the substitution of a crucial arginine residue (Arg39) with histidine [132]. At least three studies [133, 134, 135] have reported the EGF domain of TMEFF2 to be dispensable for the proposed functions of this protein. Moreover, the same studies also reported that the follistatin-like domains were crucial for the proper functionality of TMEFF2. The roles of the transmembrane domain, stalk region and the intracellular domain in the function(s) of TMEFF2 are particularly interesting. These are discussed in detail later but, very briefly, TMEFF2 is shedded from the surface of the plasma membrane through the action of membrane bound proteases (in a manner similar to that of TNFα).

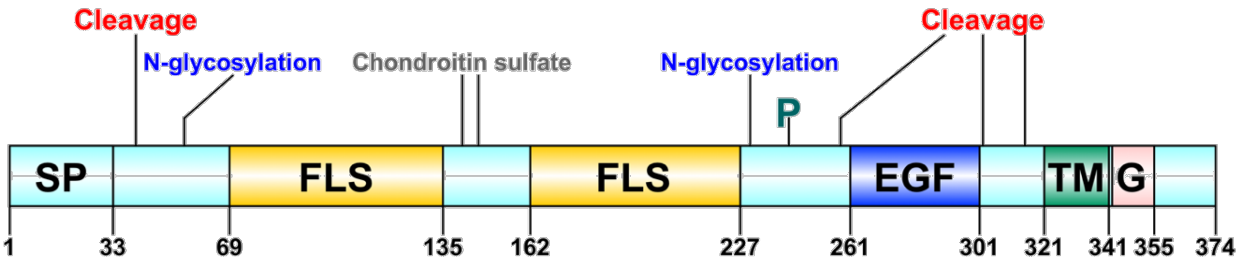


***Figure 5.1*: TMEFF2 protein features**. SP: signal peptide. FLS: follistatin-like domain. EGF: EGF-like domain. TM: transmembrane domain. G: G-protein activating motif. P: phosphorylation site. A serine/glycine rich region between the two follistatin-like domains constitutes a glycosaminoglycans (Chondroitin sulfate) attachment site. A potential tyrosine kinase phosphorylation site at residue 242, N-linked glycosylation sites around residues 55 and 203 and several potential protease cleavage sites (only sites reported through hydropathy analysis are represented). Cartoon generated using DOG v.2.0 [136].

The reported functions of TMEFF2 span across a wide range of the physiological and pathological spectra, ranging from corticotropin release hormone (CRH) stimulation in the anterior pituitary gland and providing neuroprotection by binding amyloid beta protein derivatives in Alzheimer's disease on one hand, to triggering the JAK-STAT pathway in colorectal cancer and modulating one-carbon

metabolism in prostate tissue, on the other. These are summarized in table 5.1. Moreover, *TMEFF2* promoter and 5' upstream region harbour a CpG island [137] which is methylated in a wide variety of cancers. The chromosomal location of TMEFF2 is mapped to 2q32-33 [137, 138]. The literature search reveals unclear and sometimes opposing information regarding the function, the tissue expression and intracellular localization of TMEFF2. A considerable portion of this chapter is devoted to a detailed literature review of TMEFF2 because it would be impossible to unpick the super-entangled web of opposing information regarding this very interesting and rather intriguing gene otherwise.

Information regarding TMEFF2 can be broadly categorized into:

I. Identification and characterization.
II. Methylation in cancers.
III. Tissue distribution.
IV. Proposed functions.
V. The role in prostate cancer.
VI. The role in gastric and colorectal cancers.

5.1.1.1 Identification and characterization

TMEFF2 was discovered and characterized by at least five independent groups within the time span of roughly a year [132, 137, 138, 139, 140]. Later, another group reported the discovery of a shorter secreted isoform of TMEFF2: TMEFF2-S [141].

In April 2000, in an attempt to identify novel EGF-like genes, Horie *et al.* [132] searched for the sequences encoding the conserved EGF-like domain using their in-house database and identified a novel gene in the human brain. Using a 704-bp probe which included the 5'-region of this novel sequence, they screened through a λ10 human foetal brain cDNA library and cloned this gene out. The predicted amino acid sequence revealed an N-terminal signal peptide, two follistatin-like domains upstream of the EGF-like domain, a putative transmembrane domain and a short C-terminal cytoplasmic domain (figure 5.1). The gene was thus named TMEFF2 (transmembrane protein with an EGF-like and two Follistatin-like domains 2). Later the group isolated *Tmeff2* from a LambdaZAPII mouse brain cDNA library. The predicted amino acid sequences for human and mouse TMEFF2 were 98.9% identical and showed a high degree of homology with the *X. laevis* protein x7365 as well. A human clone H7365 was found to be sharing 35.8% homology with *TMEFF2* and was named TMEFF1.

The EGF-like domain is typically characterised by the consensus motif CX7CX3-5CX10-12CXCX5GXRC, where the spacing of cysteine residues is maintained (C: cysteine, G: glycine, R: arginine, X: other amino acids) [140]. According to Horie *et al.* [132], the glycine and cysteine residues typical of EGF-like domains are conserved in both TMEFF1 and TMEFF2; however, a critical arginine residue, Arg39, which is crucial for the binding of EGF-like domain to its receptors, is replaced with histidine in both TMEFF1 and TMEFF2. The mutation of this arginine in human EGF reduces the affinity of EGF for its receptor to less than 0.5% of normal EGF [132]. Yonezawa *et al.*[139], on the other hand, reported TMEFF2 to mildly activate the HER4 receptor in MKN28 gastric cancer cells.

Similar to Horie *et al.*, Yonezawa *et al.* [139] (study initially published in December 1999 [142] and republished in 2000) cloned a gene from a human brain cDNA library while searching for novel proteins with the EGF-like domain. The probe used to screen this gene was constructed from cDNA extracted from stomach fibroblasts. They named the gene tomoregulin and reported three splice variants: TRa (368 a.a), TRb (418 a.a) and TRc (379 a.a). Upon subjecting TRa to hydropathy analysis (for its predicted amino acid sequence), in addition to the protein structure described previously, the analysis revealed a potential G-protein activating motif on the C-terminus and three potential protease cleavage sites – two adjacent to the transmembrane domain in the stalk sequence and another one towards the N-terminal portion of EGF-like domain. Their hydropathy analysis also revealed two potential sites for N-linked glycosylation and two potential glycosaminoglycan attachment sites on the ectodomain. The group later verified the N-linked glycosylation by treating the TRc stable CHO cells with tunicamycin (which prevents N-linked glycosylation), which resulted in a 6 kDa decrease in the protein size. They also demonstrated the presence of the protease cleavage sites on both N- and C-terminal portions of EGF-like domain.

Later in October 2001, Glynn-Jones *et al.* [140] reported the discovery and characterization of TENB2 while attempting to identify differentially expressed genes between androgen-dependent and androgen-independent stages of prostate cancer. The TEN12 prostate cancer xenograft model was employed for this study. The parent clone of the TEN12 xenograft is androgen-dependent. The group developed the androgen-independent TEN12-F and TEN-12-C by passaging the parent TEN12 through castrated and female mice. λcDNA libraries from TEN12 and TEN12-F were developed and upon sequencing, TENB2 was found among the genes which were consistently upregulated upon development of androgen independence in the TEN12 prostate cancer xenograft. Interestingly, the authors described TENB2 as a splice variant of tomoregulin stating [140], '*all three isomers of*

tomoregulin share 100% sequence identity with TENB2 over the 5' coding region (bp 316–1,473) but not over the 3'-end of the sequence (bp 1,474–1,814) encoding the C-terminal end of the putative cytoplasmic domain and the 3'-untranslated region' (p.182). Importantly, they reported that the isoform expressed in prostate tissue is exclusively TENB2 and not tomoregulin. However, Zhao *et al.*[143] later reported tomoregulin to be present in prostate cancer and normal prostate tissue. In addition to the traits designated to this protein in the two previously mentioned studies, Glynn-Jones *et al.* suggested a potential tyrosine kinase phosphorylation site preceding the EGF-like domain at residue 242 and an additional potential protease cleavage site at residues 40/41. Moreover, the group reported the two follistatin-like domains in TMEFF2 to exhibit 40-50% sequence similarity with the follistatin-like domains of other proteins. Glynn-Jones *et al.* [140] also demonstrated that TENB2 protein was heavily glycosylated with both N-linked short oligosaccharides and O-linked long chondroitin sulfate (bimeric repeat polymer) glycosaminoglycans.

Around the same time (September 2000), Liang *et al.* [137] identified TPEF (transmembrane protein containing EGF and follistatin domains) during a search for the differentially methylated fragments in colon, bladder and prostate cancers [144]. TPEF was mapped to human chromosome 2q33. Other salient features reported were: (i) the promoter of *TPEF* is TATA-less and hosts a CpG island predicted to be -650 to +1 bp from transcription initiation site, (ii) the most CpG rich region is -1100 to +750 from transcription initiation site, and (iii) the 3'-UTR includes a destabilizing ATTTA motif and two potential polyadenylation signals AATAAA followed by a polyadenylic acid stretch.

In December 2000, Young *et al.* [138] identified a novel gene named HPP1 that was differentially hypermethylated in colonic neoplasms and the premalignant hyperplastic polyps of colon. The gene was later cloned out using 3'-RACE (rapid amplification of cDNA ends) and showed 95% homology with TMEFF2, tomoregulin, TENB2 and TPEF. Employing fluorescent *in situ* hybridization, HPP1 was found to be located on chromosome 2 in region 2q32-q33. A second isoform of HPP1 called HPP1-B was also reported in the same study. HPP1-B had an early stop codon which could cut the protein short of cytoplasmic tail. Interestingly, HPP1-B was only found in normal mucosa and not in the cancerous tissues but, intriguingly, investigators were unable to isolate tomoregulin from pooled normal mucosa by 3'-RACE.

Lastly, in 2006 Quayle *et al.* [141] reported a short isoform of TMEFF2 (TMEFF2-S) in the LNCaP prostate cancer cell line. The TMEFF2-S is comprised of the first four exons of TMEFF2 followed by 2.9

kb sequence of the fourth intron. The predicted amino acid sequence suggests a 175-amino acid ORF consisting of the first 146 residues of TMEFF2 followed by a novel sequence of 29, mostly hydrophobic residues. TMEFF2-S lacked one follistatin-like domain, the EGF-like domain, and the transmembrane domain. As the secretion signal peptide was still intact, TMEFF2 was found to be secreted in the medium of TMEFF2-S-GFP stably transfected LNCaP cells. Moreover, the TMEFF-S-GFP exhibited a diffused cytoplasmic profile when examined by fluorescence microscopy as compared to the full-length TMEFF2-GFP, which was mostly localized at the cell membrane of LNCaP cells. It is important to note that TMEFF2 was not observed in the media of LNCaP cells expressing the full-length TMEFF2-GFP fusion construct, suggesting that TMEFF2 is not a secretory protein on its own.

5.1.1.2 *TMEFF2* promoter methylation in cancers

Accumulation of multiple genetic mutations is the requirement and a hallmark of neoplastic transformation. Various mechanisms that the cancerous cells employ to acquire these mutations include copy number variations (CNV), aneuploidy, chromosomal rearrangements and, especially pertaining to tumour suppressor genes, loss of heterozygosity (LOH). In addition to the above-mentioned canonical mechanisms, epigenetic alterations have recently attracted focus as alternative means of acquiring and maintaining the desired traits by cancer cells [145].

As mentioned earlier, *TMEFF2* harbours a CpG island in its promoter region. The CpG rich region spans across -650 to +1 bp from transcription initiation site [137]. During the HPP1 identification and characterization study by Young *et al.* [138], RT-PCR revealed that HPP1 is expressed in 28/30 samples in normal mucosa but only 7/30 samples in cancer tissues. In the same study, LOH analysis revealed that only 1/22 colorectal cancers, 3/6 adenomas, 1/22 hyperplastic polyps exhibited HPP1 LOH. On the other hand, more than 30% DNA methylation was observed in 46/55 colorectal cancers, 6/9 adenomas and 17/27 hyperplastic polyps on the 49 CpG sites which spanned 5'-UTR and first exon of HPP1. A significant inverse correlation was also observed between the levels of promoter methylation and mRNA expression levels of HPP1. Indeed, upon treatment of LoVo, HCT116 [146] and HT29 adenocarcinoma cells with 5-Aza-2-Deoxycytidine (a DNA methyltransferase inhibitor), HPP1 mRNA expression was restored, indicating that DNA methylation rather than LOH is the main contributing factor in silencing HPP1 in colorectal cancer. Similarly, TPEF was found to be heavily methylated in 9 out of 11 cell lines tested by Liang *et al.* and the mRNA expression also matched the methylation status of TPEF in these cells. HCT15, LoVo, HCT 116, SW480, SW837, HT29, 3657 and DU145 cells

harboured heavily methylated TPEF (> 80%) and expressed almost no mRNA for the gene. SK-Mel-28 and LNCaP cells, on the other hand, showed less methylation (< 20%) and expressed detectable TPEF mRNA levels [137]. A more recent study by Lee *et al.* [147] reported that high levels of TMEFF2 mRNA were expressed by the cell lines with unmethylated alleles (H522, H1793, H2009, and H1703), while its expression was remarkably reduced in the cell lines where *TMEFF2* promoter was methylated (A549, H157, H226, and H1299). Again, treatment of the non-expressing cells with 5-Aza-2-Deoxycytidine was able to significantly restore TMEFF2 mRNA expression. An important confirmation of methylation being the primary source of TMEFF2 gene silencing came from a gene-wide search by Young *et al.* for the somatic and germline mutations in colorectal cancer [148]. The study revealed that there were no pathogenic mutations in the HPP1 gene when 63 tumours and the corresponding normal tissues were scanned for sequence variations.

When compared to the adjacent normal tissue or the corresponding tissue from the normal individuals, a considerable number of primary publications have reported the differential *TMEFF2* methylation in a variety of cancers including breast [149, 150, 151, 152, 153, 154], prostate [137, 153], lung [147, 150, 153, 155], bladder [137], colon and rectal [137, 138, 146, 150, 154, 155, 156, 157], gallbladder [158], renal [159], oesophageal [160], cardiac [160], stomach/gastric [103, 160, 161], ovarian [155], multiple myeloma [153], glioblastoma [155] and mesothelioma [162]. Several studies [103, 137, 138, 146, 147, 155, 161] among these are indicative of a direct correlation between TMEFF2 mRNA downregulation and its promoter methylation and few also reported promoter methylation to be related to the stages, response to therapy and survival outcomes of the corresponding cancers [149, 150, 153, 155, 161].

In breast cancer, *TMEFF2* promoter methylation increased progressively along the potential developmental path of invasive ductal carcinoma (IDC) from the pre-invasive lesions like flat epithelial atypia (FEA), atypical ductal hyperplasia (ADH), and ductal carcinoma *in situ* (DCIS) [151, 152]. Similar was the case in colorectal cancer where Young *et al.* [138] reported HPP1 to be differentially hypermethylated in 84% of colon carcinomas, 66% of adenomas and 63% of premalignant hyperplastic polyps, again showing a progressive methylation profile during the putative development path of colorectal cancer [138]. In case of ulcerative colitis-associated colorectal cancer, *TMEFF2* methylation was observed in 40% of dysplasia and 50% of carcinomas but no methylation was detected in the non-neoplastic mucosae [146]. Sun *et al.* [163] presented similar findings in gastric cancer where TMEFF2 showed higher expression in 44.6% of normal gastric epithelium, 29.2%

of intestinal metaplasia, 18.5% of dysplasia and in 7.7% of gastric cancer. The group had already established an inverse correlation between TMEFF2 expression and its promoter methylation along the developmental stages of gastric cancer in an earlier study [103]. They also reported the lower expression levels of TMEFF2 to be directly related to poor survivability in gastric cancer. Together these observations point towards TMEFF2 having onco-suppressive properties.

Another very interesting observation from the data provided by Know *et al.* [151] was that *TMEFF2* methylation was higher in the normal tissue surrounding invasive breast cancer tissue from cancer patients as compared to the normal tissue from non-cancer patients. Similarly, Sabbinioni *et al.* [164] reported that *TMEFF2* was not only methylated in colorectal cancer and the premalignant lesions but also the tissue from the area around the premalignant lesions (adenomas). These studies clearly indicate *TMEFF2* methylation and subsequent silencing as an early event during the development of breast and colorectal cancers.

Promoter DNA methylation of *TMEFF2* in breast and colorectal cancers is worth a deeper look. For example, in a small cohort of 110 tumours and 21 normal breast tissue samples, *TMEFF2* promoter was found to be hypermethylated in recurrent breast cancer compared to non-recurrent cancer (and the normal breast tissue) [149]. In a subsequent study, the same group [150] proposed *TMEFF2* methylation as a biomarker for the diagnosis of breast cancer in patients' serum; TMEFF2 was found to be among the top 10 out of 2674 genes differentially hypermethylated in the serum of breast cancer patients when compared to the sera from the normal individuals. Moreover, *TMEFF2* methylation in sera was reported to be at the same level as in the metastatic breast cancer lesions, i.e., the metastatic tumours in various parts of the body maintained the methylation state of *TMEFF2*. In colorectal cancer, Sabbinioni *et al.* [164] proposed the detection of methylated *TMEFF2* in blood as a diagnostic marker to identify patients with advanced colorectal cancer as well as with early tumorous lesions. This implies that TMEFF2 not only has prognostic value, but its methylation can also be used as a diagnostic tool.

The epigenetic silencing of TMEFF2 is not limited to its promoter methylation. Want *et al.* [165] demonstrated restoration of HPP1 expression in HCT116 and DLD-1 colorectal cell lines upon treatment with histone deacetylase (HDAC) inhibitors. Treatment of cells with HDAC siRNAs also resulted in HPP1 re-expression. Histone deacetylation is an epigenetic mechanism of gene silencing where chromatin remodelling leads to repression of the target genes. *HPP1* promoter region harbours

an E-box that is recognized by cMYC (a known repressor of HPP1). Employing co-immunoprecipitation (IP), chromatin IP (ChIP), and sequential ChIP experiments, Want *et al.* [165] demonstrated that cMYC and HDAC3 colocalized with the *HPP1* promoter region. They proposed a model where cMYC recruits HDAC3 to silence the HPP1 expression in colorectal cancer cells.

Compared to other methods of acquiring and maintaining the desired traits by cancer cells, epigenetic alterations present a unique opportunity in cancer therapy as these can be reversed. Being epigenetically silenced in a multitude of cancers and having both prognostic and diagnostic value, TMEFF2 not only presents itself as a possible onco-suppressor, but the gene also deserves further attention as a potential therapeutic target.

It is also worth mentioning that *TMEFF2* promoter methylation was neither related to the stage of breast cancer, nor to its p53 status or Ki64 levels [151]. Moreover, compared to other genes included in these studies, *TMEFF2* methylation levels were relatively higher even in the normal breast tissue, which brings us to the tissue distribution of TMEFF2.

5.1.1.3 Tissue Distribution

TMEFF2 has probably not attracted the attention it deserves because of an apparently limited tissue distribution. Most of the studies reported a baseline TMEFF2 expression exclusively in brain and prostate tissue [132, 139, 140, 166, 167], but several contradictions exist. For example, Glynne-Jones *et al.* [140] used multi-tissue northern arrays and reported TENB2 to be present only in prostate tissue and brain (adult and foetal) except in the pituitary gland; Liang *et al.* [137] used similar tissue arrays and depicted TPEF to be predominantly present only in brain and prostate tissues, but a lower expression was present in a variety of other tissues including stomach, colon, testis, ovary, pancreases, adrenal gland, thyroid gland, salivary gland, liver, small intestine, lung, trachea, placenta, foetal brain, foetal heart, foetal kidney, bladder, mammary gland, kidney and pituitary gland. In addition, Labeur *et al.* [134] found ACTH-producing cells of the anterior pituitary to be TMEFF2 immuno-positive. Young *et al.* [138] employed *in situ* hybridization in colonic mucosae and HPP1 was observed in myofibroblasts, stromal cells, epithelial cells, B-lymphocytes, gut-associated lymphoid tissue, scattered fibroblasts, ganglion cells and endothelial cells. Using RT-PCR Zhao *et al.* [143] observed low levels of tomoregulin mRNA in kidney tissue and minimal levels in testis but these could not be confirmed through immunohistochemistry. The story is similar when it comes to cell lines: the majority of studies report no expression of TMEFF2 in most cell lines except for a few from prostate

and nervous tissue origin [137, 139, 140, 166, 167]. A few important exceptions exist here as well. For example, Grey *et al.* [166] found a low level of expression of TMEFF2 in HOS and MG63 osteosarcoma cells. Yonezwa *et al.* [139] reported weak expression in stomach fibroblasts, RD human embryonal rhabdomyosarcoma cells and PC12 rat pheochromocytoma cells in addition to the A172 glioblastoma and CHO-K1 cells. Labeur *et al.* [134] used AtT20 pituitary corticotrope tumour cells for TMEFF2 knockdown studies. Moreover, Torrecilla *et al.* [168] cloned *TMEFF2* by retrotranscription from total RNA of the human lung carcinoma-derived cell line A549, and Liang *et al.* [137] used RNA from human placental cDNA library. They also observed expression of the gene in the SK-Mel-28 human melanoma cell line pointing towards a more widespread tissue expression.

Interestingly, although several studies have proven the expression of TMEFF2 in prostate tissue [137, 140, 143, 166, 167], when Zhao *et al.*[143] tested for the expression of tomoregulin mRNA using RT-PCR and Northern Blot in five prostate cell lines (PrEC (primary culture of human normal prostatic epithelial cells), BPH1 (T-antigen-immortalized benign prostatic hyperplasia cells), and the human prostate cancer cell lines LNCaP, PC-3, and DU145), only LNCaP cells expressed significant levels of tomoregulin mRNA. The expression of TMEFF2 in LNCaP cells and its absence in PC3 and DU145 were confirmed by Gery *et al.* [166] who also reported TMEFF2 mRNA expression in the CWR22 prostate cancer xenograft. The most comprehensive study regarding the normal and diseased tissue and cell line expression of TMEFF2 came from Afar *et al.* [167]. They used custom microarrays followed by the TaqMan™ analysis confirmation and reported TMEFF2 expression to be exclusive in LNCaP and CWR22 prostate cancer cell lines in addition to the glioblastoma cell lines U118 and U187MG. However, a closer observation of the data provided by Afar *et al.* reveals a very low level of expression of TMEFF2 in breast, colon and lung cancer cell lines as well. It appears that the relative expression levels of TMEFF2 mRNA in brain and prostate tissues are so high that they mask the data from other tissues in comparison. This premise is supported by the fact that upon lowering the expression threshold of the tissue expression data for TMEFF2 on the TGCA expression atlas [169], a low to moderate baseline expression of TMEFF2 in numerous human tissues was revealed – including breast (at least in two projects *vis-à-vis* Genotype-Tissue Expression GTEx Project [170] and the Illumina Body Map [171]). Similar was the case when RNA-Seq data from the 'Human Protein Atlas' [77, 78, 79] was checked. In addition, the Human Protein Atlas pathology data revealed that the breast tissue was not only positive for TMEFF2 mRNA but also its protein product (through immunohistochemistry).

Briefly, among other species a similar TMEFF2 pattern of expression is reported, i.e., mainly in brain and nervous tissue [132, 143, 172, 173] with a few interesting exceptions. Yonezawa *et al.* [139, 142] observed tomoregulin immunoreactivity in gastric mucosa, mesenchymal cells and fibroblasts localized in the lamina propria of rat. Chen *et al.* [174] reported Tmeff2 expression in the white adipose tissue of mouse. They also registered significant *Tmeff2* expression in the mouse prostate using PALP-staining in *Tmeff2* knockout mice where the first coding exon of *Tmeff2* was replaced with cDNA encoding the human placental alkaline phosphatase (hPLAP). However, when Afar *et al.* [167] used immunostaining with antibodies for the TMEFF2 ectodomain, they found mouse prostate to be negative for Tmeff2. The absence of Tmeff2 in mouse prostate was also reported by Corbin *et al.* [172].

5.1.1.4 Proposed functions

As mentioned earlier, the physiological role of TMEFF2 remains largely elusive. The conflicting reports regarding growth-promoting versus growth-retarding effects add to the problem. Tomoregulin mRNA was reported to be expressed in mouse embryo at E-11 (earliest tested) and increased gradually till E-15 and stabilized through to E-17 [139], pointing towards a possible role in development. The Tmeff2 knockout mice were born normal, maintained the Mendelian ratio and did not show any structural defects. However, they were smaller in size, failed to gain weight and died around weaning age of approximately 3 weeks [172, 174].

The conserved EGF-like domain and the two follistatin-like domains have attracted the most attention. EGF was the first growth factor ever to be discovered [1] (by Dr. Stanley Cohen in 1962 [175]). The EGF-like protein family is comprised of several growth factors, including transforming growth factor alpha (TGF-α), neuronal growth factor (NRG), amphiregulin (AR), epiregulin (EPR), betacellulin (BTC), etc. All of these proteins can bind and activate epidermal growth factor receptors (EGFRs). The EGFR, more commonly known as HER1 or erbB1, belongs to the EGFR/ErbB superfamily which includes three other receptor tyrosine kinases: ErbB2/Neu/HER2, ErbB3/HER3, and ErbB4/HER4. EGFRs are expressed on the surface of a variety of tissues and play vital roles during embryonic development, differentiation and tissue regeneration and repair [175]. Being pleotropic developmental proteins, it is not surprising that they are involved in transformation, development and sustenance of several cancers [176].

Although the replacement of Arg39 in EGF-like domain of TMEFF2 gravely reduces its affinity for EGFRs, Yonezawa *et al.* [139] observed that TMEFF2 mildly activated HER-4 receptors in MKN28 gastric cancer cells. Moreover, Ali and Knauper [133] reported that the addition of soluble TMEFF2 ectodomain to HEK293T cells led to ERK1/2 phosphorylation and that this effect vanished dramatically upon treating the cells with the HER1 inhibitor AG1478. On the other hand, the same study also reported that the treatment of HEK293T cells with a TMEFF2 mutant lacking the EGF-like domain could still increase cellular proliferation without any loss of this effect. At least two more studies have reported the same, i.e., the deletion of EGF-like domain of TMEFF2 had no effect on its reported function [134, 135].

Turning the attention to the follistatin-like domain, follistatin was initially reported as an activin binding protein in the developing oocyte [177]. Activin is one of the 33 members of the transforming growth factor beta (TGF-β) family known mainly for its role in augmenting the gonadotropin-releasing hormone (GnRH) mediated release of follicle stimulating hormone (FSH) from the anterior pituitary. It is worth mentioning here that TGF-β family does not exclusively comprise of growth factors: inhibin, which binds and neutralizes activin in a negative feedback loop, also belongs to the TGF-β family [178]. The role of TGF-β during tumorigenesis is very interesting as it suppresses the proliferation of normal cells but promotes the acquisition of invasive properties by the transformed cells [1]. Follistatin was originally known to be a non TGF-β family member which could bind and inhibit the actions of activin in a fashion similar to that of inhibin. Later, it was discovered that follistatin is expressed by a number of tissues throughout the body and also binds other members of the TGF-β family. Also, it plays a crucial part in the embryogenesis and follistatin-null mice are smaller in size and die soon after birth due to impaired breathing [177]. However, the role of the follistatin-like domain in the follistatin module-containing proteins is not clear. One follistatin module-containing protein (osteonectin) has been reported to bind and neutralize both platelet derived growth factors (PDGFs) and vascular endothelial growth factor (VEGF) [155]. Similar to osteonectin, Lin *et al.* [155] reported that TMEFF2 could physically bind PDGF-AA both in the solution and on cell surface. They also showed this binding to prevent PDGF-AA from inducing cellular proliferation in NR6 fibroblasts. Also, the ability of TMEFF2 to prevent PDGF-AA induced cellular proliferation was abolished upon deletion of its follistatin-like domains.

Huang *et al.* [179] likewise reported binding of TMEFF2 with PDGF-AA in HEK293T cells. In this very interesting study, Tmeff2 was found to be specifically upregulated in cells of oligodendrocyte lineage

during the differentiation stage. As PDGF-AA had already been shown to prevent this differentiation, it was hypothesized that TMEFF2 could induce oligodendrocyte differentiation by binding and neutralizing PDGF-AA. But this hypothesis was proven wrong as the *Tmeff2* knockout mice did not exhibit any defects in oligodendrocyte differentiation. Also, the addition of TMEFF2 to the media containing HEK293T cells could not prevent the PDGF-AA induced phosphorylation of ERK1/2. On the other hand, TMEFF2 overexpression in HEK293T cells itself caused ERK1/2 phosphorylation, pointing towards a pro-proliferative role. Similarly, in another study [132], the addition of recombinant TMEFF2 ectodomain to primary cultures of hippocampal and mesencephalic neurons was also reported to enhance their proliferation in a dose-dependent manner.

On the other hand, TMEFF2 expression was not only downregulated in glioma but its mRNA levels decreased with increasing grade and worsening treatment outcomes [155]. In support of the onco-suppressive role of TMEFF2, Labeur *et al.* [134] reported that TMEFF2 protein levels decreased in cotricotropinomas in comparison to the normal tissue. TMEFF2 overexpression also caused a marked decrease in proliferation of AtT20 human glioma cells when tested through both short term (WST-1 assay) and long term (soft agar colony formation) cell proliferation assays. Furthermore, TMEFF2 inhibited the production of adrenocorticotropic hormone (ACTH) in response to CRH (corticotropin releasing hormone) stimulation in AtT20 glioma cells; this action was elaborated to be the result of TMEFF2-imparted action at multiple levels of CRHR1 signalling. Notably, while the EGF-like domain, the intracellular domain and transmembrane domains were dispensable, the follistatin modules were essential for the ability of TMEFF2 to impart the above-mentioned effects.

Another very interesting role of TMEFF2 in the nervous tissue was reported by Hyun-Seok *et al.* [135]: TMEFF2 could bind Amyloid-β protein oligomers (AβO) with high affinity and protected the N2a cells from the neurotoxic effects of AβO. Amyloid-β protein (Aβ) self-aggregates into amyloid fibrils and causes amyloid plaque formation which is a hallmark of Alzheimer's disease. Amyloid-β protein is derived from the proteolytic processing of the amyloid-β protein precursor (AβPP). In addition to fibrillar forms, Amyloid-β protein exists in various smaller assemblies, called amyloid-β oligomers (AβO). TMEFF2 binding was not limited to only the oligomers (AβO) but it (TMEFF2) could also directly bind the precursor AβPP and mature Aβ. Moreover, this binding was observed in solution, intracellularly as well as in the mouse brain. The immunostaining revealed that Tmeff2 was upregulated in neurons exhibiting excessive amyloid-β intraneuronal accumulation, therefore suggesting a neuroprotective role of the protein *in vivo*. Considerable TMEFF2 immunostaining was

also observed in astrocytes and *in vitro*, TMEFF2 showed a granular cytoplasmic profile rather than a membranous localization in MC6 neuroblastoma cells. When a mutant form of TMEFF2 lacking the EGF domain was used, it did not have any effect on the neuroprotective role of TMEFF2.

5.1.1.5 TMEFF2 in prostate cancer

A substantial proportion of literature regarding TMEFF2 focuses on its role in prostate cancer. Such studies greatly enhance the knowledge about TMEFF2, but at the same time, they add to the confusion about its function, the oncogenic *vs* onco-suppressive nature of TMEFF2, the range of tissue expression as well as the possibility of TMEFF2 being a biomarker and/or therapeutic target. Several groups have attempted to solve these dilemmas and they have given some fascinating insights which I shall discuss later.

Many studies on TMEFF2 in prostate cancer revolve around the progression of prostate cancer from androgen-dependent to androgen-independent stages. As I have mentioned earlier, Glynn-Jones *et al.* [140] identified TENB2 while studying the differential gene expression between androgen-dependent and androgen-independent prostate cancer xenograft models. In addition, they also reported that TMEFF2 expression was markedly increased in prostate carcinoma *vs* benign prostate hyperplasia and that its mRNA levels increased in higher grade prostate cancers. These findings point towards an oncogenic role of TMEFF2 in prostate cancer. Such inference is reinforced by several subsequent studies: for example, Bhaskar *et al.* [167] reported that both protein and mRNA levels of TMEFF2 increased significantly in prostate cancer compared to normal tissue. The group also developed an antibody-drug conjugate (ADC) called Pr1-vcMMAE (using an in-house raised monoclonal antibody (mAb) against TMEFF2 ectodomain). Compared to the control ADC, Pr1-vcMMAE effectively reduced the tumour size in LNCap prostate cancer xenografts with minimal toxicity. The humanized version of ADC, hPr1-vcMMAE, gave similar results when used against the CWR22 prostate cancer xenografts. In what appears to be a follow-up study, Zhao *et al.* [143] used 2H8 tomoregulin mAb (developed by Uchida *et al.* [142]) radiolabelled with β-emitting isotope ^{90}Y and reported it to impart a significant reduction in the LNCaP xenografts (compared to radiolabelled IgG1). The chemotherapy was replaced by radiotherapy under the pretext that the TMEFF2 expression in the human prostate cancer tissue was found to be very heterogeneous (both within and among samples) [143]. This might impair the effectivity of an antibody drug conjugate because the ADC needs to be bound to the cell surface epitope and get internalized for successful action, whereas

a radiolabelled isotope would give a 'bystander effect', thus killing the cells in the immediate vicinity as well. A substantial upregulation of tomoregulin mRNA and protein in prostate cancer tissue compared to normal tissue was also reported in this study. The injection of TMEFF2 mAb on its own did not have any effect on xenografts and there was no difference in proliferation observed in tomoregulin-stable PC3 prostate cancer cells, nor was there any difference in proliferation observed in LNCap or PC3 cells when the extracellular domain of tomoregulin was added to the cell culture. Similarly, Zhao *et al.* [180] complexed the TMEFF2 mAb (2H8) with a goat anti-mouse IgG antibody linked to the saporin toxin (Mab-ZAP). When added to TMEFF2 overexpressing PC3 prostate cancer cells (TMEFF2-PC3), Mab-ZAP induced a 50% reduction in the proliferation of TMEFF2-PC3 cells while the native PC3 cells remained unaffected. Investigators attribute this efficacy to the ability of TMEFF2 mAb to get internalized upon binding the cell surface epitope. 2H8 mAb was also reported to localize in lysosomes after internalization. Notably, TMEFF2 does not contain the typical internalization motifs found in type I and II receptor proteins which trigger their normal turnover through lysosomal degradation. In order to rule out the possibility that the TMEFF2 internalization is attributed to 2H8 antibody, additional TMEFF2 antibodies were used. Similar internalization and localization profiles were observed, which points to the possibility that TMEFF2 may contain a novel signal for cell-surface protein internalization. Again, TMEFF2-stable PC3 cells showed no difference in cell proliferation compared to the control PC3 cells. It is worth mentioning here that in order to be an effective target for ADC, the epitope has to be (i) specific for prostate tissue, (ii) upregulated in prostate cancer, (iii) minimally expressed, or not at all, in vital organs, (iv) absent in blood and (v) membrane bound and internalized upon ADC binding. The above-mentioned studies point towards TMEFF2 being an oncogene, the expression of which is increased in prostate cancer *vs* normal tissue and increased further upon developing androgen independence. Moreover, ectopic overexpression of TMEFF2 had no detrimental effect whatsoever on the proliferation of prostate cancer cells in culture or in xenografts.

Several other studies tell an altogether opposite story. Gery *et al.* [166] reported that TMEFF2 was expressed in androgen-dependent LNCap cells but not in androgen-independent DU145, and PC3 prostate cancer cells. Moreover, reduced TMEFF2 mRNA levels were observed in the androgen-independent prostate cancer xenograft models LAPC9-AI, LAPC3-AI and LAPC4-AI when compared to their androgen-dependent counterparts. Also, upon treatment of LNCaP cells with dihydrotestosterone, TMEFF2 expression increased in both time and dose-dependent manners. In

addition, when an androgen-dependent prostate cancer xenograft model, CWR22, was propagated in castrated male mice, TMEFF2 expression vanished. Importantly, when TMEFF2 was ectopically expressed in DU145 and PC3 cells, it inhibited their proliferation significantly – although TMEFF2 was upregulated by androgens, it exhibited antiproliferative effects in prostate cancer.

The regulation of TMEFF2 by androgens was further elaborated in a study by María J. Ruiz-Echevarría's group [181]. The mRNA of TMEFF2 contains several potential upstream open reading frames (uORFs) in its leader region. The uORFs are short open reading frames present upstream of the translation initiation sites in 5-10% of cellular mRNAs. uORFs can modulate the translation of mRNAs by hampering the translation of downstream coding regions and, more intriguingly, under cellular stress conditions they can promote the translation of the coding region. The latter involves the activity of the eukaryotic translation initiation factor 2 (eIF2α). Under normal conditions, the GTP bound eIF2α finds the translation initiation codon efficiently, whereas under cellular stress conditions, eIF2α is phosphorylated, which inhibits the exchange of eIF2-GDP to eIF2-GTP. The GDP bound eIF2α is impaired in its ability to find the start codon, thus leading to a generalized decrease in translation initiation, but the mRNAs containing uORFs keep translating actively, i.e., the uORFs aid the translation of certain mRNAs under stress conditions. There are two proposed mechanisms for this subjectively favourable translation. First, in the case that the mRNA contains a single uORF, the inefficient eIF2α-P misses the start codon of the uORF and goes straight to translating the target ORF because of better Kozak consensus sequence. Second, when several uORFs are present, translation re-initiation gets more efficient because, if eIF2α-P does initiate translating one uORF, the next ones are bypassed as the start codon recognition capacity of eIF2α-P is impaired. Examples for the latter are well documented in the case of activating transcription factor 4 (ATF4) mRNA [182], and a study by Overcash *et al.* [181] proves it to be the case in TMEFF2 mRNA translation. As the group had previously reported TMEFF2 as an androgen-dependent onco-suppressor in prostate cancer, they focused their attention on the uORFs in TMEFF2 mRNA. The group established *Gaussia* luciferase (GLuc) constructs where *TMEFF2* 5'-uORFs were deleted one by one and placed under a CMV promoter. In both the androgen-dependent 22Rv1 prostate cancer cells and the androgen-independent PC3 cells, the removal of all 4 uORFs enhanced the translation of TMEFF2 mRNA many fold, whereas when dihydrotestosterone (DHT) was added to the 22Rv1 cell culture, a concentration-dependent increase in luciferase activity was only observed in the construct harbouring all 4 uORFs. That is to say, that although the uORFs in TMEFF2 mRNA hampered the constitutional transcription

of the gene, they were necessary for the androgen-dependent production of TMEFF2. The endogenous TMEFF2 protein and mRNA levels increased proportionally upon treatment with the increasing concentrations of dihydrotestosterone in 22Rv1 and LNCaP cells, whereas when the androgen receptor (AR) inhibitor bicalutamide was used in conjunction with dihydrotestosterone, this effect was dampened considerably. Especially in 22Rv1 cells, when the androgen receptor was knocked down using siRNA (siAR), dihydrotestosterone failed to induce the expression of TMEFF2. This suggested not only that the TMEFF2 was an androgen regulated gene in prostate cancer, but this regulation was dependent on androgen receptor signalling. A more interesting observation came when 22Rv1 cells were treated with actinomycin D and dihydrotestosterone in a time course experiment. Dihydrotestosterone induced the increase in both the protein and mRNA levels of TMEFF2 as expected, but when transcription was inhibited using actinomycin D, the increase in TMEFF2 mRNA upon dihydrotestosterone treatment was hampered while the protein levels still kept increasing in a time-dependent manner. This implied that the androgen receptor signalling not only induced the TMEFF2 transcription but also regulated the translation of the gene. They also reported an enhanced EIF2α dependent translation of TMEFF2 upon treatment with thapsigargin, a drug used to induce ER stress. These observations suggested that the androgen signalling enhanced not only the transcription of *TMEFF2* but also the translation of TMEFF2 mRNA through the phosphorylation of eIF2α. The question that arises is why an oncogenic phenomenon like androgen signalling would lead to an increase in transcription and the translation of a potential onco-suppressor. As the phosphorylation of eIF2α leads to a generalized decrease in the translation during stress conditions, enhancement of certain genes in these conditions could either help cellular adaptation or promote the initiation of senescence or apoptosis (in case the adaptation fails), meaning that TMEFF2 could either be an adaptive protein, an apoptotic protein or both – the premise may offer an explanation regarding the oncogenic *vs* onco-suppressive role of TMEFF2.

Vera Knauper's group [133] provided another possible explanation in this regard (i.e., the oncogenic vs onco-suppressive functioning of TMEFF2). The study group specializes in investigating a family of membrane-bound and secreted metalloproteases known as ADAMs (a disintegrin and metalloprotease). ADAMs are membrane-anchored modular glycoproteins that typically contain a pro-domain, a metalloprotease domain, a disintegrin domain, a cysteine rich domain, an EGF-like domain followed by the transmembrane and cytoplasmic domains. Functionality-wise, ADAMs have been implicated in diverse physiological, developmental and pathological process, including (but not

limited to) sperm-egg interactions, axon guidance and neuronal development, cell migration and immunity and inflammation. A very interesting function of ADAMs is protein ectodomain shedding, a phenomenon where extracellular portions of membrane proteins are cleaved off by proteases and released into the surrounding ECM or even blood. The release of ectodomains can lead to either functional activation or inactivation of substrate proteins. In the case of membrane bound ligands, chemokines and cytokines, the ectodomain shedding leads to the release of the ligand from the cell surface enabling them to interact with receptors either through bloodstream or in autocrine and paracrine manners. A prime example of this phenomenon is the well documented ectodomain shedding of TNFα ligand by ADAM17 [183]. Interestingly, the ectodomain shedding is required even for the autocrine and juxtracrine signalling of some ligands like EGFR ligand [133]. On the other hand, the ectodomain release of juxtracrine ligands may serve to downregulate their function by reducing their proximity to the receptors. Similarly, the ectodomain shedding of receptor proteins could result in their functional inactivation. This inactivation is aided further where the shedded receptors act as decoys and sequester the ligands away [183]. An example is that of TNF-receptor-associated periodic febrile syndrome where the mutations in TNFR1 render them unsheddable, which in turn results in heightened TNFα signalling and a consequent hyperinflammatory response [184]. Ectodomain shedding can also lead to the activation of receptors through a phenomenon known as regulated intramembrane proteolysis (RIP). Here, after the cleavage of the ectodomain, the remaining stub is cleaved within the membrane itself through an intramembrane-cleaving protease. The resulting fragment thus generated participates in signalling. Upon the intramembrane cleavage of Notch receptor, for example, the liberated intracellular domain is capable of entering the nucleus and regulating the specific target genes [184]. The research in ADAMs was greatly highlighted upon the discovery and characterisation of ADAM17 as the TNFα converting enzyme (TACE) in 2008 [183]. Since then the list of ADAM17 substrates has been growing and it has emerged as the protease with broadest substrate range among the ADAMs family of proteases. Especially, ADAM17 is involved in the activation of almost all of the known EGFR ligands, which include EGF, HB-EGF, betacellulin, epiregulin, amphiregulin and the transforming growth factor α (TGFα). Moreover, the EGF receptors HER4 and HER2 are also substrates for ADAM17 and ADAM10 respectively. As TMEFF2 contains an EGF-like domain, it is a monomeric transmembrane proteoglycan and ectodomain shedding of TMEFF2 was reported upon treatment of A172 glioma cells with TNFα and interleukin1 [185], Knauper's group investigated TMEFF2 as a substrate of ADAMs. When phorbol ester α (PMA) was used to induce the ADAM activity, it resulted in the release of TMEFF2 ectodomain in the conditioned

media of TMEFF2-stable HEK293T cells. The PMA-induced ectodomain shedding of TMEFF2 was either reduced or abolished upon simultaneous treatment with metalloproteinase inhibitor GM6001, ADAM17 inhibitors TIMP3 and TAP-1 (TNF-α protease inhibitor), and also upon treatment with ADAM17 siRNA. The inhibitors for ADAM10, ADAM12 and ADAM9 siRNA failed to reduce TMEFF2 shedding. In the same study, TMEFF2 was shown to increase the proliferation of HEK293T cells both through overexpression and through the addition of soluble TMEFF2 ectodomain in the media. The simultaneous addition of ADAM17 inhibitor (TAP-1) and TMEFF2 overexpression reversed this increase in HEK293T cells proliferation, i.e., the pro-proliferative effect of TMEFF2 was exclusive to its shedded ectodomain. Importantly, the removal of EGF-like domain from TMEFF2 did not have any effect on its pro-proliferative function. On the other hand, the same study also reported that TMEFF2 activated the EGF receptor HER-1 which resulted in the subsequent phosphorylation of ERK1/2 in HEK293T cells. A very important finding was the role of proteases in processing the remaining membrane bound stub of TMEFF2 after its ectodomain was shed. When cell lysates of HA-TMEFF2-V5 stable HEK293 cells were immunoblotted for V5, three distinct bands were observed: one at 10 kDa, another one at 14 kDa (major one) and a very minor one at 22 kDa. Interestingly, upon treatment with both metalloproteinase and γ-secretase inhibitors the 10 kDa band disappeared. The experiment was repeated in PC3 prostate cancer cells and CHO cells with the same results except for one important difference: upon the treatment with γ-secretase inhibitor, the 14 kDa band increased in intensity. Here an important aspect to note is that metalloproteinase activity is a prerequisite for γ-secretase action. It suggests that both the 14 kDa and 22 kDa bands were the products of metalloproteinase activity, whereas the 10 kDa band was a product of subsequent γ-secretase action. The finding is interesting in the sense that TMEFF2 is possibly processed in the plasma membrane in a manner similar to that of Notch receptor. This could have functional implications in different cell and/or tissue types and microenvironments. This aspect is proven in a rather recent study [186] by the same group. In order to broaden the scope of proteases that could potentially target TMEFF2, an additional class of membrane bound proteases was introduced in the investigation. Type II transmembrane serine proteases (TTSPs) are well documented for their involvement in the proliferation and metastasis of various cancers [187]. TTSPs are characterized by their C-terminal catalytic domain, variable stem region and short transmembrane domain near the N-terminal. TTSPs have six subfamilies in chordates, two of which I shall concentrate on – matriptase and hepsin. Matriptases are probably the most well-studied in terms of cancer proliferation and spread. Matriptases aid the progression of breast cancer through the activation of c-MET pathway [187] and

also correlate with Gleason Score in prostate cancer [188]. Also, hepsin overexpression is a common occurrence in prostate tumours. Matriptase-1 and hepsin were shown to cleave TMEFF2 ectodomain on their own while matriptase-2-induced shedding of TMEFF2 was dependant on the activity of ADAM17. The experiments involved the alkaline phosphatase (AP) activity assay of the conditioned media of AP-TMEFF2-His stably transfected HEK293T cells. Importantly, in addition to shedding of TMEFF2 ectodomain by matriptase-1 and hepsin, additional C-terminal bands were observed in the cell lysates unique to each protease. Besides the previously described ADAM17 cleaved 14 kDa band (reported at 17 kDa in subsequent study), matriptase-1 cleavage resulted in a 24 kDa and a 28 kDa fragment, whereas hepsin cleavage resulted in a prominent novel 20 kDa major fragment and a unique 15 kDa minor fragment. When the experiment was later repeated using ADAM-cleavage-resistant AP-Δ303-320 TMEFF2 (which lacked the ADAM specific stalk sequence), matriptase-1 and hepsin were still able to shed TMEFF2 but the location of bands on the immunoblot of the cell lysate changed. The deleted residue (303-320) also contained two potential cleavage sites for TTSPs. This implies that both matriptase-1 and hepsin were still able to cleave TMEFF2 at some novel sites. In short, the mechanisms revolving around the shedding of TMEFF2 ectodomain and the further processing of the remaining intramembranous stub could be the deciding factors in subsequent functioning of TMEFF2, i.e., it was proposed that the fragments produced through the action of various proteases would function differently. The authors presented it as a mechanism which affects the oncogenic versus onco-suppressive activities of TMEFF2 depending on cell type and/or the cellular and tissue microenvironment.

Another indication supporting this phenomenon came from the study by Chen *et al.* [189] from Ruiz-Echeverría's group. When full-length TMFF2-MYC-HIS was overexpressed in HEK293T cells, it could bind and modulate the activity of sarcosine dehydrogenase (SARDH), an enzyme that converts the amino acid sarcosine into glycine. When only the ectodomain of TMEFF2 (lacking intracellular and transmembrane domains) was overexpressed, it neither bound to nor affected the activity of SRDH. Moreover, full-length TMEFF2 overexpression reduced the proliferation of HEK293T cells (transient transfection), sensitised the HEK293T cells for apoptosis by staurosporine (stable transfection) and reduced the colony formation in soft agar by ~5 fold (stable transfection). In contrast, when only TMEFF2 ectodomain was added to media of HEK293T cells and RWPE1 prostate cancer cells, it led to an increase in cell proliferation. Notably, the overexpression of TMEFF2 ectodomain in HEK293T cells did not cause a decrease in the colony formation in soft agar, i.e., the intracellular domain was shown

to be necessary for the typical onco-suppressor effects of TMEFF2 while the shedded extracellular domain appeared to be a ligand which enhances the growth of the same cells. Conversely, the endogenous TMEFF2 was also reported to bind SRDH in the LNCaP and 22Rv1 prostate cancer cells. Sarcosine has been reported as a marker for prostate cancer progression and metastasis and the knockdown of SRDH in prostate epithelial cells increases their mobility and invasion. In a doxycycline-inducible TMEFF2-RWPE1 cell line, upon adding sarcosine and doxycycline together, TMEFF2 reversed the cell migration enhancement effect of sarcosine [189].

In a related study [190], Ruiz-Echeverría's group observed that the TMEFF2-induced reduction in cell migration was not merely due to (the enhanced SRDH and the resultant) decrease in the intracellular sarcosine concentration, rather the metabolism of sarcosine was the key factor in this process, more specifically the one-carbon metabolism. One-carbon metabolism refers to a group of chemical reactions where one carbon group is transferred from one metabolite to another. During this process, the carbon group is too unstable to survive on its own and needs a carbon carrier – tetrahydrofolate being the most common one. Sarcosine and dimethylglycine are two of the main one-carbon donors in one-carbon metabolism pathways. When the antifolate drug methotrexate (MTX) was used to treat the TMEFF2 knockdown cells, it abolished the increase in cell migration induced by TMEFF2-knockdown [190]. TMEFF2 knockdown also reduced the mRNA levels of dimethylglycine dehydrogenase (DMGH) and glycine decarboxylase (GLDC) which are important contributors to one-carbon metabolism. Another interesting finding in this study was that TMEFF2 co-localized with β-actin and α-tubulin in 22Rv1 and freestyle 293-F cells. Also, upon knockdown of SRDH, the interaction of TMEFF2 with the cytoskeleton was enhanced.

Thus far, I have examined two proposed explanations regarding the oncogenic *vs* onco-suppressive role of TMEFF2 in prostate cancer, i.e., TMEFF2 being an adaptive and/or an apoptotic protein (owing to the role of uORFs in TMEFF2 mRNA) and the differences in the proteolytic processing of TMEFF2 depending upon the cellular and tissue microenvironment. One more study worth mentioning here investigated yet another feature of TMEFF2, i.e., the potential GPCR signalling motif in its C-terminal portion. Chen *et al.* [191] cloned a deletion mutant lacking 13 consecutive basic-rich amino acids in the C-terminus of the protein (TMEFF2_ΔGA) and compared it to the full-length FL_TMEFF2. When both forms were expressed in the prostate cancer cell lines RWPE1 and RWPE2, FL_TMEFF2 showed a significant inhibition of cell migration but TMEFF2_ΔGA did not. Additionally, in this study, the overexpression of FL_TMEFF2 in RWE2 prostate cancer cells (i) reduced the expression of integrins

αv, β1 and β3, (ii) inhibited RAS homolog family member A (RHOA) activation, (iii) reduced formation of stress fibres and focal adhesions by impairing phosphorylation of the focal adhesion kinase (FAK), and (iv) promoted cell rounding (thus preventing lamellipodia protrusions leading to decreased migration of cells). On the other hand, the overexpression of TMEFF2_ΔGA failed to accomplish any of the above. In an interesting experiment the group generated transgenic mice with inducible TMEFF2 expression in prostate tissue. These were crossed with a mouse model of prostate cancer TRAMP (transgenic adenocarcinoma of mouse prostate). The expression of TMEFF2 in TRAMP mouse led to a decrease in expression of integrins αv, β1 and β3.

Summarizing the role of TMEFF2 in prostate cancer, one could draw the following deductions: (i) the proteolytic processing of ectodomain and the remaining intramembranous and intracellular portions of TMEFF2 can affect its functions, (ii) the involvement of eIF2α in the regulation of TMEFF2 expression, enhanced expression of TMEFF2 under ER stress and induction of TMEFF2 shedding by reactive oxygen species (ROS) point towards TMEFF2 being a stress-related protein, (iii) although there are opposing reports regarding the oncogenic *vs* onco-suppressive roles of TMEFF2, there is more evidence in support of the latter, and (iv) the physiological function of TMEFF2 remains unknown.

5.1.1.6 TMEFF2 in gastric and colorectal cancers

Further corroborating evidence regarding the onco-suppressive role of TMEFF2 came from the studies regarding its involvement in gastric and colorectal cancers. The findings on the proposed functions of TMEFF2 in these studies are not only interesting but also important in the context of my own findings regarding the gene. As mentioned earlier, Sun *et al.* [103] reported significant downregulation of TMEFF2 through various stages of gastric cancer, which directly correlated with its promoter methylation and worsening patient survivability. This study reported that TMEFF2 interacted with SHP1 (tyrosine-protein phosphatase non-receptor type 6) in AGS and MKN45 gastric cancer cells resulting in cell cycle arrest, induction of apoptosis, and a decrease in cell proliferation both *in vitro* and *in vivo* (xenograft models). This interaction between TMEFF2 and SHP1 was abolished upon deletion of intracellular domain of TMEFF2. Moreover, no mutations were found in the TMEFF2 intracellular domain in gastric cancer patients (TGCA data). These findings corroborate with the previously mentioned importance of the TMEFF2 intracellular domain and also the methylation rather than mutation being the primary cause of TMEFF2 loss of function. The

investigators also proposed that TMEFF2 had a role in maintaining genomic integrity. TMEFF2 knockdown in GES-1 gastric cancer cells along with enhanced proliferation also led to higher DNA damage. DNA damage related genes were found to be enriched when (i) TMEFF2 overexpressing AGS cells were subjected to cDNA microarray and subsequent gene ontology analysis, and (ii) gastric tumours with higher TMEFF2 expression were compared with those with lower expression of the gene.

In a later study [163], the same research group investigated the role of TMEFF2 in *Helicobacter pylori*-induced gastric cancer development. The *H. pylori* infection resulted in both upregulation and phosphorylation (activation) of STAT3. STAT3 was reported to downregulate TMEFF2 through binding onto its promoter, i.e., *H. pylori* infection resulted in downregulation of TMEFF2 through STAT3 (in cell cultures, mouse models and human gastric tissue pathology). On the other hand, TMEFF2 overexpression decreased STAT3 phosphorylation in AGS gastric cancer cells both in culture and in AGS xenografts (TMEFF2 overexpression also resulted in a dramatic decrease in the tumour size). The TMEFF2 induced downregulation of p-STAT3 was also dependent on its interaction with SHP-1.

Further evidence for the involvement of STATs in TMEFF2 function came from David Shibata's group. Firstly, Elahi *et al.* [192] reported that TMEFF2 induced apoptosis in HCT116 colorectal cancer cells and reduced their proliferation through activation of STAT1 pathway. Interestingly, when TMEFF2 stable HCT116 colorectal cells were subjected to microarray analysis, 18 of the 42 upregulated genes were found to be interferon-inducible. These interferon-inducible genes exhibited the highest fold change among the list of upregulated genes. The treatment of HCT116 cells with STAT1 siRNA reversed the TMEFF2-induced reduction of proliferation both on treated surface and soft agar. Interferons in general (and interferon alpha, IFN-α, in particular) are a known inducer of apoptosis in malignant and virus harbouring cells [193]. Although interferon signalling is very complex, according to a general rule of thumb, interferon gamma (IFN- γ) leads to STAT1:STAT1 homodimers which later translocate to the nucleus and trigger the transcription of the genes harbouring GAS sites in their promoter. The interferon alpha (IFN-α) activation leads to the formation of STAT1:STAT2 heterodimers and transcription of genes with ISRE sites in their promoters. In a subsequent study by the same group, Hernandez *et al.* [194] reported that the TMEFF2 overexpression in HCT-116 cells led to upregulation of genes with both GAS and ISRE harbouring promoters (GAS to a higher degree than ISRE). Additionally, phosphorylation and subsequent nuclear localization of both STAT1 and STAT2 were observed in TMEFF2 stable HCT-116 cells. STAT phosphorylation is mediated by four kinases of

the Janus family (JAK1, JAK2, JAK3 and TYK2). Hernandez *et al.* showed that the TMEFF2 overexpression results in the upregulation of JAK1 and JAK2 but not the JAK3 and TYK2 kinases. HCT-116 cell line is resistant to interferon therapy (like several other cancers). An interesting observation from this study was that TMEFF2 overexpression marginally sensitized the HCT-116 to INF-α but not INF-γ.

To summarize the discussion on TMEFF2, one can confidently claim that TMEFF2 has an onco-suppressive activity in several cancers along with being a potential prognostic and diagnostic marker. The onco-suppressive *vs* oncogenic activities of TMEFF2 do not appear to be cell or tissue specific, rather the cumulative evidence points more towards the extent and duration of overexpression and the differential post-translational processing of the protein to be the main factors in this regard. Conversely, an aspect which could be of interest in future studies is the TMEFF2 glycosylation. As the protein has both O-linked glycosaminoglycans and N-linked glycans attached, it could very well be possible that differential glycosylation of TMEFF2 affects its function. Having such diverse roles in a multitude of pathologies, the interest in TMEFF2 appears to be gravely underrated.

***Table 5.1*: Reported functions of TMEFF2**

Author and Reference	Function	Oncogenic Vs Onco-suppressive	Disease
Lin, Taylor *et al*. 2011[155]	TMEFF2 binds and inhibits PDGF-AA through its follistatin-like domains.	Onco-suppressive	Cancers (Several)
Horie, Mitsumoto *et al*. 2000[132]	TMEFF2 promotes survival of Hippocampal and Mesencephalic neurons in primary culture.	Neither but pro-growth	Neuronal development
Yonezawa, Uchida et al.2000[139]	Tomorgeulin mildly activates erbB-4 receptors in MKN28 gastric cancer cells	Oncogenic	Gastric cancer
Labeur, Wolfel *et al*. 2015 [134]	TMEFF2 inhibits Akt phosphorylation and production of ACTH in response to CRH stimulation in AtT20 glioma cells. TMEFF2 decreases the proliferation of AtT20 cells.	Onco-suppressive	Glioma
Huang, Teng *et al*. 2017 [179]	(i)*Tmeff2* is specifically upregulated in OPCs at the onset of cell differentiation but it is dispensable in the process. (ii)TMEFF2 weakly interacts with PDGFA but does not prevent the action of PDGFA on its receptor in HEK293T cells. (iii)TMEFF2 overexpression increases	Neither but pro-growth	Development

	ERK1/2 phosphorylation in HEK293T cells.		
Gery, Sawers *et al*. 2002[166]	TMEFF2 is regulated by androgens in prostate cancer and inhibits the proliferation of DU145 and PC3 prostate cancer cells.	Onco-suppressive	Prostate Cancer
Afar, Bhaskar *et al*. 2004[167]	TMEFF2mAb conjugated cytotoxic drug Pr1-vcMMAE effectively reduced tumour size in LNCap and CWR22 xenograft with minimal toxicity.	Potentially Oncogenic	Prostate cancer
Zhao, Schneider *et al*. 2005 [143]	β-emitting isotope 90Y labelled TMEFF2 mAB 2H8 effectively reduces LNCaP xenograft size and is well tolerated.	Potentially Oncogenic	Prostate cancer
Zhao, Liu *et al*. 2008[180]	TMEFF2 is expressed on cell surface and it is internalized and recycled in lysosomes. TMEFF2 mAB is complexed with cytotoxin saporin, it causes death of TMEFF2 stable PC3 cells.	Potentially Oncogenic	Prostate cancer
Hong, Maezawa *et al*. 2015 [135]	TMEFF2 binds Amyloid-β protein, its precursor AβPP and the neurotoxic oligomeric forms of Amyloid-β protein AβPOs *in vitro* and *in vivo*. TMEFF2 protects the N2a cells from neurotoxicity induced by AβPO.	NONE	Alzheimer's Disease
Ali and Knauper 2007 [133]	TMEFF2 ectodomain shedding is triggered by TNFα, is brought about by ADAM10 and ADAM17 and increases proliferation of HEK293T cells. TMEFF2 causes phosphorylation of ERK1/2.	Oncogenic	Prostate cancer
Gawel-Beben, Ali *et al*. 2018[186]	TMEFF2 is a substrate for membrane anchored serine proteases (TTSPs) matriptase-1 and hepsin.	Depends on PTM	Prostate cancer
Chen, Overcash *et al*. 2011 [189]	Full-length TMEFF2 reduces cell proliferation and sensitises the HEK293T cells for apoptosis, binds and augments SRDH activity, reduces sarcosine levels leading to inhibition of cell migration. TMEFF2 Ectodomain does the opposite.	Ectodomain: Oncogenic Full-length: Onco-suppressive	Prostate cancer
Chen, Corbin *et al*. 2014 [191]	Full-length TMEFF2 (i) reduces the expression of integrins αv, β1 and β3, (ii) inhibits RHOA activation, (iii) inhibits phosphorylation of the focal adhesion kinase (FAK), (iv) reduces cell migration. Deletion of putative G-protein domain prevents all of above.	Onco-suppressive	Prostate cancer
Green, Chen *et al*. 2013[190]	TMEFF2 reduces invasiveness of 22Rv1 prostate cancer cells through modulation of one-carbon metabolism. TMEFF2 interacts with the cytoskeleton.	Onco-suppressive	Prostate cancer

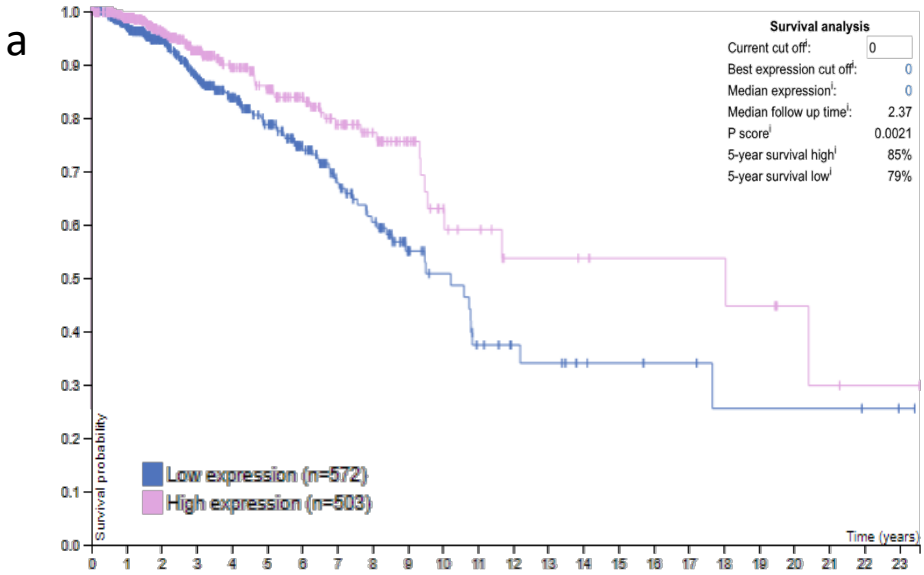
Elahi, Zhang et al. 2008 [192]	HPP1 induces apoptosis and causes changes in morphology and reduction in cell proliferation in HCT116 colorectal cells through activation of STAT1 pathway.	Onco-suppressive	Colorectal cancer
Hernandez *et al*. 2015[194]	HPP1 activates JAK-STAT interferon pathways to induce apoptosis and reduce proliferation of HCT-116 cells. TMEFF2 overexpression marginally sensitized the HCT-116 to INF-α induced cell death.	Onco-suppressive	Colorectal cancer
Sun, Tang *et al*. 2015 [163]	TMFF2 decreases STAT3 phosphorylation through its interaction with SHP1 in gastric cancer. STAT3 directly binds to the TMEFF2 promoter and represses its transcription	Onco-suppressive	Gastric cancer
Sun, Tang *et al*. 2015[103]	TMEF2 binds SHP-1 in gastric cancer and induces cell cycle arrest, apoptosis and prevents DNA damage.	Onco-suppressive	Gastric cancer

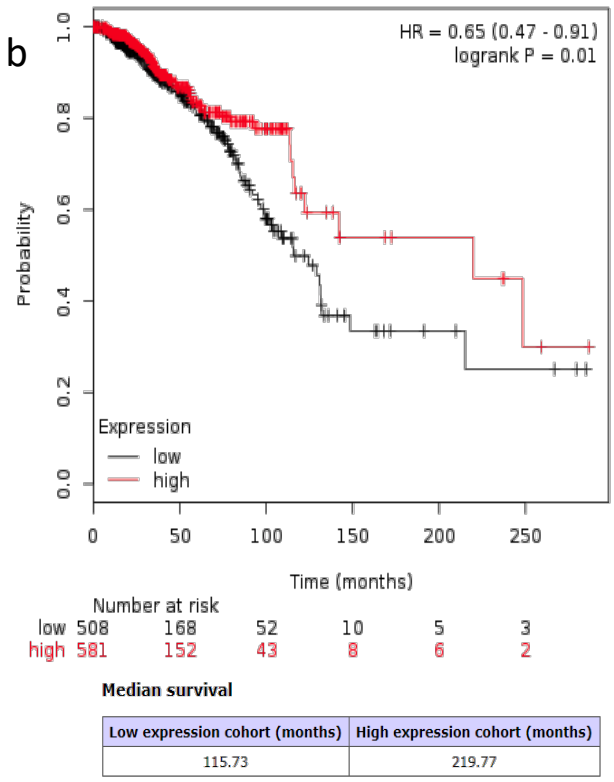
5.2 Results

5.2.1 TMEFF2 has an onco-suppressive role in breast cancer

As described previously, *TMEFF2* promoter methylation has been reported to increase progressively along the developmental path of breast cancer [151, 152], *TMEFF2* promoter was hypermethylated in recurrent breast cancer, compared to non-recurrent breast cancer [149] and TMEFF2 was among the top 10 in the list of 2674 differentially hypermethylated genes found in the serum of breast cancer patients [150]. In order to confirm the downregulation of TMEFF2 in breast cancer, I used the data mining platform Oncomine™ [195, 196]. The results are presented in supplementary figure S.3. Using the data available from the TGCA, the analysis by Oncomine™ [196] clearly shows a significant downregulation of TMEFF2 mRNA in breast cancer patients in comparison to the normal individuals. Moreover, Oncomine™ ranks TMEFF2 among the top 10% of the genes that are downregulated in breast cancer tissue when compared to the normal breast tissue (based on reporter A_23_P125382). I have already shown the Kaplan-Meier plot depicting a positive correlation between TMEFF2 mRNA expression and the breast cancer patient survival in Chapter 3, where the TMEFF2 mRNA expression levels were measured using the Affymetrix cDNA microarray (probe set ID, 233910_at). In order to confirm these results, I checked further from two additional sources, both of which used different cohorts of patients and a different technique (RNA-Seq) for measuring the mRNA levels: (i) Kaplan-

Meier Plotter Pan-cancer RNA-seq [72, 80] and (ii) Human Protein Atlas [77, 78, 79, 197]. Results are depicted in figure 5.2. The data from all three sources showed a significant correlation between low expression of TMEFF2 mRNA and poor breast cancer patient survival. Together these results are indicative of an onco-suppressive activity of TMEFF2 in breast cancer.





Median survival

Low expression cohort (months)	High expression cohort (months)
115.73	219.77

***Figure 5.2*: Kaplan-Meier plots showing the correlation between TMEFF2 mRNA expression and breast cancer patient survival**. mRNA expression levels were measured using RNA-Seq. Sources: (a) The Human Protein Atlas (https://www.proteinatlas.org). P represents log-rank p value. Cut off values are in FPKM. (b) Kaplan-Meier Plotter Pan-cancer RNA-seq (https://kmplot.com/analysis/index.php?p=service&default=true).

5.2.2 TMEFF2 overexpression causes apoptosis in MCF-7 cells

In order to investigate the type of cell death induced by TMEFF2 in MCF-7 cells, I employed four different methodologies. Firstly, 48 hours after transfection, MCF-7 cells were subjected to the caspase-3 and caspase-7 activity assay (Caspase-Glo® 3/7, Promega®). The results in figure 5.3 (a) show a significant and substantial increase in caspase activation upon TMEFF2 overexpression. As MCF-7 cells are caspase-3 deficient [198], these data depict the activity of caspase-7. Since the activated caspases can still be neutralized by the IAPs, I overexpressed TMEFF2 in MCF-7 cells and simultaneously treated the cells with the pan-caspase inhibitor Z-VAD-FMK. Luciferase was used as negative control while tBID was used as positive control in this experiment. Cells were harvested 48 hours post-transfection and subjected to TMRE / DRAQ7™ combined staining and flow cytometry. TMRE (tetramethylrhodamine, ethyl ester) is a cell-permeable, positively-charged, red-orange dye that readily accumulates in the active mitochondria due to their negative charge. The depolarized mitochondria have decreased membrane potential and fail to sequester TMRE. A decrease in TMRE staining therefore represents the loss of mitochondrial membrane potential – an indication of early apoptosis [199]. On the other hand, DRAQ7™ is a far-red fluorescent dye that only stains nuclei in dead cells with the permeabilized plasma membrane and nuclear envelope. An increase in DRAQ7™ staining therefore represents late stage apoptosis. As it is evident from figure 5.3 (b) and figure 5.3 (c) (c.iii - c.iv), although there was a small (but significant) decrease in total apoptosis upon treatment of TMEFF2 transfected cells with Z-VAD-FMK, the DRAQ7™ staining decreased dramatically, i.e., cells lost their mitochondrial membrane potential but because of the inhibition of caspase activity, their plasma membrane and/or nuclear envelope remained intact.

a.

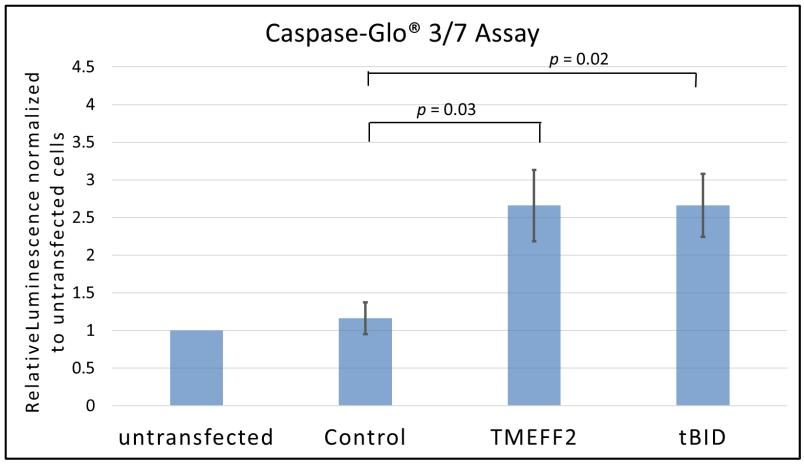


***Figure 5.3* (a): Caspase 3/7 activity upon TMEFF2 overexpression in MCF-7 cells.** Caspase 3/7 activity was measured through the relative luminescence (Caspase-Glo® 3/7 assay) upon TMEFF2 overexpression normalized to the untransfected cells. GFP was used as negative 'control' and tBID as positive control. Caspase-Glo® 3/7 reagent was added 48 hours post-transfection of MCF-7 cells, incubated for 1 hour and luminescence was measured. Bars represent average from three wells of a 96-well plate, each with an independently transfected population of cells. Error-bars represent the +/- standard deviation. P values were calculated using two-sample t-test assuming unequal variances.

b.

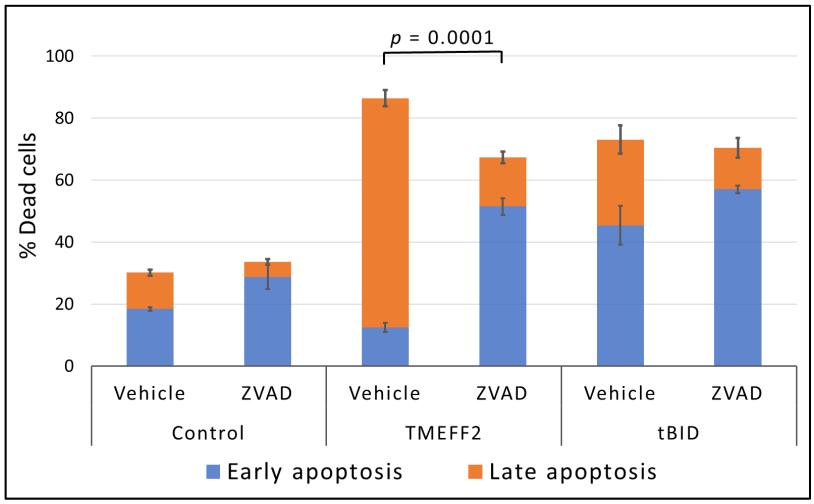


***Figure 5.3:* (b) Early and late apoptosis in TEMFF2 overexpressing MCF-7 cells upon inhibition of caspases.** Early and late apoptosis was measured using TMRE and DRAQ7™ staining respectively through flow cytometry. MCF-7 cells were transiently transfected with TMEFF2, luciferase (negative 'control') and tBID (positive control). 4 hours post-transfection, Opti-MEM® media was replaced with DMEM +10% FBS which contained either 50 μM Z-VAD-FMK or DMSO 'vehicle'. Cells were harvested 48 hours post-transfection, stained with TMRE/DRAQ7™ and subjected to flow cytometry. Clustered bars represent average from three different wells of the 24-well plate each with an independently transfected population of cells. Error-bars represent the +/- standard deviation. P value is for the total apoptosis = Early apoptosis + Late apoptosis and it was calculated using two-sample t-test assuming unequal variances.

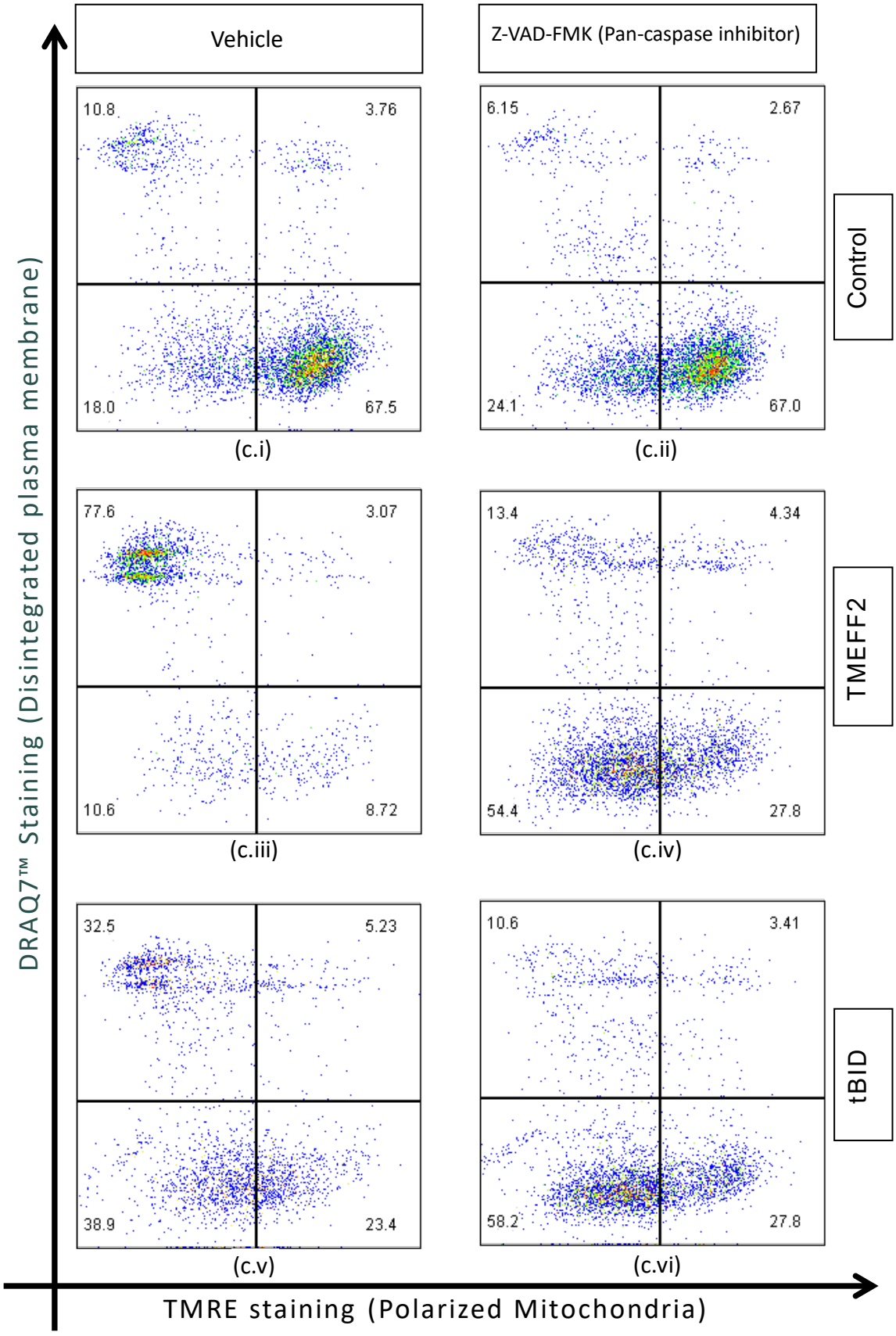

Vehicle
Z-VAD-FMK (Pan-caspase inhibitor)
Control
TMEFF2
tBID
10.8
3.76
18.0
67.5
(c.i)
6.15
2.67
24.1
67.0
(c.ii)
77.6
3.07
10.6
8.72
(c.iii)
13.4
4.34
54.4
27.8
(c.iv)
32.5
5.23
38.9
23.4
(c.v)
10.6
3.41
58.2
27.8
(c.vi)
DRAQ7™ Staining (Disintegrated plasma membrane)
TMRE staining (Polarized Mitochondria)

***Figure 5.3*: (c) TMRE (x-axis) and DRAQ7™ (y-axis) staining upon TMEFF2 overexpression in MCF-7 Cells:** Representative single scatter plots for the data presented in graph 5.3 (b). Bottom right quadrant: Live cells – TMRE +ve and DRAQ7™ -ve cells with intact plasma membranes and polarized mitochondria. Bottom left quadrant: Early apoptosis TMRE -ve and DRAQ7™ -ve cells with intact plasma membrane but depolarized mitochondria. Top left quadrant: Late apoptosis TMRE -ve and DRAQ7™ +ve cells with depolarized mitochondria and disintegrated plasma membrane. (c.i) MCF-7 cells transfected with Luciferase 'control' and treated with DMSO 'vehicle'. (c.ii) MCF-7 cells transfected with Luciferase 'control' and treated with 50 μM Z-VAD-FMK. (c.iii) MCF-7 cells transfected with TMEFF2 and treated with DMSO 'vehicle'. (c.iv) MCF-7 cells transfected with TMEFF2 and treated with 50 μM Z-VAD-FMK. (c.v) MCF-7 cells transfected with tBID and treated with DMSO 'vehicle'. (c.ii) MCF-7 cells transfected with tBID and treated with 50 μM Z-VAD-FMK. The percentage population in each quadrant is indicated.

In order to probe further for apoptosis being the type of cell death induced by TMEFF2 in MCF-7 cells, I performed immunoblotting for some key apoptotic factors. The results are presented in figure 5.4. The cleavage of poly ADP-ribose polymerase-1 (PARP-I), a target of active caspases, is routinely used as an apoptotic marker [200]. When an antibody against both full-length (~ 116 kDa) and cleaved (~89 kDa) PARP-1 was used, a dramatic increase in the intensity of cleaved PARP-I was observed upon TMEFF2 overexpression. Moreover, although the activation of caspase-7 had already been observed through the Caspase-Glo® 3/7 assay (figure 5.3 (a)), the immunoblot of caspase-7 showed an increase in both its activation and expression in TMEFF2 transfected cells. As for the anti-apoptotic BCL2, two close bands were observed just above the 25 kDa molecular marker. The expected band for BCL2 is reported to appear at ~ 26 kDa, assuming the lower band was the right one, the BCL2 protein level decreased rapidly upon TMEFF2 overexpression. Furthermore, MCL-1, another anti-apoptotic member of BCL2 family, has two splice isoforms with opposite functions. The long isoform (MCL-1L) is the canonical anti-apoptotic protein that harbours all three BH1, BH2 and BH3 domains along with a transmembrane domain. A shorter splice variant of MCL-1 (MCL-1S) is a BH3 domain-only pro-apoptotic protein [201]. When an antibody which detects both long and short isoforms of MCL-1 was used, upon the overexpression of TMEFF2, the pro-apoptotic MCL-1S band increased in intensity while the anti-apoptotic MCL-1L remained largely unchanged. TP53, more commonly known as p53, is regarded as the master regulator of apoptosis [1]. Upon TMEFF2 overexpression, a substantial increase in TP53 was observed. PUMA, another pro-apoptotic BCL2 family member, falls under the direct transcriptional regulation of p53. Upon TMEFF2 overexpression, PUMA also increased substantially. Although the intention of this experiment was to confirm the induction of apoptosis by

TMEFF2 rather than identifying the mechanism behind it, an increase in TP53 and PUMA was worth further investigation. MDM2, an E3 ubiquitin ligase, is one of the main regulators of p53. MDM2 not only marks p53 for proteasomal lysis through its ligase activity, but also actively represses its transcriptional activity. However, upon cleavage by caspase-2 and the subsequent loss of C-terminal RING domain, MDM2 loses its E3 ligase ability and the N-terminal portion of MDM2 stabilizes TP53 instead [1]. Upon TMEFF2 overexpression, a slight increase in the full-length MDM2 was observed. As MDM2 lies under direct transcriptional control of p53, this increase was not surprising. However, no difference in protein level of the cleaved ~60 kDa fragment of MDM2 (p60) was observed between the control and test lysates. This result alleviated the possibility of MDM2 cleavage being the reason behind TMEFF2 induced increase in TP53. Besides, to check for the possibility of autophagy being an additional form of cell death induced by TMEFF2, I also assayed for any changes in SQSTM1 (p62) with the aim of measuring autophagic flux. No difference in SQSTM1 was detected upon TMEFF2 overexpression (supplementary figure S.4). Together these results provided further evidence of apoptosis being the primary mode of cell death in MCF-7 cells upon the overexpression of TMEFF2.

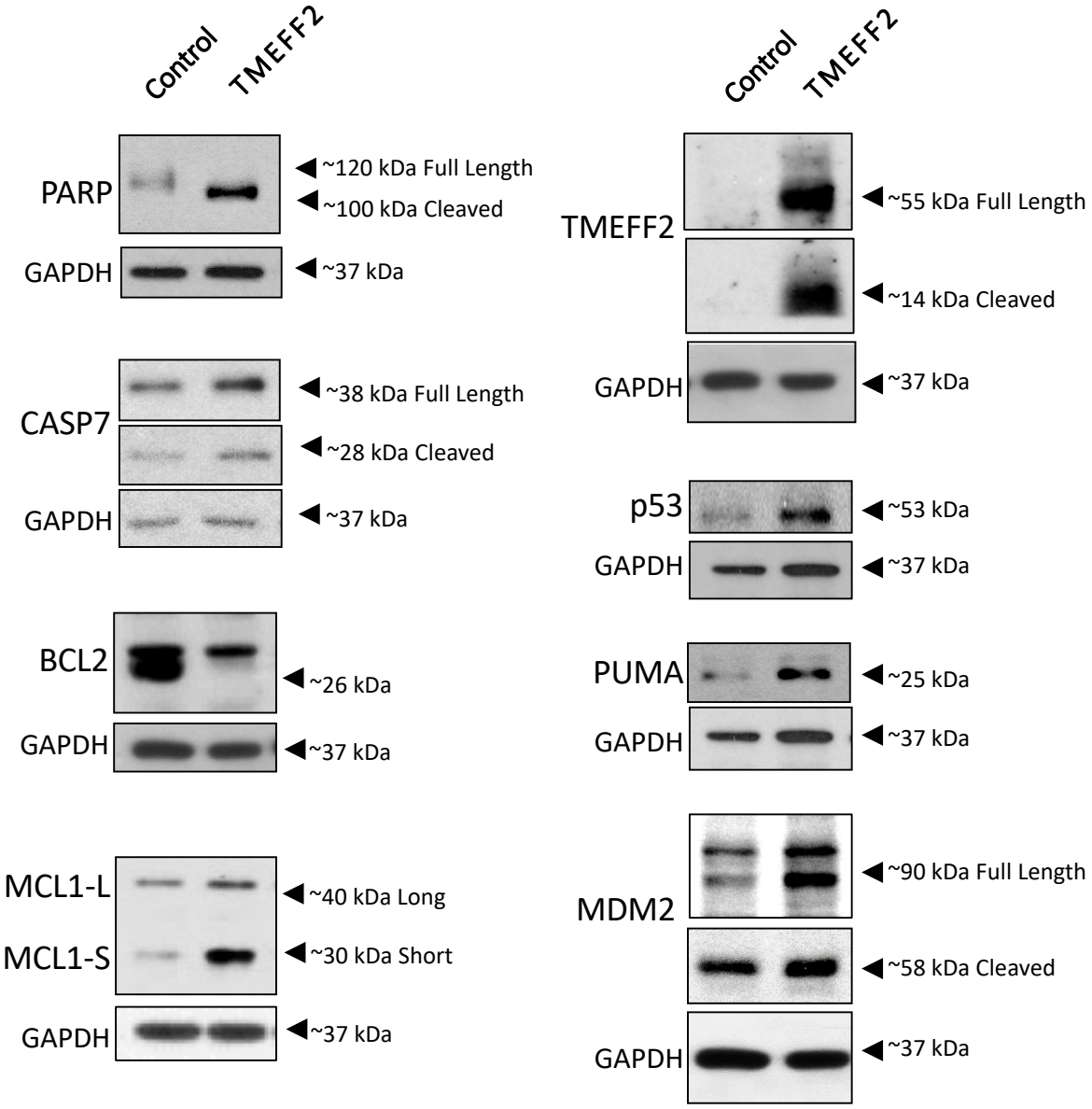
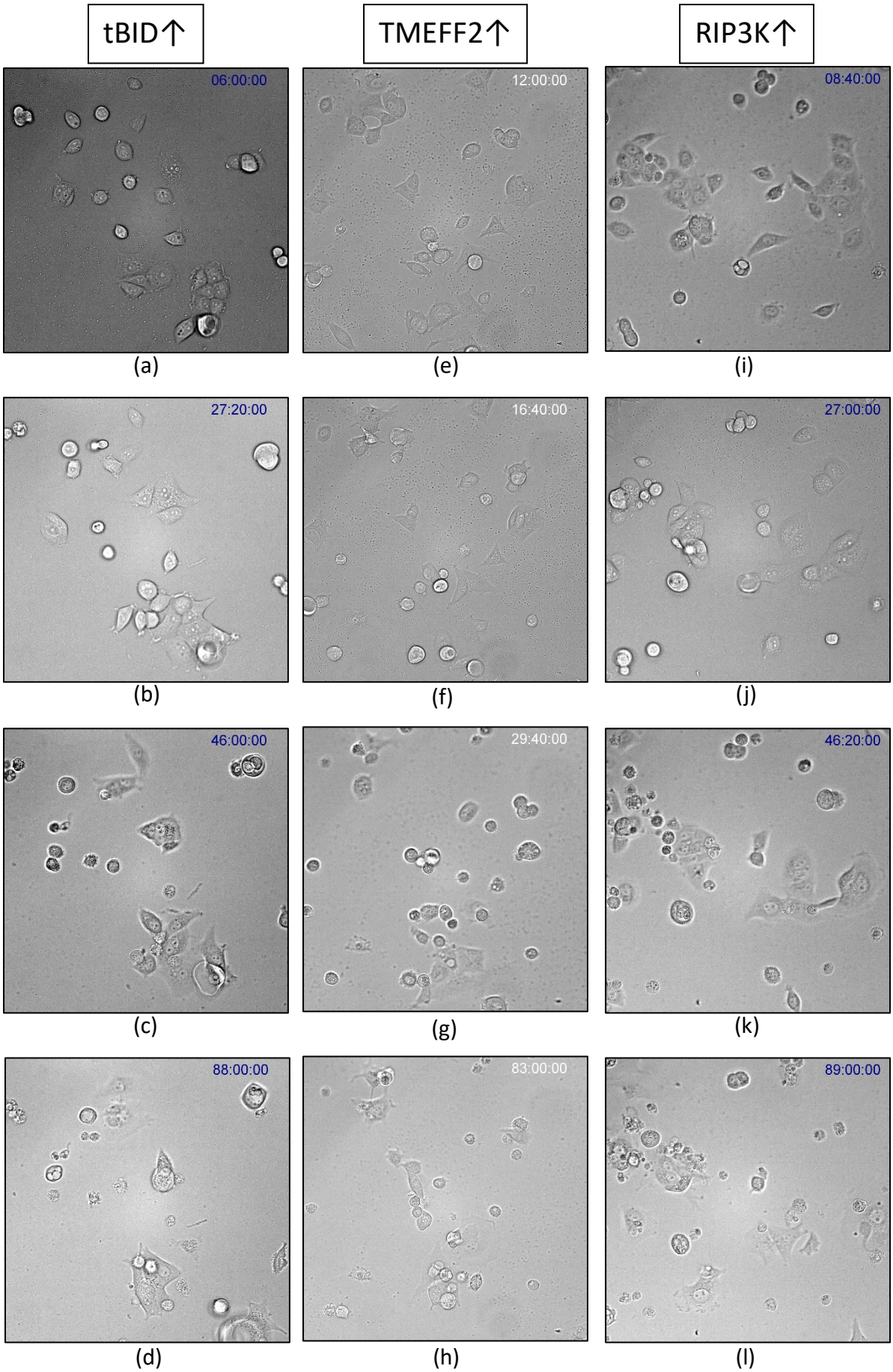


***Figure 5.4*: Immunoblots for key apoptotic factors upon TMEFF2 overexpression**. MCF-7 cells were transfected with TMEFF2 and GFP (negative 'control') in 10 cm dishes and harvested 30 hours post-transfection. Proteins were separated through SDS-PAGE under reducing conditions and electroblotted onto nitrocellulose membrane. The list of antibodies used and immunoblotting protocol are explained in the Materials and Methods section. GAPDH was used as a loading control. Immunoblotting (for the provided cell pellets) was performed by Master student Miss Ottilie Swinyard.

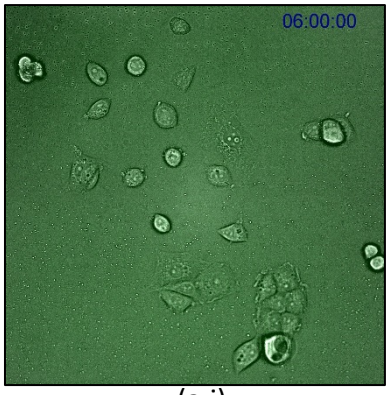
As I have already explained in the introduction, for the classification of cell death sub-routines, the 2012 NCCD guidelines put a significant emphasis on the morphological changes during the cell death process. In view of the above, I performed the time-lapse microscopy on MCF-7 cells transfected with

TMEFF2. tBID was used as the positive control for apoptosis while RIP3k was used as the positive control for necroptosis. In order to get an estimate of transfection efficiency, cells were co-transfected with GFP as a reporter in the ratio of 1:4 (GFP: test plasmid). The excitation laser was kept at the lowest intensity possible in order to avoid phototoxicity. Several fields were micrographed over a number of repeated experiments. The micrographs from two of these fields (the representative results) are presented in figure 5.5. In addition, the hyperlinks for the time lapse microscopy movies are provided in table 5.2. (NB: time lapse movies present a more accurate estimation of morphological changes over time). The morphological features of TMEFF2 transfected cells greatly resembled those transfected with tBID. Cells transfected with TMEFF2 and those transfected with tBID showed some typical features of apoptosis including rounding of cells, detachment and the most vivid ones being the cytoplasmic shrinkage and plasma membrane blebbing. However, formation of apoptotic bodies was not observed either in tBID or TMEFF2 overexpressing cells. Also, cells transfected with tBID retained the plasma membrane integrity for longer as compared to those transfected with TMEFF2. In some instances, the tBID transfected cells displayed membrane blebbing without losing the plasma membrane integrity till the end of experiment, i.e., 92 hours post-transfection. Furthermore, some vacuolization was also observed in tBID and TMEFF2 overexpressing cells. For the cells transfected with RIP3K, by and large, cells exhibited rounding and distension which was followed by an abrupt rupture. A salient feature observed in RIPK3 transfected cells was the formation of large swelling bubbles around the plasma membrane, most probably indicating the loss of plasma membrane integrity. Substantial cell death was observed towards the end of the experiment in all conditions.

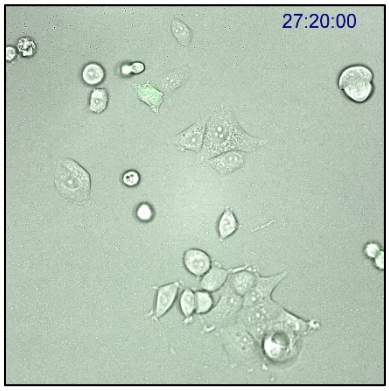
To summarise these experiments, upon TMEFF2 overexpression in MCF-7 cells, I observed (i) loss of mitochondrial membrane potential, (ii) abolition of late apoptosis upon inhibition of caspases, (iii) upregulation of some key apoptotic factors and (iv) morphological features of apoptosis. Together these results suggest that apoptosis is the primary mode of cell death induced by TMEFF2 in MCF-7 cells. From this point forward, TMEFF2-induced cell death in MCF-7 cells will be referred to as apoptosis.


tBID↑
TMEFF2↑
RIP3K↑
06:00:00
12:00:00
08:40:00
(a)
(e)
(i)
27:20:00
16:40:00
27:00:00
(b)
(f)
(j)
46:00:00
29:40:00
46:20:00
(c)
(g)
(k)
88:00:00
83:00:00
89:00:00
(d)
(h)
(l)

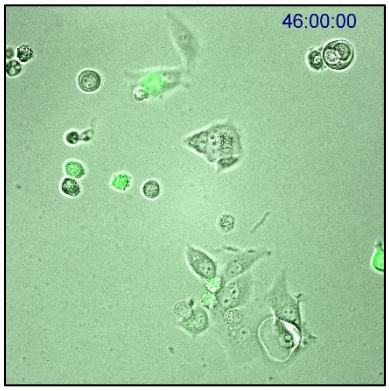
tBID↑



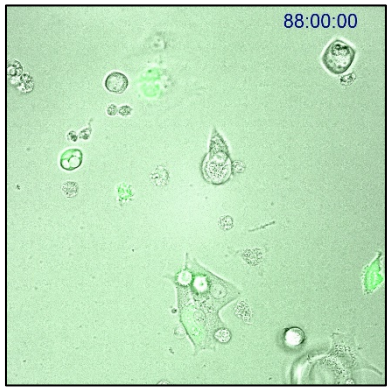
(a.i)



(b.i)



(c.i)

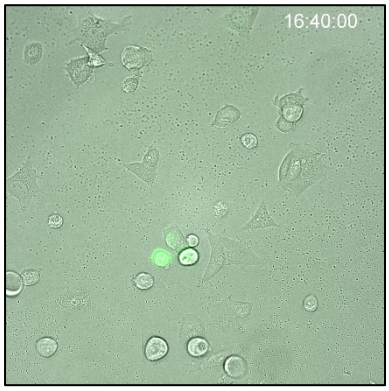


(d.i)

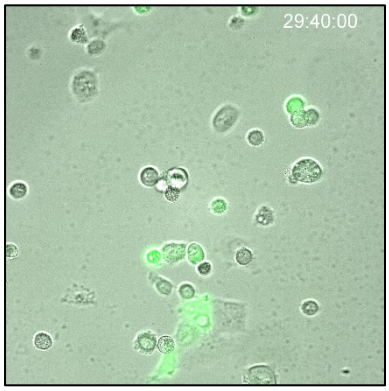
TMEFF2↑



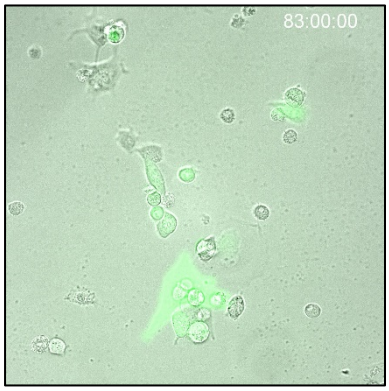
(e.i)



(f.i)



(g.i)

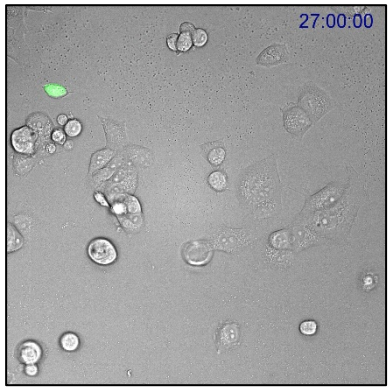


(h.i)

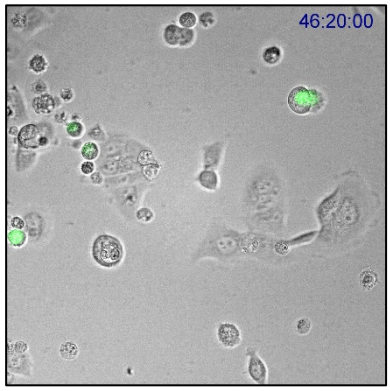
RIP3K↑



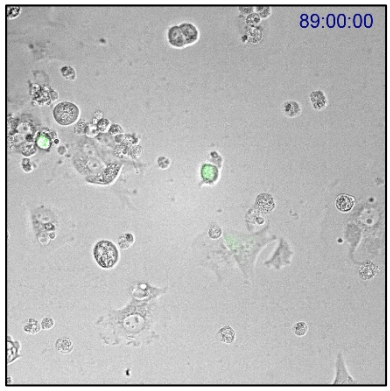
(i.i)



(j.i)

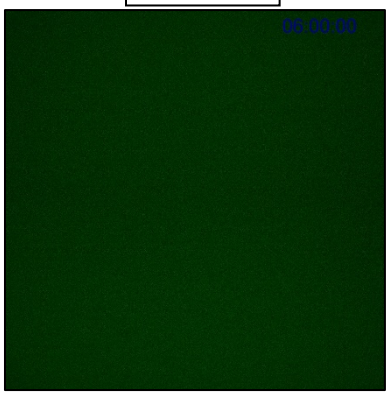


(k.i)

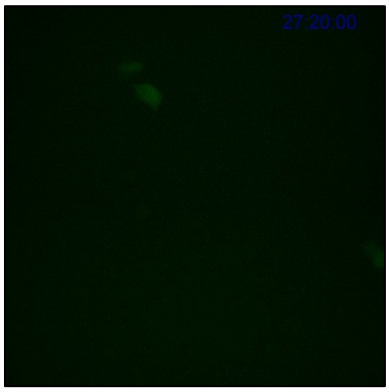


(l.i)

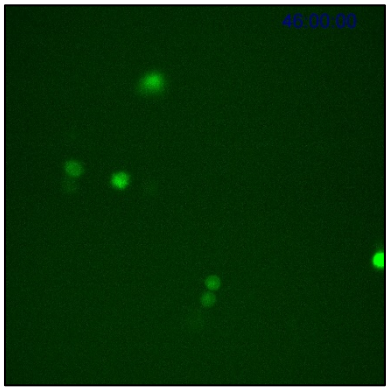
tBID↑



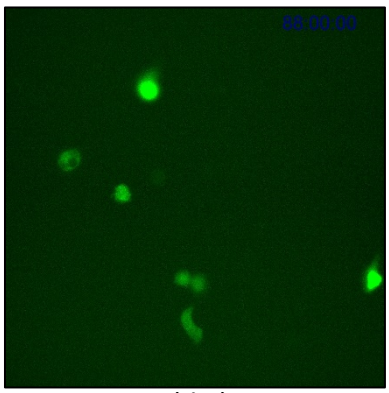
(a.ii)



(b.ii)

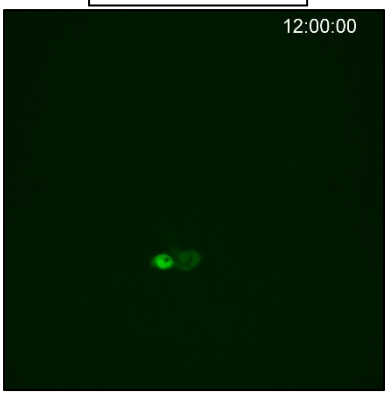


(c.ii)

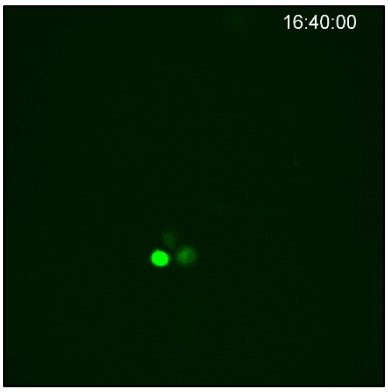


(d.ii)

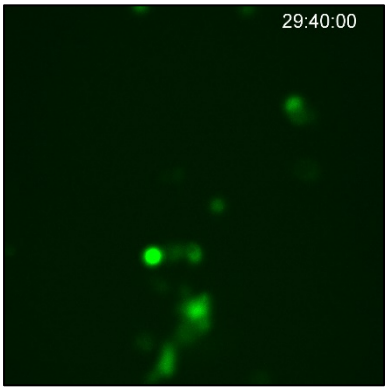
TMEFF2↑



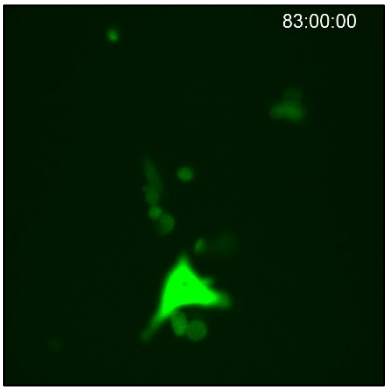
(e.ii)



(f.ii)

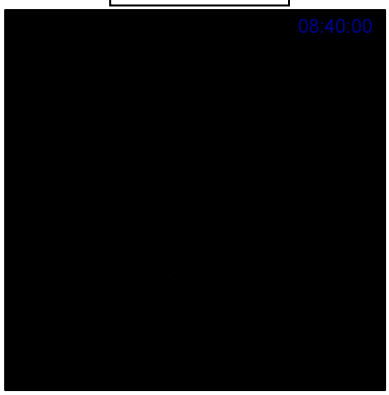


(g.ii)

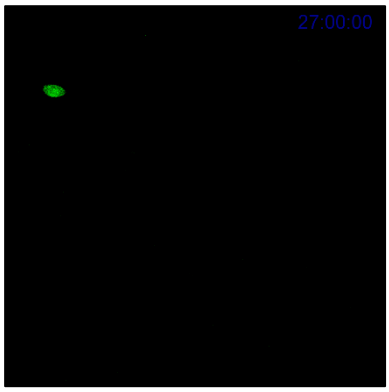


(h.ii)

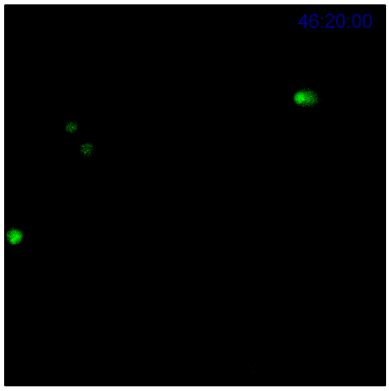
RIP3K↑



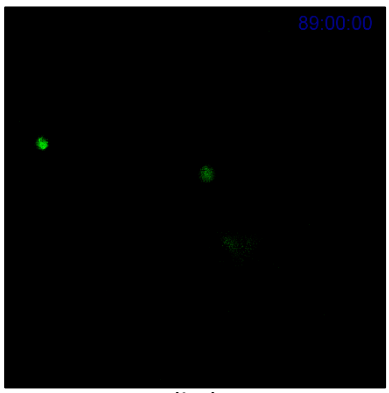
(i.ii)



(j.ii)

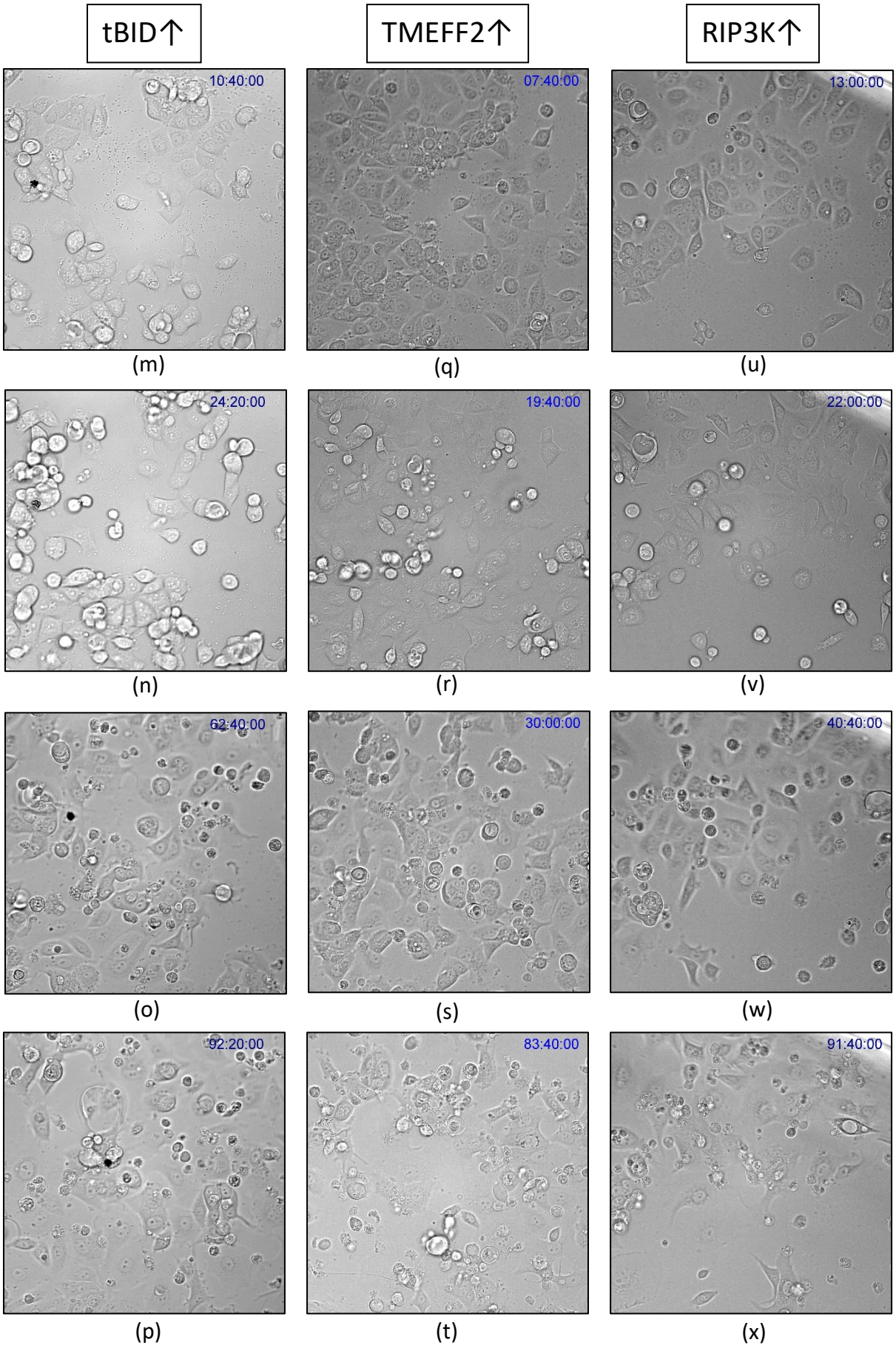


(k.ii)

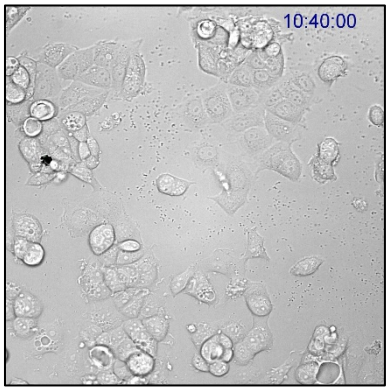


(l.ii)

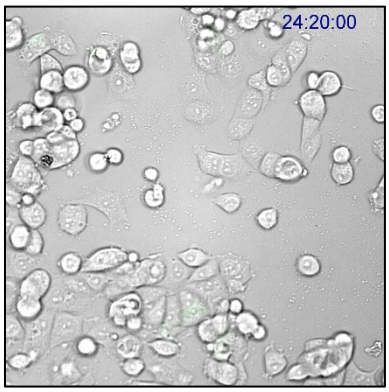
***Figure 5.5*: Time-lapse micrographs of the MCF-7 cells: Imaging Field 1:** MCF-7 cells were co-transfected with GFP and either TMEFF2, tBID or RIP3K (separately) in ratio of 1:4 in an Ibidi® 35mm glass bottom dish coated with poly L-lysine. A four-way divider insert was placed in the dish before seeding the cells. Imaging was started 6 hours post-transfection. Images were captured at 20 minutes interval until 92 hours post-transfection. Cells were kept at 37°C and 5.0 % CO_2 in high relative humidity during imaging. Image processing was performed using Fiji – 'batteries-included' image processing package. Imaging Field 1: (a) Bright Field: GFP + tBID overexpressing, before the onset of cell death. (b) Bright Field: GFP + tBID overexpressing, early cell death. (c) Bright Field: GFP + tBID overexpressing, mid cell death. (d) Bright Field: GFP + tBID overexpressing, late cell death. (e) Bright Field: GFP + TMEFF2 overexpressing, before the onset of cell death. (f) Bright Field: GFP + TMEFF2 overexpressing, early cell death. (g) Bright Field: GFP + TMEFF2 overexpressing, mid cell death. (h) Bright Field: GFP + TMEFF2 overexpressing, late cell death. (i) Bright Field: GFP + RIPK3 overexpressing, before the onset of cell death. (j) Bright Field: GFP + RIPK3 overexpressing, early cell death. (k) Bright Field: GFP + RIPK3 overexpressing, mid cell death. (l) Bright Field: GFP + RIPK3 overexpressing, late cell death. (a.i) Bright Field and GFP overlay: GFP + tBID overexpressing, before the onset of cell death. (b.i) Bright Field and GFP overlay: GFP + tBID overexpressing, early cell death. (c.i) Bright Field and GFP overlay: GFP + tBID overexpressing, mid cell death. (d.i) Bright Field and GFP overlay: GFP + tBID overexpressing, late cell death. (e.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, before the onset of cell death. (f.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, early cell death. (g.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, mid cell death. (h.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, late cell death. (i.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, before the onset of cell death. (j.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, early cell death. (k.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, mid cell death. (l.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, late cell death. (a.ii) GFP: GFP + tBID overexpressing, before the onset of cell death. (b.ii) GFP: GFP + tBID overexpressing, early cell death. (c.ii) GFP: GFP + tBID overexpressing, mid cell death. (d.ii) GFP: GFP + tBID overexpressing, late cell death. (e.ii) GFP: GFP + TMEFF2 overexpressing, before the onset of cell death. (f.ii) GFP: GFP + TMEFF2 overexpressing, early cell death. (g.ii) GFP: GFP + TMEFF2 overexpressing, mid cell death. (h.ii) GFP: GFP + TMEFF2 overexpressing, late cell death. (i.ii) GFP: GFP + RIPK3 overexpressing, before the onset of cell death. (j.ii) GFP: GFP + RIPK3 overexpressing, early cell death. (k.ii) GFP: GFP + RIPK3 overexpressing, mid cell death. (l.ii) GFP: GFP + RIPK3 overexpressing, late cell death. NB: Imaging was performed in collaboration with Dr Adam Cawte.


tBID↑
TMEFF2↑
RIP3K↑
10:40:00
07:40:00
13:00:00
(m)
(q)
(u)
24:20:00
19:40:00
22:00:00
(n)
(r)
(v)
62:40:00
30:00:00
40:40:00
(o)
(s)
(w)
92:20:00
83:40:00
91:40:00
(p)
(t)
(x)

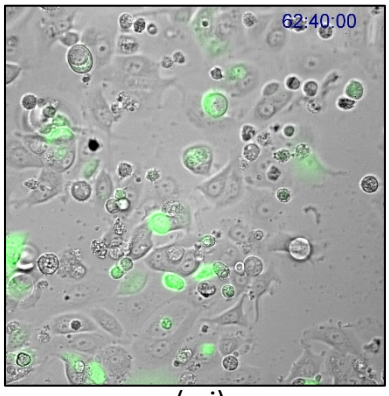
tBID↑



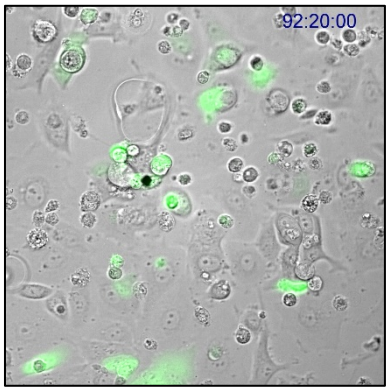
(m.i)



(n.i)

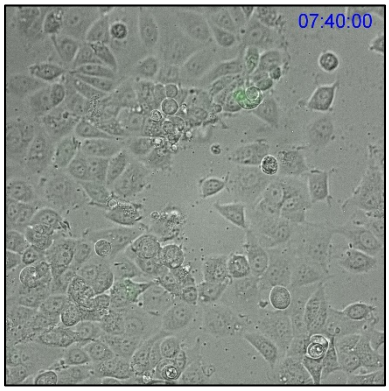


(o.i)

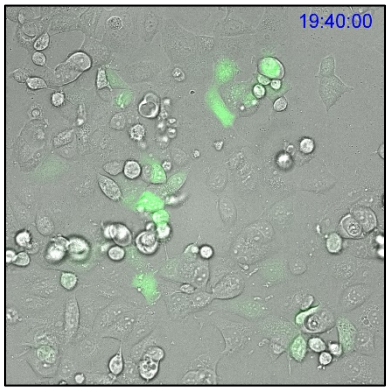


(p.i)

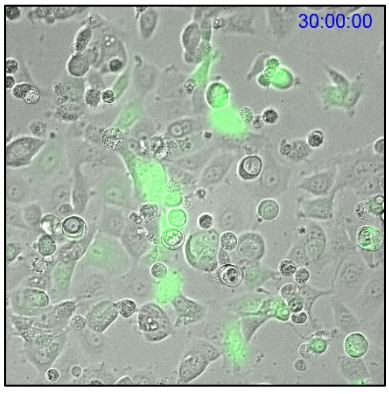
TMEFF2↑



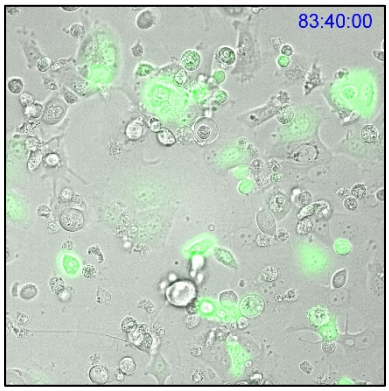
(q.i)



(r.i)

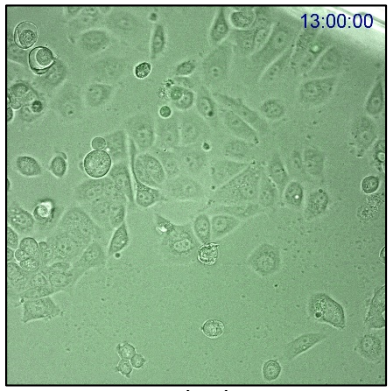


(s.i)

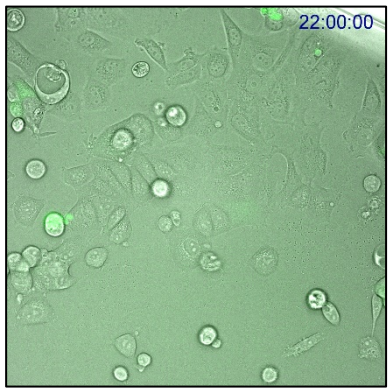


(t.i)

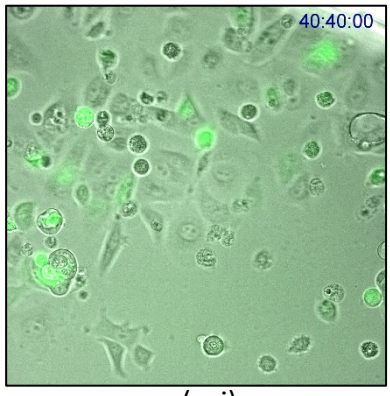
RIP3K↑



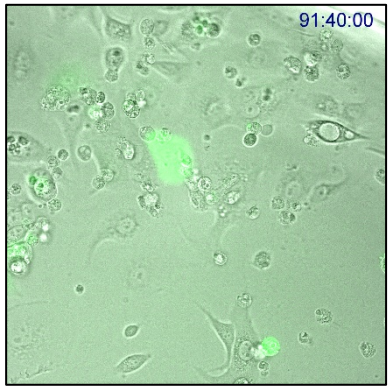
(u.i)



(v.i)



(w.i)



(x.i)

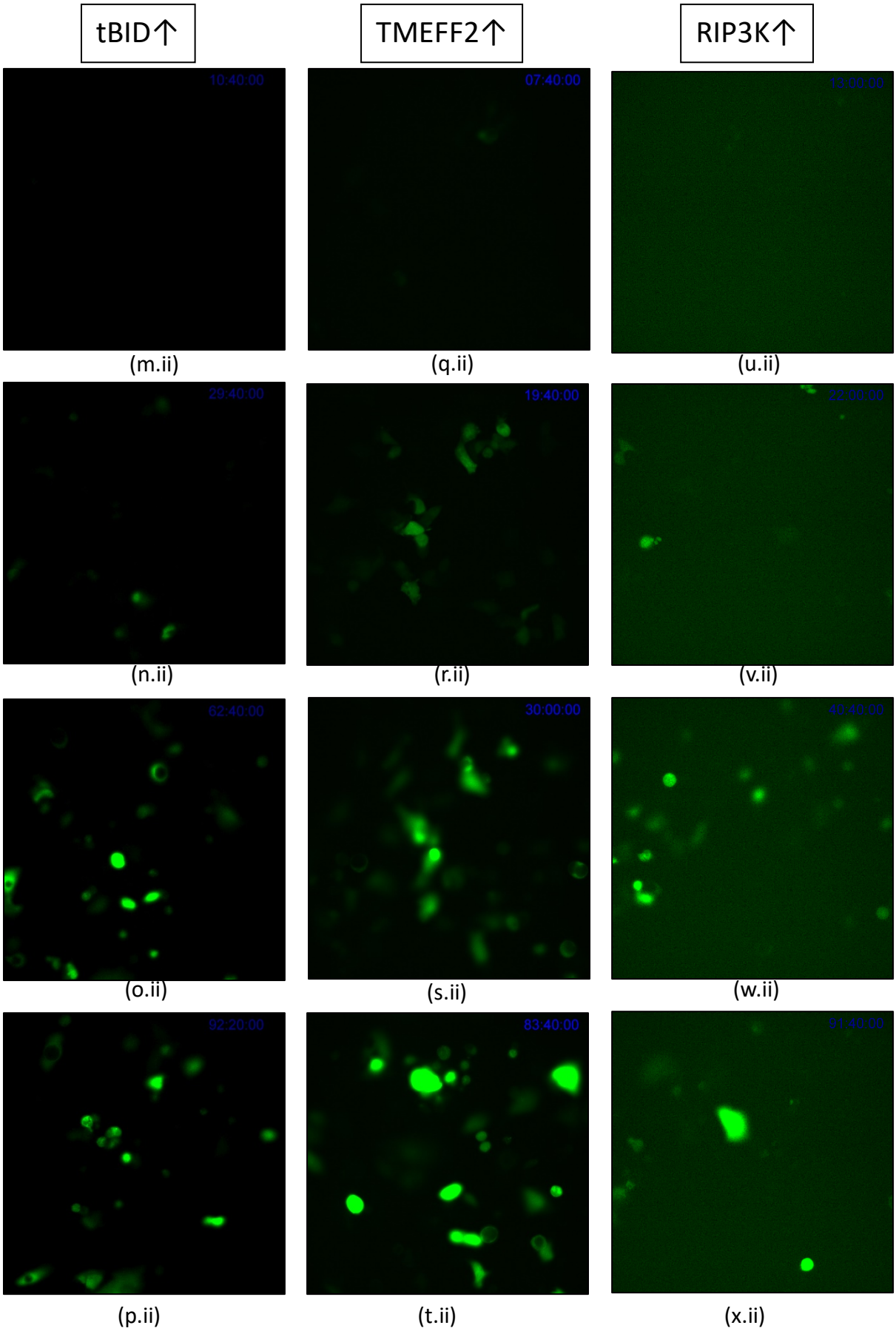

tBID↑
TMEFF2↑
RIP3K↑
10:40:00
07:40:00
13:00:00
(m.ii)
(q.ii)
(u.ii)
29:40:00
19:40:00
22:00:00
(n.ii)
(r.ii)
(v.ii)
62:40:00
30:00:00
40:40:00
(o.ii)
(s.ii)
(w.ii)
92:20:00
83:40:00
91:40:00
(p.ii)
(t.ii)
(x.ii)

***Figure 5.5*: Time-lapse micrographs of the MCF-7 cells: Imaging Field 2:** MCF-7 cells were co-transfected with GFP and either TMEFF2, tBID or RIP3K (separately) in ratio of 1:4 in an Ibidi® 35mm glass bottom dish coated with poly L-lysine. A four-way divider insert was placed in the dish before seeding the cells. Imaging was started 6 hours post-transfection. Images were captured at 20 minutes interval until 92 hours post-transfection. During this time, cells were kept at 37°C and 5.0 % CO_2 in high relative humidity during imaging. Image processing was performed using Fiji – 'batteries-included' image processing package. (m) Bright Field: GFP + tBID overexpressing, before the onset of cell death. (n) Bright Field: GFP + tBID overexpressing, early cell death. (o) Bright Field: GFP + tBID overexpressing, mid cell death. (p) Bright Field: GFP + tBID overexpressing, late cell death. (q) Bright Field: GFP + TMEFF2 overexpressing, before the onset of cell death. (r) Bright Field: GFP + TMEFF2 overexpressing, early cell death. (s) Bright Field: GFP + TMEFF2 overexpressing, mid cell death. (t) Bright Field: GFP + TMEFF2 overexpressing, late cell death. (u) Bright Field: GFP + RIPK3 overexpressing, before the onset of cell death. (v) Bright Field: GFP + RIPK3 overexpressing, early cell death. (w) Bright Field: GFP + RIPK3 overexpressing, mid cell death. (x) Bright Field: GFP + RIPK3 overexpressing, late cell death. (m.i) Bright Field and GFP overlay: GFP + tBID overexpressing, before the onset of cell death. (n.i) Bright Field and GFP overlay: GFP + tBID overexpressing, early cell death. (o.i) Bright Field and GFP overlay: GFP + tBID overexpressing, mid cell death. (p.i) Bright Field and GFP overlay: GFP + tBID overexpressing, late cell death. (q.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, before the onset of cell death. (r.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, early cell death. (s.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, mid cell death. (t.i) Bright Field and GFP overlay: GFP + TMEFF2 overexpressing, late cell death. (u.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, before the onset of cell death. (v.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, early cell death. (w.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, mid cell death. (x.i) Bright Field and GFP overlay: GFP + RIPK3 overexpressing, late cell death. (m.ii) GFP: GFP + tBID overexpressing, before the onset of cell death. (n.ii) GFP: GFP + tBID overexpressing, early cell death. (o.ii) GFP: GFP + tBID overexpressing, mid cell death. (p.ii) GFP: GFP + tBID overexpressing, late cell death. (q.ii) GFP: GFP + TMEFF2 overexpressing, before the onset of cell death. (r.ii) GFP: GFP + TMEFF2 overexpressing, early cell death. (s.ii) GFP: GFP + TMEFF2 overexpressing, mid cell death. (t.ii) GFP: GFP + TMEFF2 overexpressing, late cell death. (u.ii) GFP: GFP + RIPK3 overexpressing, before the onset of cell death. (v.ii) GFP: GFP + RIPK3 overexpressing, early cell death. (w.ii) GFP: GFP + RIPK3 overexpressing, mid cell death. (x.ii) GFP: GFP + RIPK3 overexpressing, late cell death. NB: Imaging was performed by Dr Adam Cawte.

***Table 5.2*:** Hyperlinks for the time-lapse movies of the MCF-7 cells. NB: the videos are 'private', in order to view them, reader has to sign in on YouTube (YouTube commonly accepts institutional details for login)

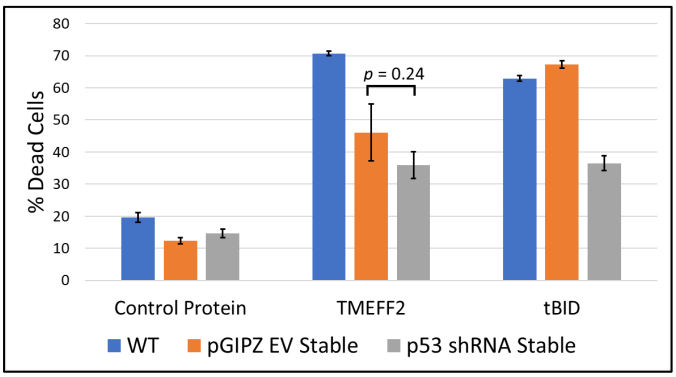
Condition[a]	Hyperlink
tBID overexpressing MCF-7 cells: Phase Contrast only: Field 1	https://youtu.be/teICdo72XmQ
tBID overexpressing MCF-7 cells: Composite: Field 1	https://youtu.be/pAMuULhe5_0

tBID overexpressing MCF-7 cells: Phase contrast only: Field 2	https://youtu.be/dmJw85_H7wQ
tBID overexpressing MCF-7 cells: Composite: Field 2	https://youtu.be/vO07TDmRGG4
TMEFF2 overexpressing MCF-7 cells: Phase Contrast only: Field 1	https://youtu.be/jc_7swgE0u4
TMEFF2 overexpressing MCF-7 cells: Composite: Field 1	https://youtu.be/Zwqbef1kExU
TMEFF2 overexpressing MCF-7 cells: Phase contrast only: Field 2	https://youtu.be/IWMJKUgk09U
TMEFF2 overexpressing MCF-7 cells: Composite: Field 2	https://youtu.be/MZdGAEuZlQ0
RIP3K overexpressing MCF-7 cells: Phase Contrast only: Field 1	https://youtu.be/xPIRk9ha50c
RIP3K overexpressing MCF-7 cells: Composite: Field 1	https://youtu.be/JyKu7_pt3uw
RIP3K overexpressing MCF-7 cells: Phase contrast only: Field 2	https://youtu.be/_g2dyPyPgJk
RIP3K overexpressing MCF-7 cells: Composite: Field 2	https://youtu.be/HglqcRx8Z4U

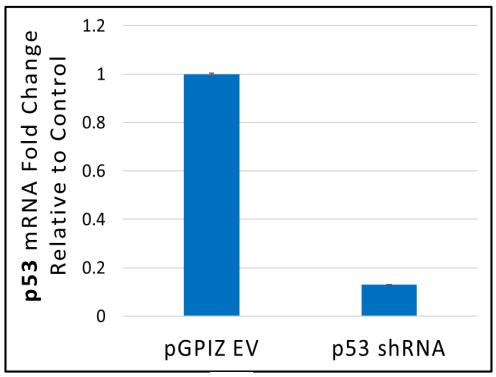
[a] MCF-7 cells were co-transfected with GFP and either TMEFF2, tBID or RIP3K (separately) in ratio of 1:4 in an Ibidi® 35mm glass bottom dish coated with poly L-lysine. A four-way divider insert was placed in the dish before seeding the cells. Imaging was started 6 hours post-transfection. Images were captured at 20 minutes interval till 92 hours post-transfection. Cells were kept at 37°C and 5.0 % CO_2 in high relative humidity during imaging. Image processing was performed using Fiji – 'batteries-included' image processing package. NB: movies present the data for 72 hours post-transfection only.

5.2.3 TMEFF2-induced apoptosis is independent of p53 activity

Observing an increase in the protein levels of p53 and those of its downstream transcription targets PUMA and MDM2 in TMEFF2-transfected MCF-7 cells led me to hypothesize that TMEFF2 could be causing apoptosis in MCF-7 cells through the upregulation and/or stabilization of p53. This premise was further supported by the fact that MCF-7 cells contain the unmutated copy of p53 [202]. In addition, the dependence of almost all of the previously reported anticancer genes on p53 had been investigated. Therefore, I tested for the difference in TMEFF2-induced apoptosis upon p53 knockdown. For this purpose, I generated p53 shRNA stable MCF-7 cells through puromycin selection. The plasmid harbouring the shRNA, pGIPZ (Dharmacon™), contained a high expression variant of GFP (TurboGFP™). As with the previous experiment, I employed TMRE/DRAQ7™ staining to measure cell death. As the emission spectra of GFP, TMRE and DRAQ7™ differ substantially (no overlap detected), it allowed me to gate and select only the stably transfected green cells for further analysis. Results are shown in figure 5.6. Upon TMEFF2 transfection, no significant change was observed in apoptosis between the p53 knockdown cells and those stably transfected with pGIPZ empty vector.


% Dead Cells
80
70
60
50
40
30
20
10
0
p = 0.24
Control Protein
TMEFF2
tBID
WT
pGIPZ EV Stable
p53 shRNA Stable

(a)


p53 mRNA Fold Change Relative to Control
1.2
1
0.8
0.6
0.4
0.2
0
pGPIZ EV
p53 shRNA

(b)

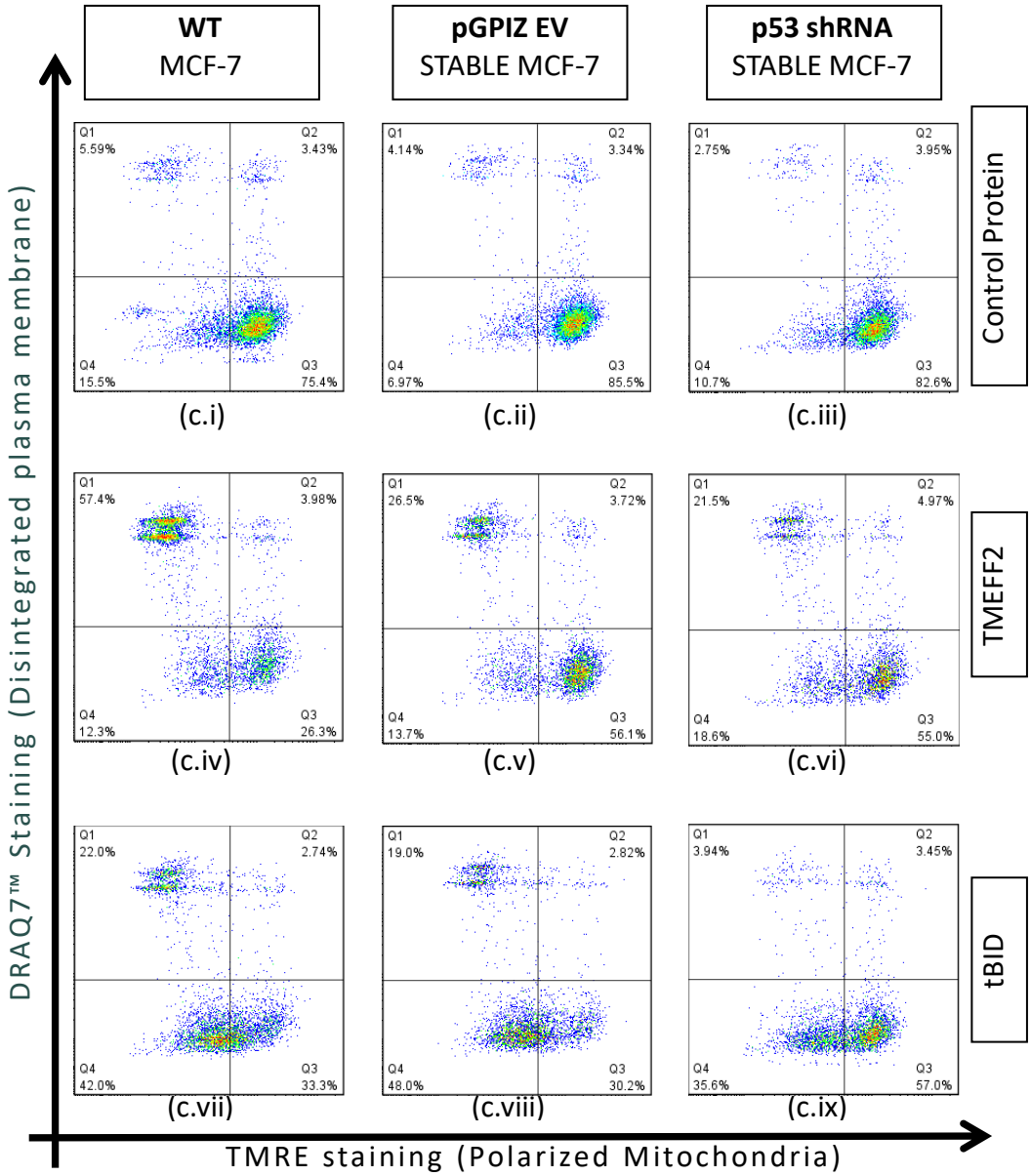

WT
MCF-7
pGPIZ EV
STABLE MCF-7
p53 shRNA
STABLE MCF-7
Control Protein
TMEFF2
tBID
DRAQ7™ Staining (Disintegrated plasma membrane)
TMRE staining (Polarized Mitochondria)
Q1 5.59%
Q2 3.43%
Q4 15.5%
Q3 75.4%
(c.i)
Q1 4.14%
Q2 3.34%
Q4 6.97%
Q3 85.5%
(c.ii)
Q1 2.75%
Q2 3.95%
Q4 10.7%
Q3 82.6%
(c.iii)
Q1 57.4%
Q2 3.98%
Q4 12.3%
Q3 26.3%
(c.iv)
Q1 26.5%
Q2 3.72%
Q4 13.7%
Q3 56.1%
(c.v)
Q1 21.5%
Q2 4.97%
Q4 18.6%
Q3 55.0%
(c.vi)
Q1 22.0%
Q2 2.74%
Q4 42.0%
Q3 33.3%
(c.vii)
Q1 19.0%
Q2 2.82%
Q4 48.0%
Q3 30.2%
(c.viii)
Q1 3.94%
Q2 3.45%
Q4 35.6%
Q3 57.0%
(c.ix)

***Figure 5.6*: TMEFF2 induced apoptosis in MCF-7 cells upon p53 knockdown.** (a) Cell death in wild type MCF-7 cells, p53 shRNA stable MCF-7 cells and MCF-7 cells stably transfected with pGIPZ empty vector upon TMEFF2 overexpression, detected using TMRE and DRAQ7™ staining through flow cytometry. Cells were transiently transfected with TMEFF2, luciferase (negative 'control') and tBID (positive control). Cells were harvested 48 hours post-transfection, stained with TMRE/DRAQ7™ and subjected to flow cytometry. Only GFP positive cells were gated for further measurements of TMRE and DRAQ7™ fluorescence, except in the wild type cells. Bars represent average from three different wells of a 24-well plate each with independently transfected population of cells. Error-bars represent the +/- standard deviation. p value was calculated using two-sample t-test assuming unequal variances. (b) qRT-PCR results for the relative p53 mRNA fold change between pools of MCF-7 cells stably transfected with pGIPZ empty vector and p53 shRNA according to ΔΔCt method using L19 as the housekeeping gene. Bars represent the average form 3 PCR-reactions. Error bars represent +/- standard deviation. (c.i) to (c.ix) Representative single scatter plots for TMRE (x-axis) and DRAQ7™ (y-axis) staining used for data presented in graph (a). (c.i) Wild type MCF-7 cells transiently transfected with Luciferase 'control'. (c.ii) pGIPZ empty vector stable MCF-7 cells transiently transfected with Luciferase 'control'. (c.iii) p53 shRNA stable MCF-7 cells transiently transfected with luciferase 'control'. (c.iv) Wild type MCF-7 cells transiently transfected with TMEFF2. (c.v) pGIPZ empty vector stable MCF-7 cells transiently transfected with TMEFF2. (c.vi) p53 shRNA stable MCF-7 cells transiently transfected with TMEFF2. (c.vii) Wild type MCF-7 cells transiently transfected with tBID. (c.viii) pGIPZ empty vector stable MCF-7 cells transiently transfected with tBID. (c.ix) p53 shRNA stable MCF-7 cells transiently transfected with tBID.

However, when compared to the wild type MCF-7 cells, both p53 knockdown cells and those stably transfected with pGIPZ empty vector showed a decrease in cell death upon TMEFF2 transfection. This change could be attributed to two factors: (i) the puromycin selection conferred a general resistance to cell death in the cells and (ii) a different gating strategy was employed for the wild type and stably transfected cells, as for the latter group only the green cells were selected to measure the TMRE and DRAQ-7 fluorescence. I used luciferase as negative control and tBID as the positive control for this experiment. Unexpectedly, the cell-death inducing capacity of tBID decreased substantially in p53 knockdown cells as compared to those stably transfected with the pGIPZ empty vector. A further confirmation was therefore needed for the TMEFF2-induced apoptosis being independent of the p53 activity. For this purpose, I used a potent small molecule inhibitor of p53, Pifithrin-alpha (PFT-α) [203]. Up to 15 μM PFT-α treatment for 1 hour had been reported for a successful inhibition of p53 activity in MCF-7 cells [204, 205, 206]. I added 20 μM PFT-α to the media containing MCF-7 cells 4 hours post-transfection; cells were harvested 48 hours post-transfection and subjected to the TMRE plus

DRAQ7™ staining and flow cytometry. Results in figure 5.7 depict no change whatsoever in cell death between the cells treated with PFT-α and those treated with the vehicle (DMSO).

a.

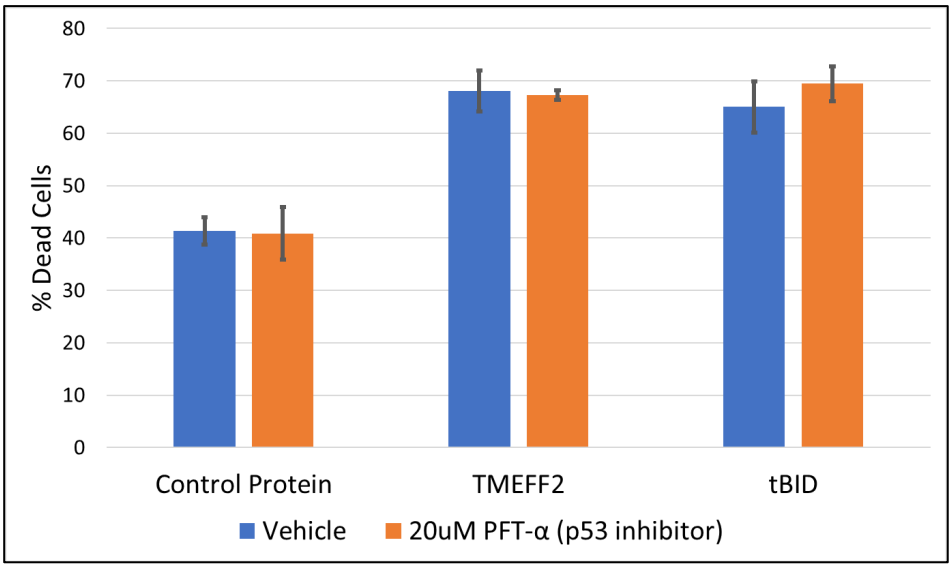


***Figure 5.7* (a): TMEFF2 induced apoptosis in MCF-7 cells upon p53 inhibition.** MCF-7 cells were transiently transfected with TMEFF2, luciferase (negative 'control') and tBID (positive control). 4 hours post-transfection, Opti-MEM® media was replaced with DMEM +10% FBS which contained either 20 μM PFT-α or DMSO 'vehicle'. Cells were harvested 48 hours post-transfection, stained with TMRE/DRAQ7™ and subjected to flow cytometry. Bars represent average from three different wells of a 24-well plate each with independently transfected population of cells. Error-bars represent the +/- standard deviation.

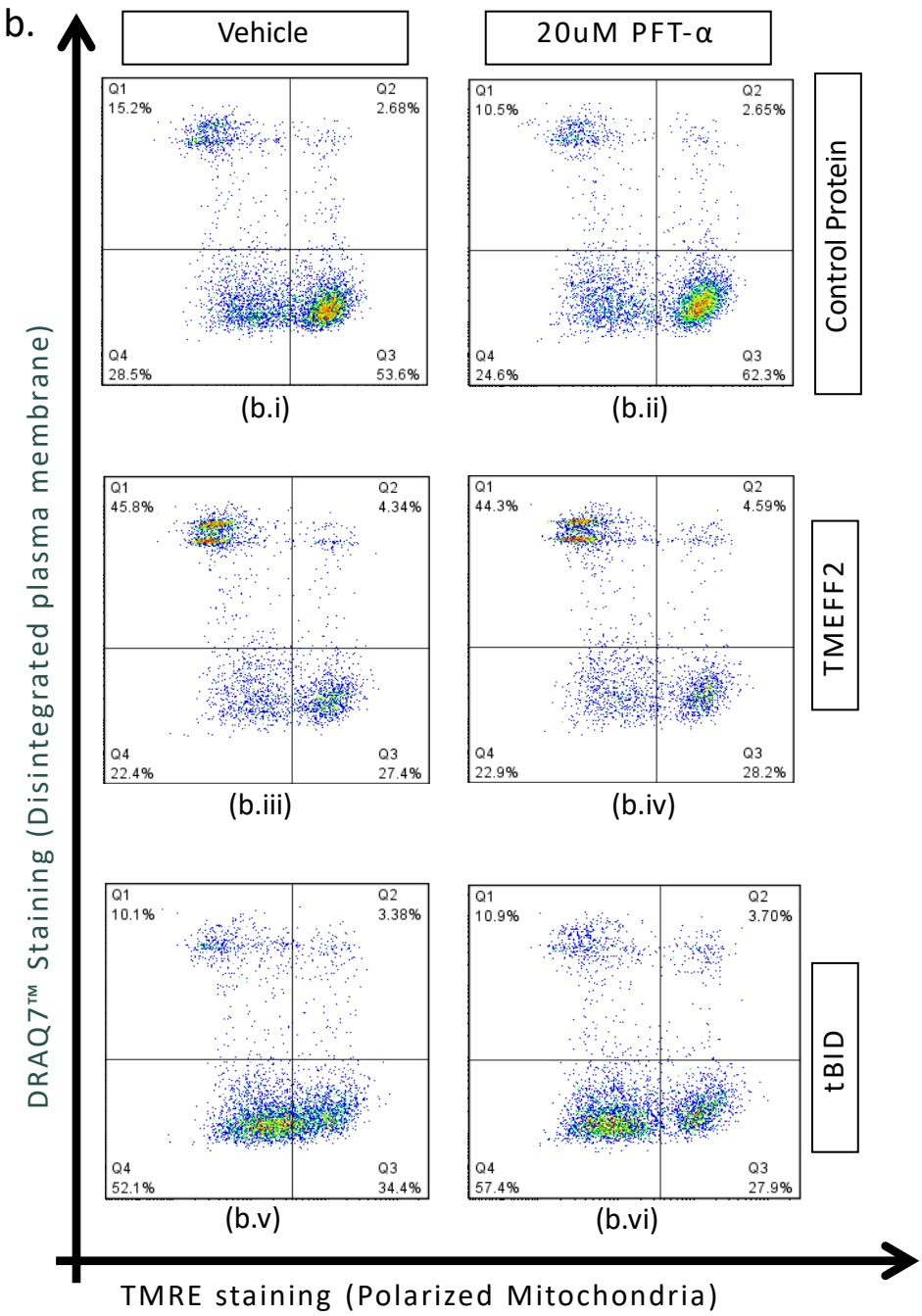


***Figure 5.7 (b)*: TMRE (x-axis) and DRAQ7™ (y-axis) staining upon TMEFF2 overexpression and treatment with p53 inhibitor.** Representative single scatter plots for the data presented in Figure 5.7(a). (b.i) MCF-7 cells transfected with luciferase 'control' and treated with DMSO 'vehicle'. (b.ii) MCF-7 cells transfected with luciferase 'control' and treated with 20 μM PFT-α. (b.iii) MCF-7 cells transfected with TMEFF2 and treated with DMSO 'vehicle'. (b.iv) MCF-7 cells transfected with TMEFF2 and treated with 20 μM PFT-α. (b.v) MCF-7 cells transfected with tBID and treated with DMSO 'vehicle'. (b.vi) MCF-7 cells transfected with tBID and treated with 20 μM PFT-α.

5.2.4 Intracellular localization of TMEFF2 in MCF-7 cells

Although TMEFF2 is a type-I transmembrane protein, numerous studies have reported several different intracellular locations for TMEFF2 (table 5.3). In order to gain insight regarding the possible mechanism behind the apoptosis induced by TMEFF2 in MCF-7 cells, I investigated its intracellular localization upon overexpression. Using the same TMEFF2 antibody that was used to confirm its overexpression through immunoblotting, I performed immunofluorescence in MCF-7 cells (after TMEFF2 overexpression through transient transfection). As a decrease in mitochondrial membrane potential is noted upon TMEFF2 overexpression, I also examined if TMEFF2 colocalized with mitochondria using an antibody against the mitochondrial ATP synthase β subunit (ATPB). Moreover, as Thomas Green *et al.* [190] had reported colocalization of TMEFF2 with the cytoskeleton during induction of its antiproliferative effects in 22Rv1 cells, I decided to probe for the colocalization of TMEFF2 with α-Tubulin in MCF-7 cells. Results are depicted in figure 5.8. Upon overexpression in MCF-7 cells, TMEFF2 gave a perinuclear, punctate cytoplasmic signal rather than a membranous location. Also, just as with the immunoblotting, the endogenous expression levels of TMEFF2 were not detectable. However, a rather aesthetically pleasing image with thick fluorescence strings emerged when the intensity of the emission channel was increased disproportionately (data now shown). No colocalization of TMEFF2 was observed either with mitochondria or α-Tubulin (figure 5.9). Interestingly, a decrease in mitochondrial staining was observed only in cells overexpressing TMEFF2. Conversely, I also investigated if there is any difference in intracellular localization of TMEFF2 between MCF-7 cells, COS-7 cells and CV-1 cells. Results are presented in supplementary figure S.11. TMEFF2 also gave a perinuclear, punctate cytoplasmic appearance in COS-7 cells. I did not observe either endogenous or overexpressed TMEFF2 in CV-1 cells (supplementary figure S.11).

***Table 5.3*:** Reported Intracellular localization of the TMEFF2.

Author and Reference	Cancer type	Cells	Antibody used	Intracellular Location
Daniel E.H. Afar *et al.* [167]	Prostate	PC3	Pr1#19	Membranous and punctate cytoplasmic upon internalization.
Xiao-Yan-Zhao *et al.* [180]	Prostate	PC3	2H8	Membranous and punctate cytoplasmic upon internalization, localizes with acidic organelles.
Hyun-Seok Hong *et al.* [135]	NONE	MC6	48G2	Perinuclear and punctate cytoplasmic, not membranous.
Ryan F. Overcash *et al.* [181]	Prostate	22Rv1	Abcam	Punctate cytoplasmic along the rim of nucleus in immunofluorescence. In cell fractionation, membranous, cytoplasmic and nuclear.
Xiaofei Chen *et al.* [189]	Prostate	HEK293T	Not mentioned	Membranous, cytoplasmic diffused in immunofluorescence but showed mitochondrial pattern when fusion constructs were used.
Thomas Green *et al.* [190]	Prostate	22Rv1	Abcam	Membranous, cytoplasmic, nuclear and colocalizes with beta actin and α-Tubulin.
Steven N. Quayle and Marianne D Sadar [141]	prostate	LnCaP	GFP fusion construct	Membranous and diffuse cytoplasmic for full length but punctate cytoplasmic for short secreted isoform.

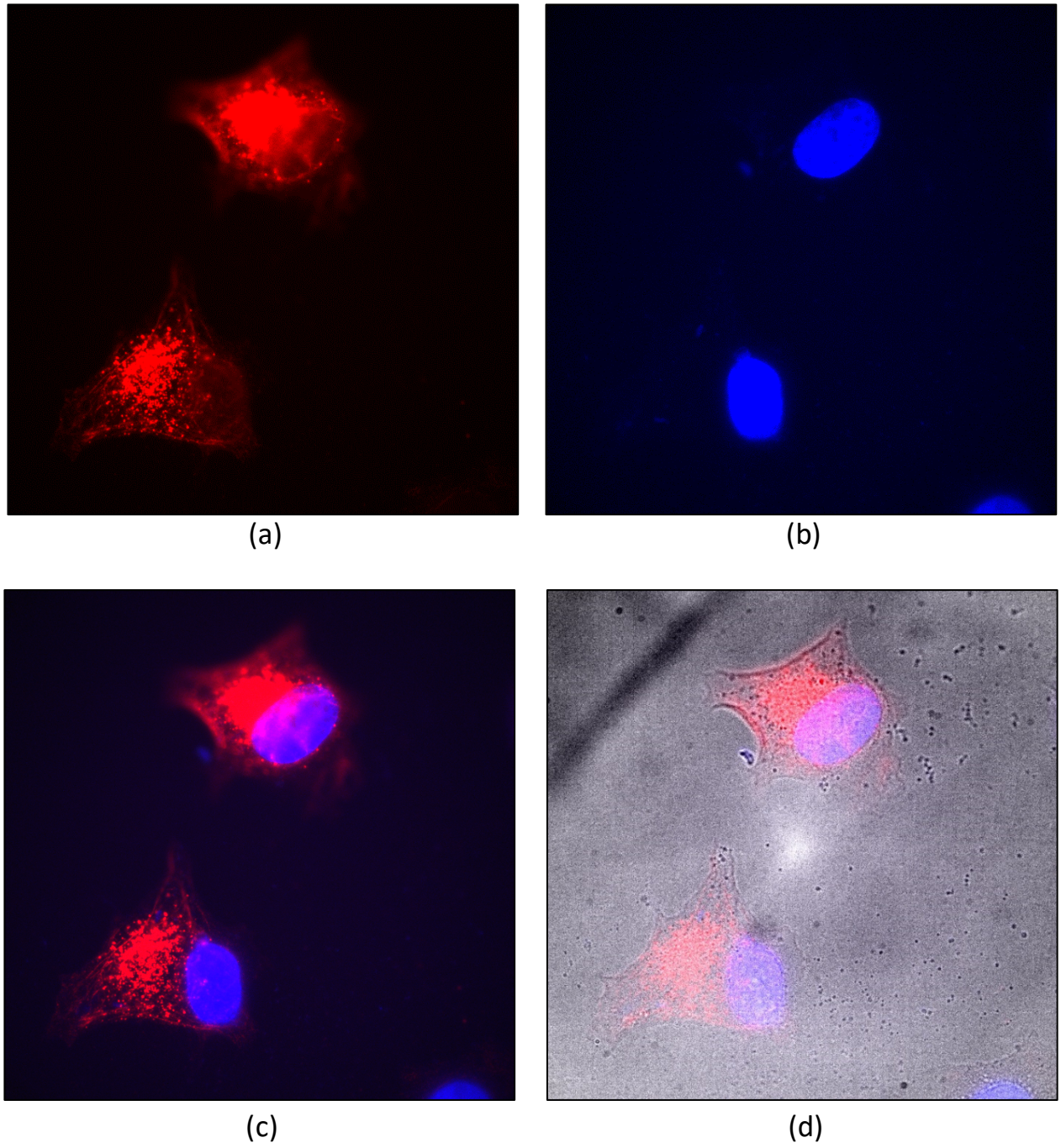
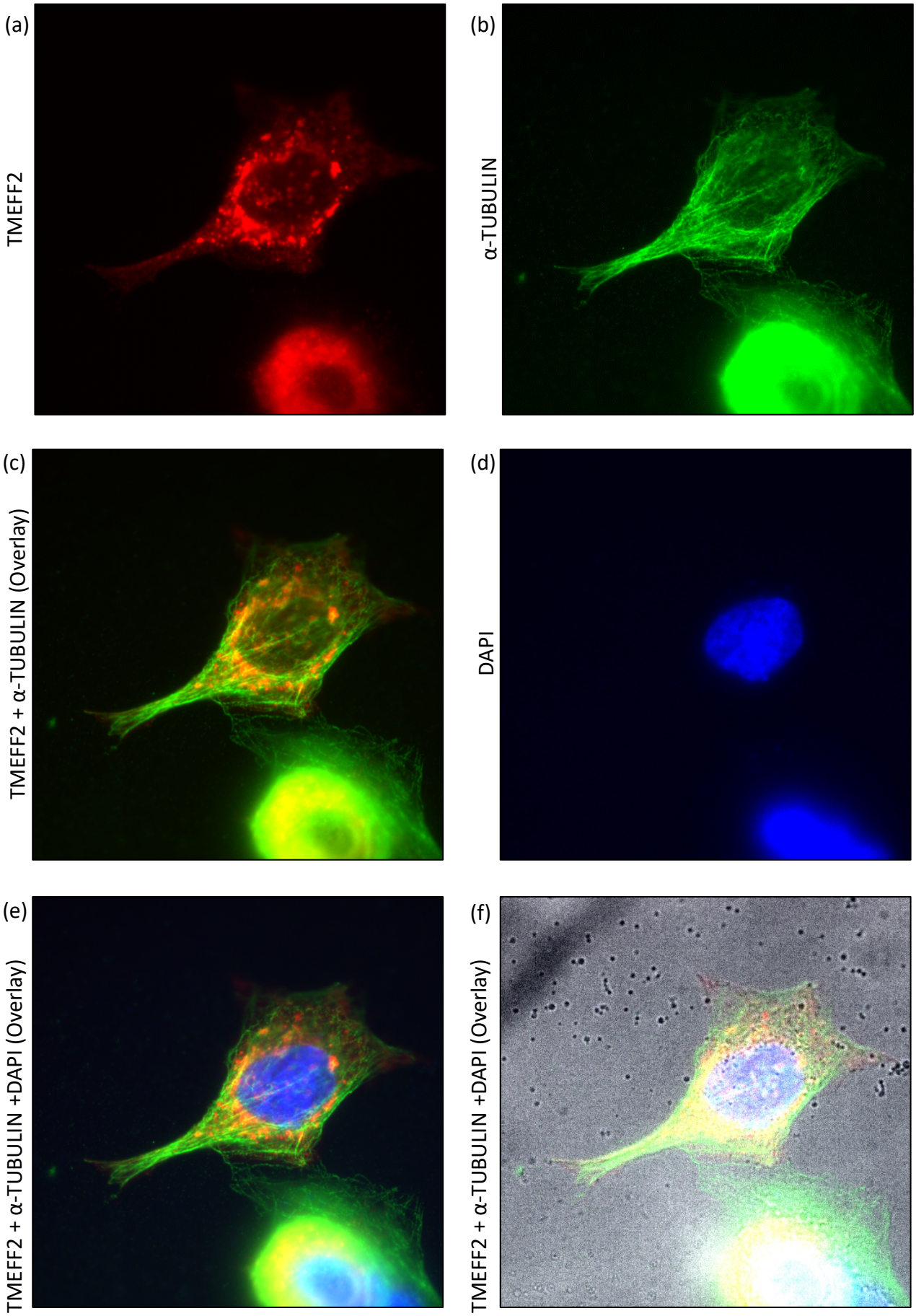


***Figure 5.8*: Intracellular localization of the experimentally overexpressed TMEFF2 in MCF-7 cells**. MCF-7 cells were transiently transfected with TMEFF2 in ibidi® 8-well glass bottom chamber coated with poly L-lysine. Cells were fixed with 4.0% paraformaldehyde 30 hours post-transfection and subjected to immunostaining (see materials and methods). (a) TMEFF2 [Alexa Fluor® 647-Red], (b) DAPI, (c) TMEFF2 and DAPI overlay, (d) TMEFF and DAPI overlay in bright field. Image processing was performed using Fiji – 'batteries-included' image processing package. NB: Microscopy was performed by Dr Adam Cawte.


(a)
TMEFF2
(b)
α-TUBULIN
(c)
TMEFF2 + α-TUBULIN (Overlay)
(d)
DAPI
(e)
TMEFF2 + α-TUBULIN +DAPI (Overlay)
(f)
TMEFF2 + α-TUBULIN +DAPI (Overlay)

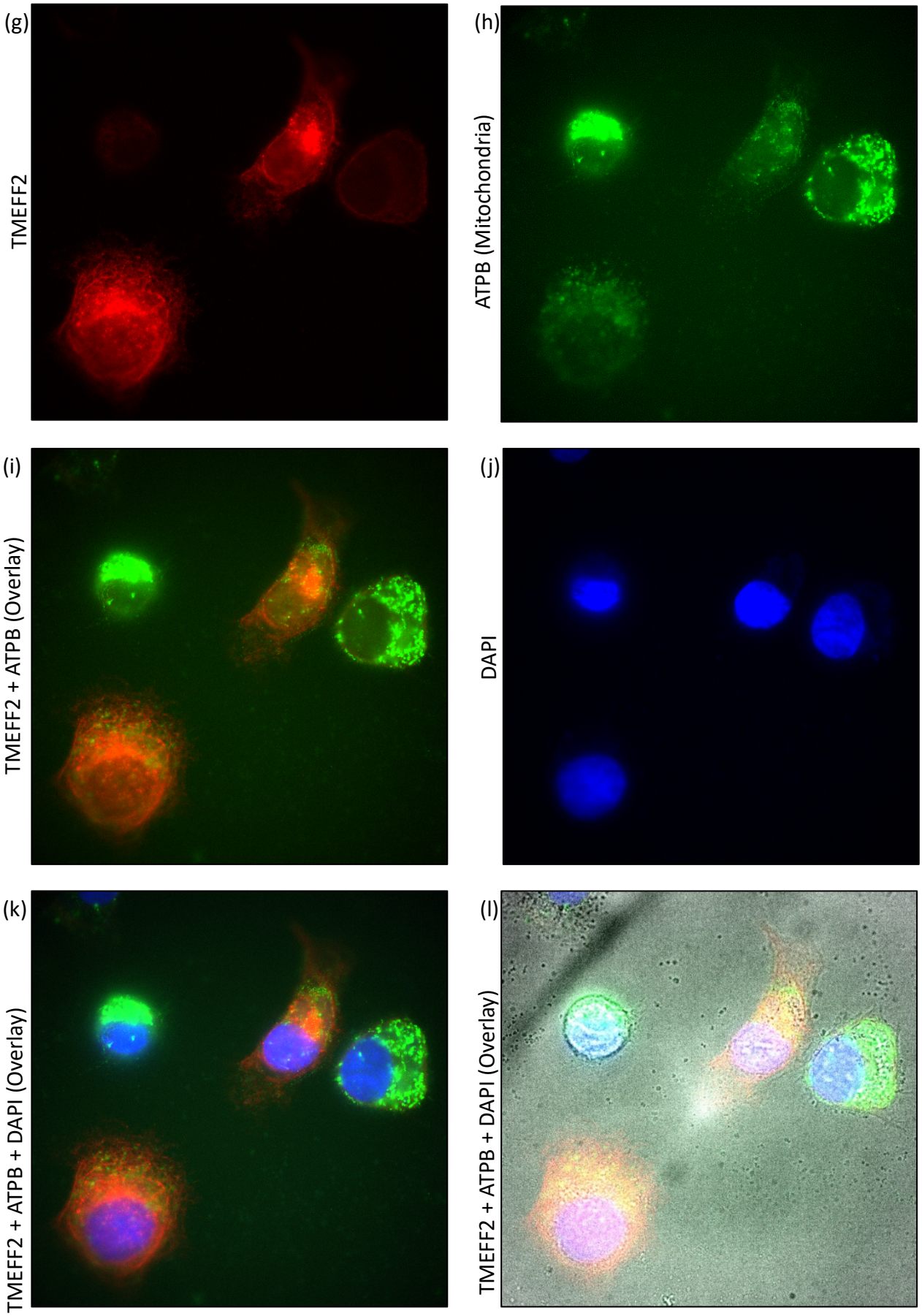
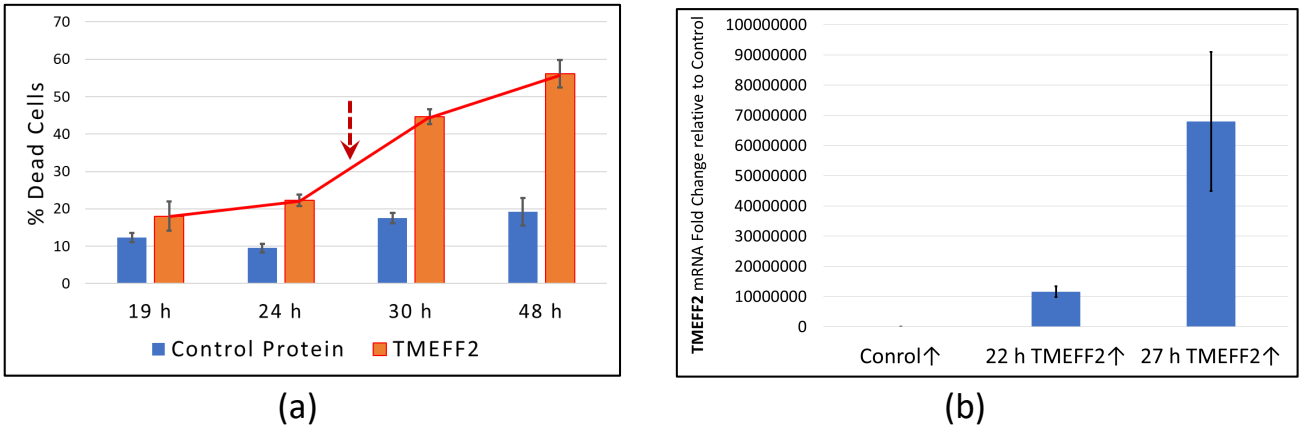

(g)
TMEFF2
(h)
ATPB (Mitochondria)
(i)
TMEFF2 + ATPB (Overlay)
(j)
DAPI
(k)
TMEFF2 + ATPB + DAPI (Overlay)
(l)
TMEFF2 + ATPB + DAPI (Overlay)

***Figure 5.9*: TMEFF2, α-Tubulin and ATPB (mitochondrial) Immunofluorescence in MCF-7 cells**. MCF-7 cells were transiently transfected with TMEFF2 in ibidi® 8-well glass bottom chamber coated with poly L-lysine. Cells were fixed with 4.0% paraformaldehyde 30 hours post-transfection and subjected to immunostaining (see materials and methods). Image processing was performed using Fiji – 'batteries-included' image processing package. (a) TMEFF2 [Alexa Fluor® 647-Red]. (b) α-Tubulin [Alexa Fluor® 488-Green]. (c) TMEFF2 and α-Tubulin overlay. (d) DAPI. (e) TMEFF2, α-Tubulin and DAPI overlay. (f) TMEFF2, α-Tubulin and DAPI overlay in bright field. (g) TMEFF2 [Alexa Fluor® 647-Red]. (h) Mitochondrial ATPB [Alexa Fluor® 488-Green]. (i) TMEFF2 and Mitochondrial ATPB overlay. (j) DAPI. (k) TMEFF2, Mitochondrial ATPB and DAPI overlay. (l) TMEFF2, Mitochondrial ATPB and DAPI overlay in bright field. NB: microscopy was performed by Dr Adam Cawte.

5.2.5 Preparation for transcriptomics analysis

As it is evident from the TMEFF2 literature review presented at the beginning of this chapter, the previously reported functions of TMEFF2 are very diverse. It was not feasible to investigate the link between every previously reported function of TMEFF2 and the apoptosis induction in MCF-7 cells. Therefore, I decided to employ the transcriptomic analysis and the subsequent bioinformatics pathway analyses. Firstly, the timing of harvesting the cells after transfection was of utmost importance. If I had harvested the cells too early, I would have compromised the degree of overexpression. In case I had harvested the cells too late, it would have resulted in extensive cell death which would firstly compromise the quality of the extracted RNA, and secondly the functional impact of TMEFF2 would have been masked by the list of genes involved in the apoptosis. Consequently, I ran first a time course experiment where cells were harvested at 19, 24, 30 and 48 hours post-transfection and subjected to TMRE/DRAQ7™ staining and flow cytometry. Results are depicted in figure 5.10 (a). Apoptosis progressed rapidly after 24 hours, almost increasing 2.5-fold till 30 hours post-transfection. Between 30 and 48 hours, the apoptosis progressed further by only ~15%. Accordingly, I decided to choose a time point between 24 and 30 hours post-transfection. In order to confirm the optimum overexpression of TMEFF2, I harvested cells at 22 and 27 hours post-transfection and subjected these to qRT-PCR for measuring the relative change in TMEFF2 mRNA. Results in figure 5.10 (b) show a rapid increase in the relative mRNA levels of TMEFF2 from 22 to 27 hours. Firstly, it explained the steep rise in apoptosis between these two time points and, secondly, it confirmed the optimum overexpression of TMEFF2. I therefore chose 26.5 hours post-transfection as the harvesting time for the further transcriptomic analysis. It is important to mention that in figure

5.10 (b), the fold change appears exceptionally high because the endogenous TMEFF2 expression level in MCF-7 cells is negligible.



(a) (b)

***Figure 5.10*: Choosing the optimum time point for harvesting cells for the transcriptomics analysis.** (a) apoptosis: MCF-7 cells were transiently transfected with TMEFF2 and Luciferase (negative 'control') and harvested 19, 24, 30 and 48 hours post-transfection. Cell death was measured through TMRE/DRAQ7™ staining using flow cytometry. Bars represent average from three different wells of a 24-well plate each with independently transfected population of cells. Error-bars represent the +/- standard deviation. Arrowhead represents the putative optimum time point for harvesting. (b) expression level**:** qRT-PCR results for the relative TMEFF2 mRNA fold change in of MCF-7 transfected with luciferase 'control' and TMEFF2. Cells were harvested at 22 and 27 hours post-transfection. Fold change was quantified according to ΔΔCt method using L19 as the housekeeping gene. Bars represent the average from 3 PCR-reactions. Error bars represent +/- standard deviation. NB: the mRNA in 'control' was undetectable, Ct value of 40 was added manually for the analysis.

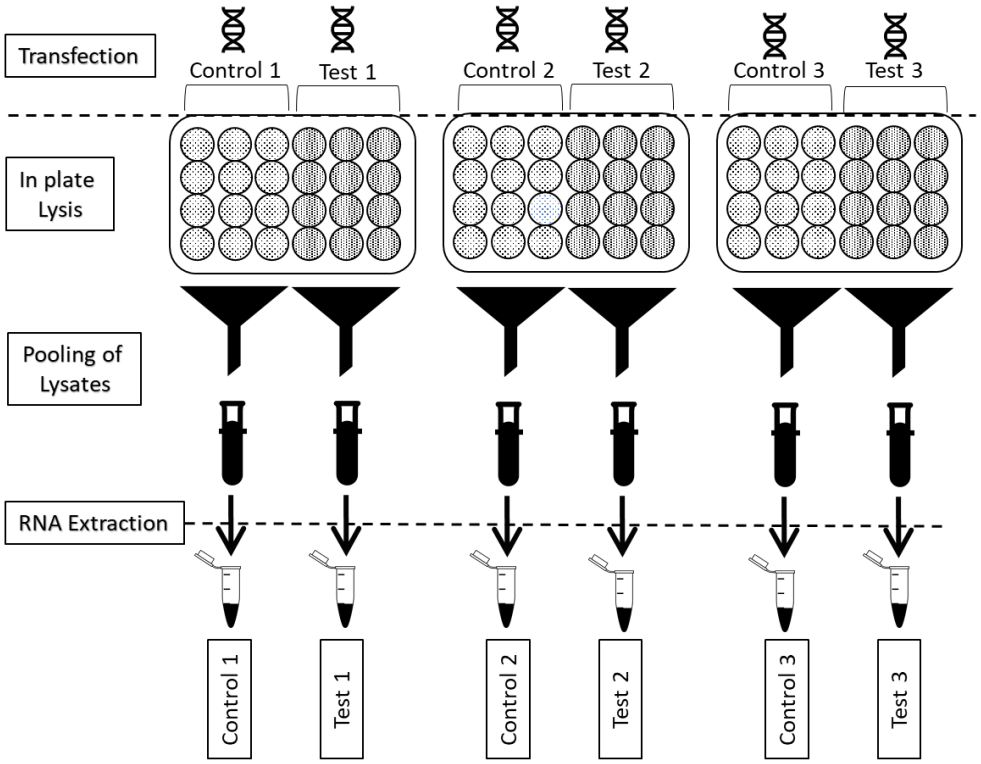


***Figure 5.11*: Schematic representation of the methodology employed in sample preparation for the transcriptomic analysis of TMEFF2 overexpressing MCF-7 cells.** MCF-7 cells were transiently transfected with TMEFF2 and luciferase (negative control) in 24-well plates. Cells were lysed in their respective wells and lysates from 12 wells of the 24-well plate were pooled together to extract one RNA sample. The whole process was repeated in triplicate where each RNA sample was extracted from a pool of 12 independently transfected populations of MCF-7 cells.

I transiently transfected MCF-7 cells with TMEFF2 in 24-well plates and used transient transfection with luciferase in the parallel as negative control. The transiently transfected cells from the 12 wells of the 24-well plate were then pooled together to extract one RNA sample. The whole process was repeated in triplicate – each RNA sample came from a pool of 12 independently transfected populations of cells, with three of these pools for the control and three for the test samples (figure 5.11). Utmost care was taken to maintain the integrity of the extracted RNA and the removal of genomic DNA contamination. For the latter, I used first gDNA Eliminator columns by QIAGEN® and later employed the On-Column DNA digestion using the RNase-Free DNase by QIAGEN®. I confirmed the TMEFF2 overexpression in each sample using qRT-PCR before subjecting them to the microarray processing (figure 5.12).

.

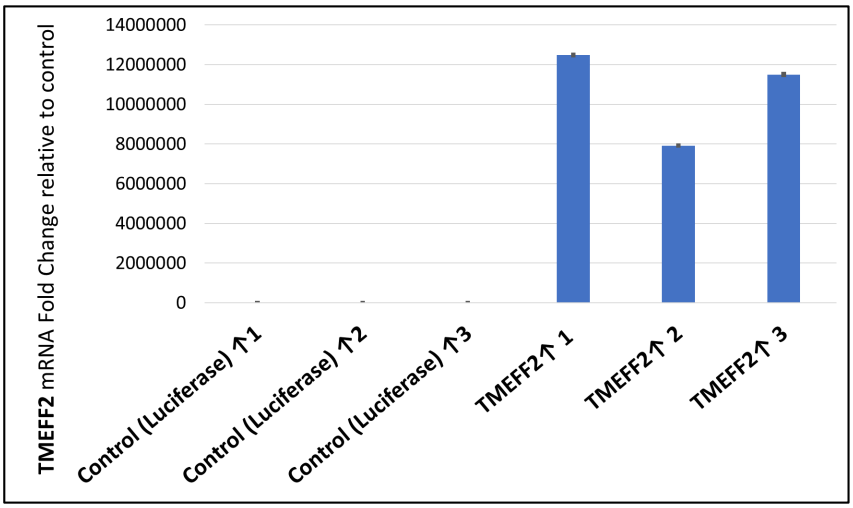


***Figure 5.12*: Confirmation of TMEFF2 overexpression in triplicate samples for microarray transcriptomic analysis**. qRT-PCR results for the relative TMEFF2 mRNA fold change in MCF-7 cells transfected with luciferase 'control' and TMEFF2 in triplicate. Fold change was quantified according to ΔΔCt method using L19 as the housekeeping gene. Cells were harvested 26.5 hours post-transfection and RNA for each sample represented was extracted after pooling cells form 12 wells of a 24-well plate, each with independently transfected population of cells. Bars represent the average readings form 3 PCR-reactions. Error bars represent +/- standard deviation.

5.2.6 Differential gene expression analysis of TMEFF2 overexpressing MCF-7 cells

In order to investigate the differential gene expression in MCF-7 cells upon the overexpression of TMEFF2, I employed Affymetrix® GeneChip™, Clariom™ D human microarrays which are capable of providing exon-level expression profiles of > 540,000 transcripts. These include coding mRNA, long non-coding RNA, microRNA, novel transcripts and also the alternative splice variants. The subsequent bioinformatic analysis and the pathway analysis were performed using the Transcriptome Analysis Console (TAC) 4.0.1 (Applied Biosystems) and Ingenuity Pathway Analysis (IPA) (Qiagen®) [86, 207]. The principle component analysis and quality control analysis plots are presented in supplementary figures S.5 and S.6. The results were filtered through a stringent criteria of fold change between 2.0 and -2.0, p value of 0.005 and false discovery rate (FDR) of 0.05. The results are summarized in figure 5.13 and table 5.4. Briefly, upon TMEFF2 overexpression, 3826 transcripts were upregulated while 1787 transcripts were downregulated. Among the differentially expressed transcripts, 768 were for

coding genes, 2004 non-coding transcripts, 256 were pseudogenes, 216 precursor microRNAs, 66 small_RNAs, 216 precursor_microRNAs, 435 novel unassigned transcripts and 1853 multiple complexes. The latter represent exome transcripts which had been reported by known databases to contain more than one type of locus, i.e., the microarray contains probe sets for multiple key individual exons for genes in order to detect splice variation between samples. However, when I looked closely at key differentially expressed genes which fell under the category of Multiple_complexes, by and large, this set of probes simply represented coding genes and, as there were multiple probes for each gene, the sensitivity of the assay increased many fold. Any non-coding transcripts in the category of Multiple_complexes were explicitly mentioned in the gene symbol and Entrez gene IDs. Moreover, these were automatically filtered in the subsequent pathways analysis because the IPA mapped them accordingly, i.e., IPA was provided with probe set IDs and it automatically mapped the multiple complexes in their relative categories.

Table 5.4: **Summary of differentially expressed transcripts in MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'**.

Group	Total	Passed Filter	Up-Regulated	Down-Regulated
NonCoding	66845	2004	1346	658
Multiple_Complex	29510	1853	1238	615
Coding	18858	768	536	232
Unassigned	10001	435	287	148
Pseudogene	4340	256	203	53
Precursor_microRNA	3297	216	147	69
Small_RNA	2386	66	56	10
Ribosomal	507	15	13	2
tRNA	6	0	0	0

Filter: -2 ≤ fold change ≥ 2, One-Way ANOVA with Empirical Bayes Statistics adjustment p value ≤ 0.005 and False Discovery Rate (FDR) calculated through Benjimini-Hochberg Step-Up FDR-controlling procedure ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome arrays. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

a.

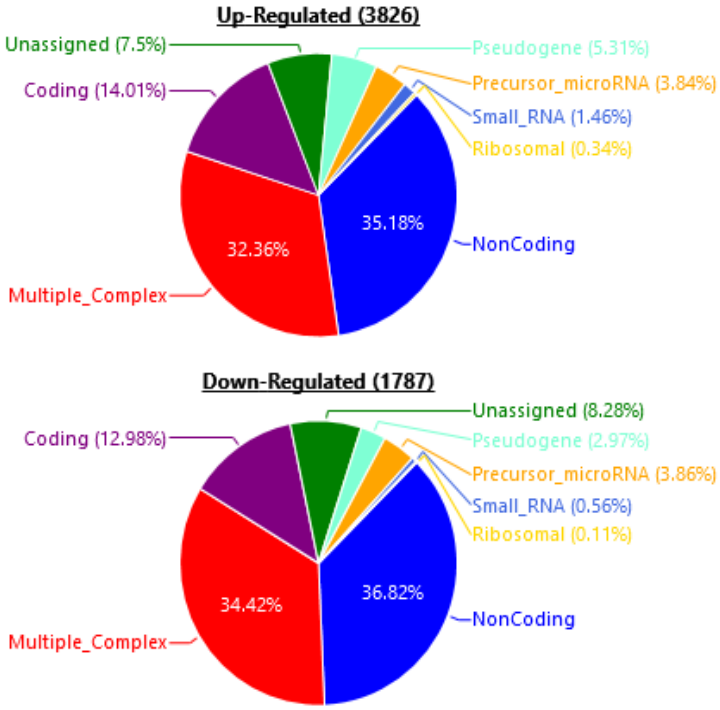


Figure 5.13 **(a): Summary of differentially expressed transcripts in MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'.** Filter: -2 ≤ fold change ≥ 2, p value ≤ 0.005 and False Discovery Rate (FDR) ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

b.

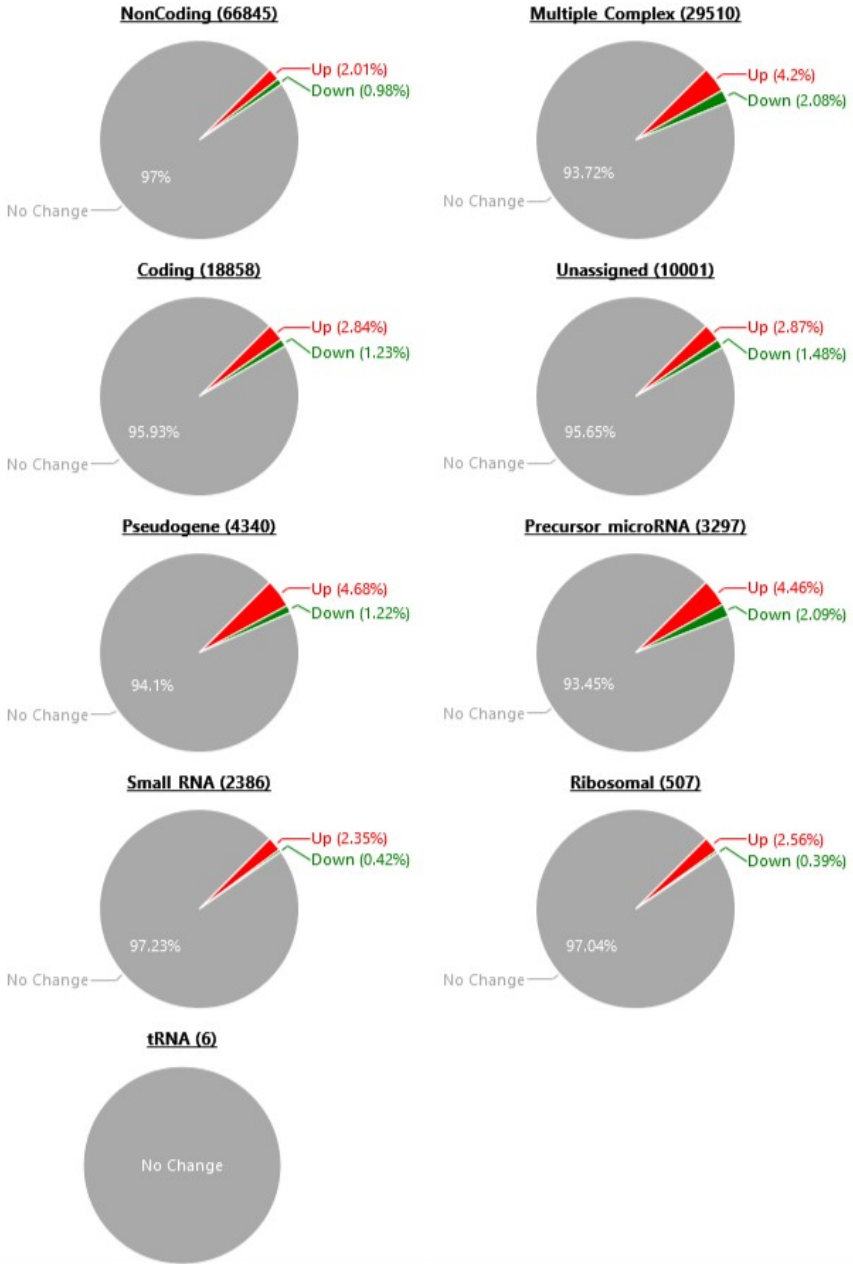


Figure 5.13 **(b): Estimation of differentially expressed transcripts as part of the analysis available transcriptome.** Differential gene analysis of MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control' which passed the Filter, -2 ≤ fold change ≥ 2, p value ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

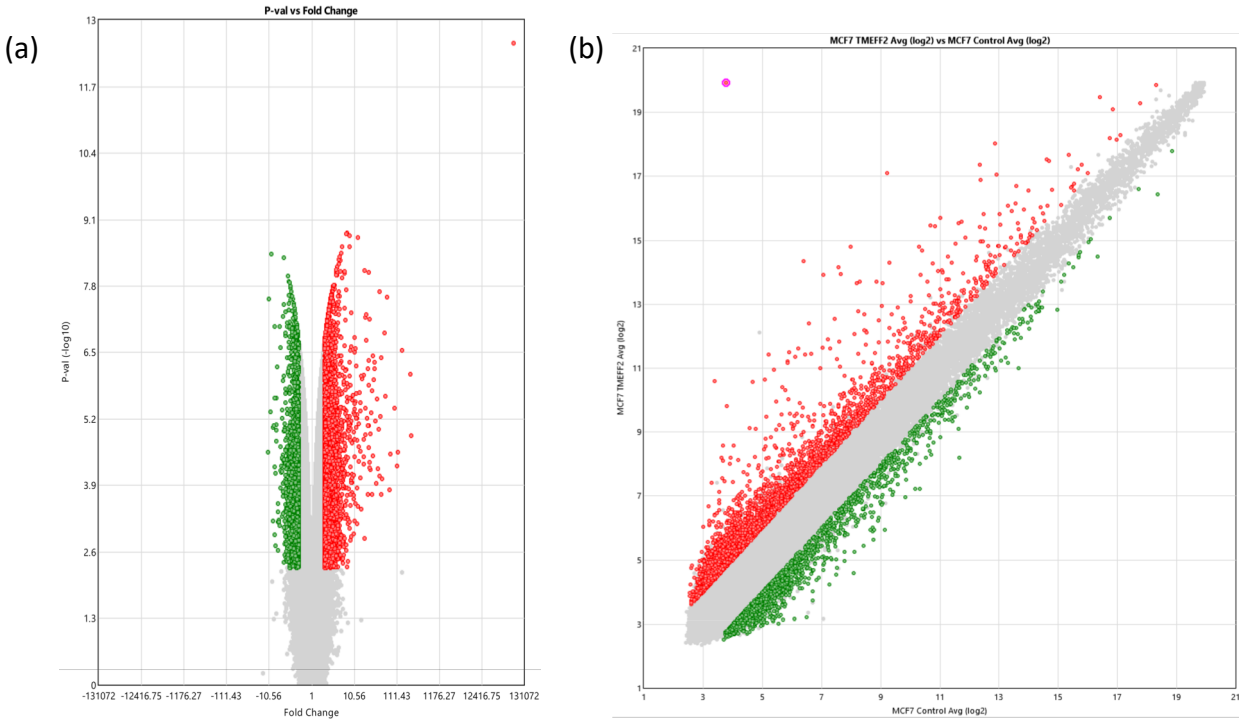


***Figure 5.14*: Volcano plot and scatter plot for differential gene expression upon TMEFF2 overexpression in MCF-7 cells:** Differential gene expression between MCF-7 cells transfected with TMEFF2 and MCF-7 cells transfected with Luciferase 'control'**.** Each dot presents one transcript: red = higher fold change, green = lower fold change, grey = did not pass the Filter: -2 ≤ fold change ≥ 2, p value ≤ 0.005, FDR ≤ 0.05. (a) p-value *vs* fold change (Volcano plot) x-axis: Fold change, left negative, right positive, y-axis: p value ($-\log_{10}$). (b) Scatter plot x-axis: MCF-7 control average signal intensity ($\log_2$), y-axis: MCF-7 TMEFF3 average signal intensity ($\log_2$). Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

When examined in the context of the whole transcriptome (available on Clariom™ D microarray), the change imparted by TMEFF2 overexpression was relatively small (see figure 5.13 (b)). Another important observation from the analysis was that the upregulated genes were prominently higher in both number and fold change than the downregulated ones (figure 5.14). TMEFF2 overexpression mainly resulted in upregulation of genes. For example, the two most upregulated genes, CCL22 and CXCL10, showed a fold change of 248.36 and 235.49 respectively. In comparison, the most downregulated genes, RP11-498C9.15 and DANCR, showed a fold change decrease of -11.3 and -10.94 respectively. Moreover, the chromosomal analysis showed that the changes in transcriptome are by and large evenly distributed across the chromosomal loci, but some concentration was observed on chromosomes 17 and 19 (figure 5.15).



***Figure 5.15*: Loci of differentially expressed transcripts in TMEFF2 overexpressing MCF-7 cells.** MCF-7 cells transiently transfected with TMEFF2 *vs* MCF-7 cells transiently transfected with Luciferase 'control'. Filter: -2 ≤ fold change ≥ 2, p value ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

Table 5.5 presents the list of 60 most upregulated and 30 most downregulated protein coding genes according to fold change which passed the filter (p value ≤ 0.005 and false discovery rate (FDR) ≤ 0.05). A clustered heatmap of a small subset of genes is presented in figure 5.16. The list also includes genes which fall under the category of Multiple_complexes. I shall discuss the functional roles of differentially expressed genes in the next section. But *prima facie*, a bird's-eye view at the list of genes highlighted three groups. Firstly, TMEFF2 overexpression resulted in a marked upregulation in interferons and their related and/or regulated genes. These include interferon beta 1 (IFNB1), interferon lambda 1 (IFNL1), interferon lambda 2 (FINL2), interferon lambda 3 (IFNL3), interferon stimulated exonuclease gene 20 kDa (ISG20), interferon alpha inducible protein 27 (IFI27), interferon induced protein with tetratricopeptide repeats 2 *(*IFIT2), interferon induced protein 44 (IFI44), interferon alpha inducible protein 6 (IFI6), interferon gamma inducible protein 16 (IFI16), interferon induced protein 44-like (IFI44L) and interferon inducible guanylate binding protein 1 (GBP1). The second noteworthy group of upregulated genes was chemokines. These include chemokine (C-C motif) ligands 2, 5 and 22, CCL2, CCL5 and CCL22 respectively, chemokine (C-C motif) ligands 8, 10, 11 and 16, CXCL8, CXCL10, CXCL11, CXCL16 respectively. Thirdly, two genes were noticeable in particular: (i) TNFSF10, otherwise known as TRAIL, showed a 25-fold increase upon TMEFF2 overexpression and (ii) tumour necrosis factor alpha-induced protein 3 (TNFAIP3 – more commonly known as A20 which is a potent inhibitor of NF-κB) increased 37-fold upon TMEFF2 overexpression.

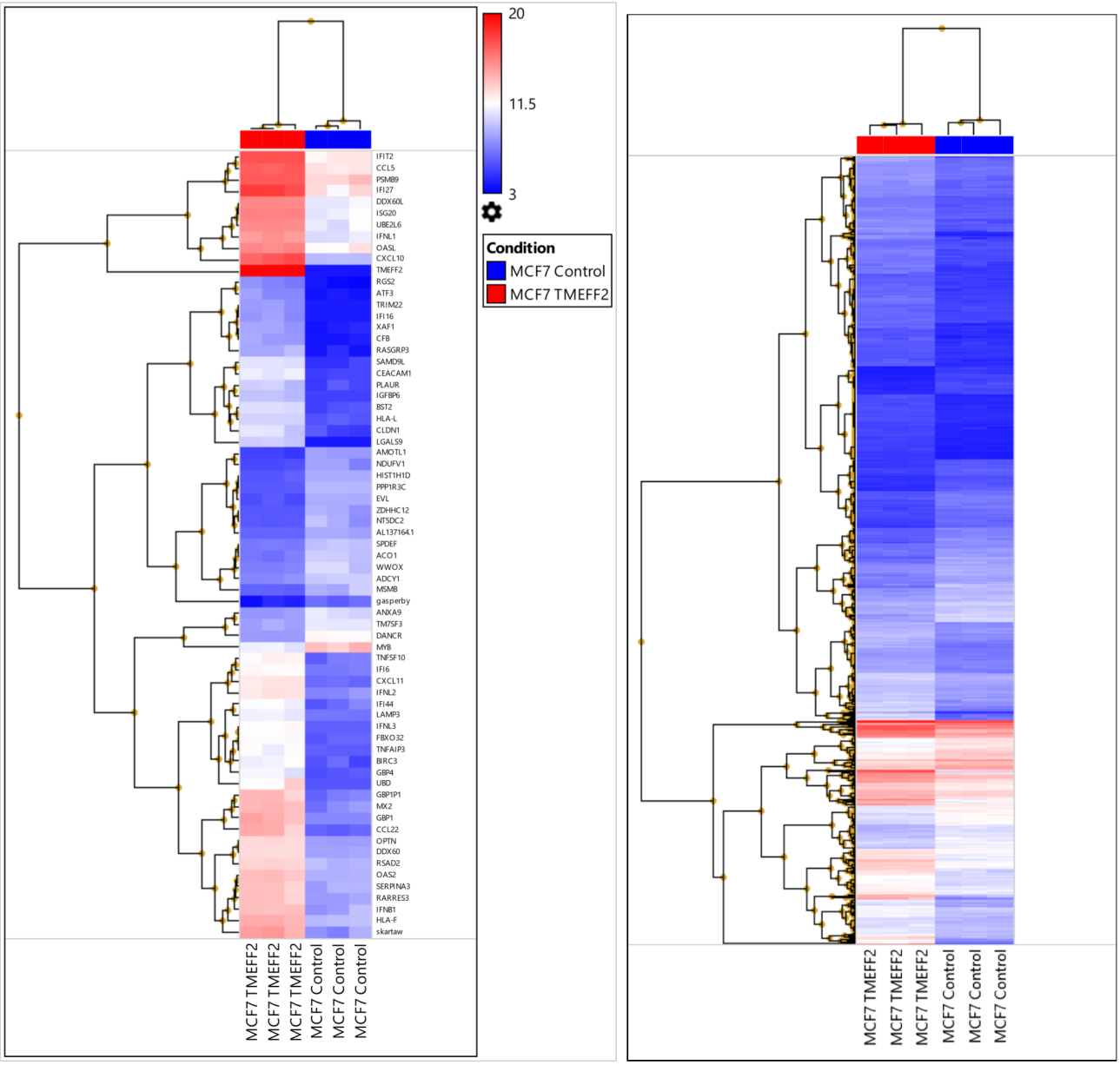


***Figure 5.16*: Heatmaps of a subset of differentially expressed genes upon TMEFF2 overexpression in MCF-7 cells:** Heatmaps of a subset of differentially expressed genes. between MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with Luciferase 'Control'**.** Right: Filter applied: -2 ≤ fold change ≥ 2, p ≤ 0.005 and FDR ≤ 0.05. Only 'Coding' and 'Multiple-Complexes' included in the heatmap. Left: representative most differentially expressed genes from the right. Filter applied: -5 ≤ fold change ≥ 15, p ≤ 0.005, FDR ≤ 0.05. Only 'Coding' and 'Multiple-Complexes' included in the heatmap. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

***Table 5.5*: List of ~60 most upregulated and ~30 most downregulated genes in MCF-7 cells upon TMEFF2 overexpression**

Gene Symbol	Description	Fold Change	P-value[a]	FDR[b] P-value
TMEFF2	transmembrane protein with EGF-like and two follistatin-like domains 2	72655.30	2.86E-13	3.88E-08
CCL22	chemokine (C-C motif) ligand 22	248.36	1.30E-05	4.00E-04
CXCL10	chemokine (C-X-C motif) ligand 10	235.49	8.33E-07	1.00E-04
GBP1	guanylate binding protein 1, interferon-inducible	96.61	3.84E-06	2.00E-04
MX2	MX dynamin-like GTPase 2	77.61	3.05E-05	7.00E-04
LGALS9	lectin, galactosidase-binding, soluble, 9	63.87	2.63E-08	5.80E-05
SAMD9L	sterile alpha motif domain containing 9-like	60.35	2.26E-06	2.00E-04
CXCL11	chemokine (C-X-C motif) ligand 11	56.67	1.28E-07	6.63E-05
UBD	ubiquitin D	48.43	9.30E-05	1.40E-03
CEACAM1	carcinoembryonic antigen-related cell adhesion molecule	47.29	3.97E-07	9.23E-05
CLDN1	claudin 1	46.97	2.00E-04	2.20E-03
RARRES3	retinoic acid receptor responder (tazarotene induced) 3	44.97	5.64E-06	3.00E-04
IFNB1	interferon, beta 1, fibroblast	42.50	8.12E-05	1.30E-03
IFNL3	interferon, lambda 3	41.93	2.02E-08	5.77E-05
GBP4	guanylate binding protein 4	41.26	9.80E-06	4.00E-04
BIRC3	baculoviral IAP repeat containing 3	39.56	5.76E-05	1.00E-03
TNFAIP3	tumour necrosis factor, alpha-induced protein 3	37.99	2.55E-06	2.00E-04
IFI27	interferon, alpha-inducible protein 27	36.49	1.00E-04	1.60E-03
IFNL2	interferon, lambda 2	34.27	6.87E-06	3.00E-04
IFIT2	interferon-induced protein with tetratricopeptide repeats 2	32.79	6.47E-07	1.00E-04
HLA-F	major histocompatibility complex, class I, F	32.09	4.82E-07	9.93E-05
FBXO32	F-box protein 32	31.74	1.06E-06	1.00E-04
RASGRP3	RAS guanyl releasing protein 3 (calcium and DAG-regulated)	30.12	3.26E-05	7.00E-04
OAS2	2-5-oligoadenylate synthetase 2	29.26	3.66E-07	9.05E-05
IFI44	interferon-induced protein 44	27.93	2.00E-04	2.20E-03
DDX60L	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60-like	27.71	1.78E-06	2.00E-04
BST2	bone marrow stromal cell antigen 2	26.48	4.71E-06	3.00E-04
ISG20	interferon stimulated exonuclease gene 20kDa	25.85	1.22E-05	4.00E-04
TNFSF10	tumour necrosis factor (ligand) superfamily, member 10	25.49	6.09E-05	1.10E-03
UBE2L6	ubiquitin-conjugating enzyme E2L 6	24.61	4.38E-05	9.00E-04
IFI6	interferon, alpha-inducible protein 6	24.30	7.11E-08	6.11E-05
SERPINA3	serpin peptidase inhibitor, clade A	24.11	5.61E-05	1.00E-03
XAF1	XIAP associated factor 1	23.21	2.27E-05	6.00E-04
OPTN	optineurin	23.12	8.43E-09	4.93E-05
IFNL1	interferon, lambda 1	22.92	8.45E-06	3.00E-04
CCL5	chemokine (C-C motif) ligand 5	22.89	6.51E-07	1.00E-04
PLAUR	plasminogen activator, urokinase receptor	22.42	2.00E-04	2.20E-03
CFB	complement factor B	21.53	5.97E-06	3.00E-04
IGFBP6	insulin like growth factor binding protein 6	20.44	1.80E-07	7.16E-05
TRIM22	tripartite motif containing 22	19.47	3.81E-06	2.00E-04
DDX60	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60	19.02	8.02E-09	4.93E-05
RGS2	regulator of G-protein signalling 2	17.55	3.37E-05	7.00E-04
PSMB9	proteasome subunit beta 9	17.52	4.73E-05	9.00E-04
IFI16	interferon, gamma-inducible protein 16	17.46	1.55E-05	5.00E-04
LAMP3	lysosomal-associated membrane protein 3	16.45	1.03E-06	1.00E-04
RSAD2	radical S-adenosyl methionine domain containing 2	15.09	4.29E-06	2.00E-04
OASL	2-5-oligoadenylate synthetase-like	15.08	8.77E-05	1.30E-03
ATF3	activating transcription factor 3	15.01	2.73E-05	7.00E-04
SLC15A3	solute carrier family 15 (oligopeptide transporter), member 3	14.83	3.00E-06	2.00E-04
GMPR	guanosine monophosphate reductase	13.56	7.79E-06	3.00E-04
SOD2	superoxide dismutase 2, mitochondrial	13.34	4.30E-05	8.00E-04
RUNX2	runt-related transcription factor 2	12.81	3.00E-04	3.10E-03
GBP3	guanylate binding protein 3	12.54	5.20E-06	3.00E-04
BATF2	basic leucine zipper transcription factor, ATF-like 2	12.52	1.66E-05	5.00E-04
APOL1	apolipoprotein L1	12.24	1.00E-04	1.80E-03
CMPK2	cytidine monophosphate (UMP-CMP) kinase 2, mitochondrial	12.11	5.94E-06	3.00E-04
IFI44L	interferon-induced protein 44-like	12.03	1.98E-06	2.00E-04
TAP1	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	11.97	5.93E-06	3.00E-04
PSMB10	proteasome subunit beta 10	11.90	5.65E-07	1.00E-04
HIST1H2AJ	histone cluster 1, H2aj	-4.34	2.02E-06	2.00E-04
C9orf116	chromosome 9 open reading frame 116	-4.34	1.60E-03	9.30E-03
RAB3A	RAB3A, member RAS oncogene family	-4.34	2.58E-05	6.00E-04
MPZL3	myelin protein zero-like 3	-4.37	2.97E-06	2.00E-04

MCEE	Methylmalonyl-CoA epimerase	-4.46	4.00E-04	3.60E-03
FAM72D	family with sequence similarity 72, member D	-4.62	1.08E-05	4.00E-04
ARMC9	armadillo repeat containing 9	-4.66	1.00E-04	1.50E-03
POTEI	POTE ankyrin domain family, member I	-4.67	4.18E-05	8.00E-04
FAM3B	family with sequence similarity 3, member B	-4.74	3.91E-05	8.00E-04
ARHGAP11A	Rho GTPase activating protein 11A	-4.76	3.66E-08	5.81E-05
TROAP	trophinin associated protein	-4.83	1.11E-06	1.00E-04
BIRC5	baculoviral IAP repeat containing 5	-5.00	4.00E-04	3.90E-03
ANXA9	annexin A9	-5.02	4.00E-04	3.80E-03
SPDEF	SAM pointed domain containing ETS transcription factor	-5.10	4.21E-06	2.00E-04
ADCY1	adenylate cyclase 1 (brain)	-5.14	3.66E-05	8.00E-04
TM7SF3	transmembrane 7 superfamily member 3	-5.31	5.13E-05	9.00E-04
MSMB	Microseminoprotein-beta	-6.06	6.00E-04	4.90E-03
MYB	v-myb avian myeloblastosis viral oncogene homolog	-6.31	6.00E-04	4.60E-03
ZDHHC12	zinc finger, DHHC-type containing 12	-6.56	4.00E-04	3.60E-03
NT5DC2	5-nucleotidase domain containing 2	-6.56	2.30E-03	1.19E-02
HIST1H1D	histone cluster 1, H1d	-7.10	2.39E-06	2.00E-04
AMOTL1	angiomotin like 1	-7.16	7.93E-06	3.00E-04
ACO1	aconitase 1, soluble	-7.28	7.31E-05	1.20E-03
NDUFV1	NADH dehydrogenase (ubiquinone) flavoprotein 1, 51kDa	-7.78	1.20E-03	7.40E-03
WWOX	WW domain containing oxidoreductase	-8.49	6.00E-04	4.80E-03
EVL	Enah/Vasp-like	-8.68	9.17E-06	4.00E-04
PPP1R3C	protein phosphatase 1, regulatory subunit 3C	-9.17	3.71E-09	4.93E-05

[a]p values were calculated using one-way ANOVA with Empirical Bayes Statistics adjustment and [b]False Discovery Rate (FDR) calculated through Benjimini-Hochberg Step-Up FDR-controlling procedure. Only 'Coding' and 'Multiple Complex' transcripts included. Excluding novel transcripts, pseudogenes and RNAs from Multiple Complex. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

As it is evident from figure 5.13 and table 5.4, non-coding transcripts comprised a considerable proportion of differentially expressed transcripts. Moreover, a large number of precursor_microRNAs and smallRNAs also passed the filter. The differential gene expression analysis also identified numerous novel and unassigned transcripts. The proportion of these non-coding transcripts increased substantially towards the downregulated end of the DGE spectrum (in general). In fact, the most downregulated transcript upon TMEFF2 overexpression was a novel long non-coding RNA, RP11-498C9.15, and the second most downregulated transcript was the differentiation antagonizing non-protein coding RNA (DANCR). Also, among the top 10 most downregulated transcripts were a novel uncharacterized non-coding transcript NONHSAT056579, a long non-coding RNA RP11-247L20.4, and the precursor-microRNA MIR31981-1. The function of the latter is largely unknown. Due to time constraints, I was not able to examine this very interesting portion of data in detail.

5.2.7 Pathway analysis

Although transcriptomics analyses reveal a great wealth of information, the methodology is limited to the detection of changes in the RNA levels. When it comes to the upstream factors that are

causative of these changes or the downstream results imparted by changes, specialist bioinformatic platforms are available. I employed a three-tier pathway analysis approach (figure 5.17). Firstly, the Transcriptome Analysis Console (TAC 4.0.1) that was used to analyse the microarray data has integrated WikiPathways [121], a pathway analysis tool. The list of genes from the differential gene expression analysis which passed the filter (Fold change ≥ 2 to -2 ≤, p value ≤ 0.005 and FDR ≤ 0.05) was thus automatically analysed through WikiPathways. However, the WikiPathways has limited functionality and the algorithm matches the molecules in the analysis with the database of only canonical pathways, that is, it overlaps the molecules available in the data with the most well-known typical signalling pathways. WikiPathways is unable to analyse the induction or inhibition of upstream changes that led to the observed changes in transcriptome, neither can it predict the downstream effects of the given transcriptomic changes. For these purposes, I employed a dedicated bioinformatic tool, the Ingenuity Pathway Analysis (IPA) [86, 207]. In addition to matching the molecules in the DGE analysis with the canonical networks, IPA employs a 'casual analysis approach' where the causative factors for the given transcriptomic changes are compared with those previously reported in literature. A similar analysis for the possible downstream effects is also provided by the IPA. The details of the database and calculation methodologies of IPA are elaborated in Ref [86]. Following the recommendation of the software provider, first a list of differentially expressed genes with a slightly loose filter (fold change ≥ 1.5 to -1.5≤, $p < 0.05$ and FDR < 0.25) was provided to the IPA. Moreover, the transcripts type provided to IPA excluded 'non-coding' transcripts. 'Multiple_complex', 'Precusor_MicroRNA',' Small_RNA', 'Pseudogenes', 'Ribosomal', 'tRNA' and 'Unassigned' transcripts were included in the list provided to IPA along with 'Coding' transcripts. From the list of 13189 transcripts provided to IPA, the programme could successfully map 7954 transcripts while 5235 remained unmapped (IPA could not identify them). In the list of 7954 transcripts, the stringent filter was applied again with fold change ≥ 2 to -2 ≤, p ≤ 0.005 and FDR ≤ 0.05. A total of 2294 transcripts passed the filter with 715 downregulated and 1579 upregulated transcripts.

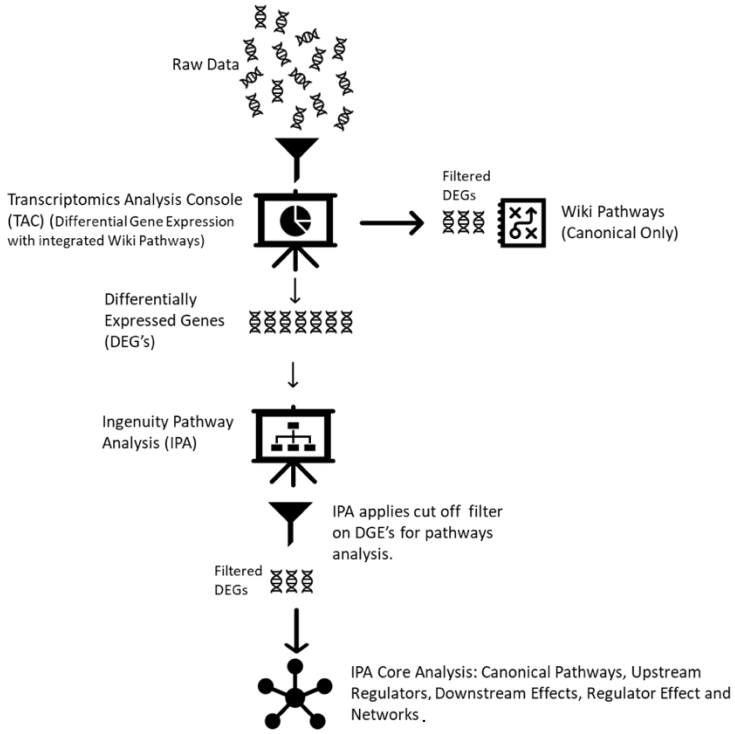
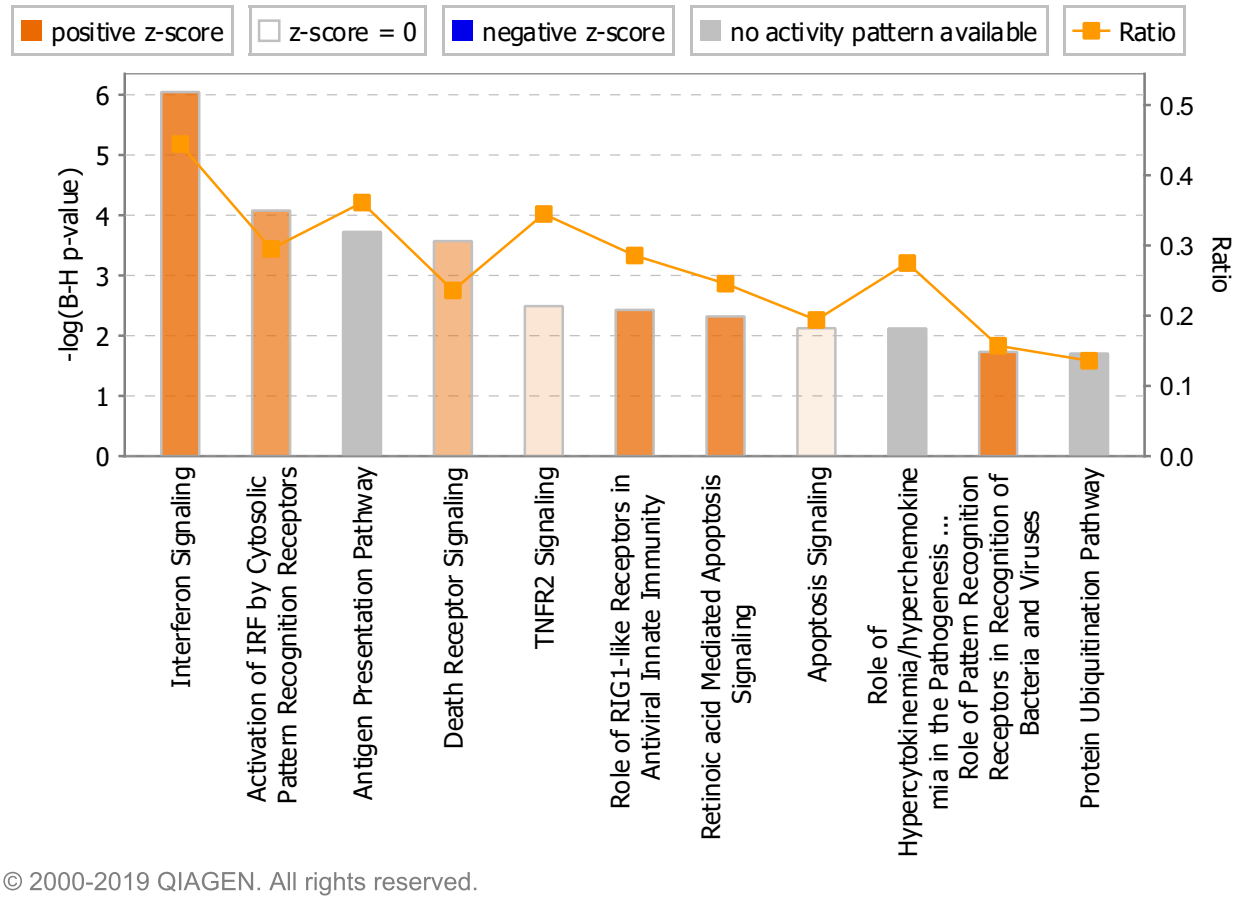


***Figure 5.17*: Schematic representation of the three-tier approach employed for the pathway analysis.** Transcriptome Analysis Console (TAC 4.0) has integrated WikiPathways analysis tool. The list of genes from differential gene expression analysis which passed the filter (fold change ≥ 2 to -2 ≤, p value ≤ 0.005 and FDR ≤ 0.05) was analysed through WikiPathways. In addition, the list of genes which passed the filter with -1.5 ≤ fold change ≥ 1.5, $p < 0.05$ and FDR < 0.25 was provided to the Ingenuity Pathway Analysis (IPA) (Qiagen®) [86] [207]. IPA filtered the list further through the cut-off values applied (-2 ≤ fold change ≥ 2, p value ≤ 0.005 and FDR ≤ 0.05) and ran the IPA Core Analysis which includes the Canonical Pathways analysis, the Upstream Regulator Analysis and the Downstream Effects Analysis. IPA also provides the Regulator Effects analysis and the Networks analysis (these two analyses were not utilized for this study).

5.2.7.1 Canonical Pathways

The top 11 IPA Canonical Pathways (according to B-H p-value of overlap) are shown in figure 5.18. Based on the B-H p-value, the activation z-score and the ratio of overlapping molecules, two cellular processes dominated the analysis results, the first being the interferon signalling and, the second, the death receptor dependent apoptosis signalling. The rest of pathways seemed to be the offshoots of these two processes. The lists of individual overlapping molecules in each pathway, along with their fold change and expected up or downregulation are depicted in tables 5.6 through to 5.13. The upregulated molecules involved in the 'interferon signalling' included interferon inducible proteins (IFI-), IFI6, IFI35, IFITI1, IFITM1, FIFITM2, IFITM3, interferon beta 1 (IFNB1), interferon regulator factor

1 and 9 (IRF1 and IRF9), 2'-5' oligoadenylate synthase 1 (OAS1), ISG15 ubiquitin-like modifier, proteasome subunit beta 8 (PSMB8) and transporter 1, ATP binding cassette subfamily (TAP1). There is only one molecule among the list of 16 which did not match the expected regulatory direction: the suppressor of cytokine signalling (SOC1), which was expected to be downregulated upon activation of interferon signalling but it is slightly upregulated in the gene set.



© 2000-2019 QIAGEN. All rights reserved.

Figure 5.18: **Canonical Pathways outlook upon TMEFF2 overexpression in MCF-7 cells.** Analysis results are based on differentially expressed genes between MCF-7 cells transiently transfected with TMEFF2 *vs* MCF-7 cells transiently transfected with luciferase 'control'. Analysis cut off set at -log(Benjimini-Hochberg p-value) ≥ 1.7. DGE and pathway Analysis filter: -2 ≤ fold change ≥ 2, p ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. DGE analysis was performed using Transcriptome Analysis Console (TAC) 4.0.1. Pathway analysis and image source: Ingenuity Pathway Analysis - IPA (QIAGEN®).

The molecules involved in the 'activation of IRF by cytosolic pattern recognition receptors' included DExD/H-box helicase 58 (DDX58), interferon alpha 4, 7 and 8 (IFNA4, IFNA7 and IFNA8), interferon beta 1 (IFNB1), interferon induced protein with tetratricopeptide repeats 2 (IFIT2), interferon induced with helicase C domain 1 (IFIH1), interferon regulatory factor 9 (IRF9), interleukin 6 (IL6), Jun proto-oncogene, AP-1 transcription factor subunit (JUN), STAT2, TANK biding kinase 1 (TBK1), tumour necrosis factor (TNF) and Z-DNA binding protein 1 (ZBP1). The molecules which did not match the expected direction of regulation are noteworthy. These are the NF-κB inhibitor alpha (NFKBIA) and NF-κB inhibitor epsilon (NFKBIE), which are expected to be downregulated upon activation of this pathway but are rather upregulated upon TMEFF2 overexpression. In other words, upon activation of interferon pathway, NF-κB is normally activated but in the case of TMEFF2 overexpression, the opposite appears to be happening. Another canonical pathway, the 'role of pattern recognition receptors in recognition of bacteria and virus', appears to be an offshoot of the pattern recognition receptors activation and the subsequent interferon induction. Noteworthy molecules in this pathway include nuclear factor kapa B subunit 2 (NFKB2), receptor interacting serine/threonine kinase 2 (RIPK2), tumour necrosis factor (TNF), TNF superfamily member 10 (TNFSF10, otherwise known as TRAIL), and TNF superfamily member 15 (TNFSF15, a potent vascular endothelial growth factor (VEGF) inhibitor). The molecules involved in this pathway ('role of pattern recognition receptors in recognition of bacteria and virus') which did not overlap with other interferon related pathways included complements C3 and C5, caspase-1 (CAPS1), C-type lectin domain containing 7A (CLEC7A), myeloid differentiation primary response 88 (MYD88), 2'-5' oligoadenylate synthase 2 and 3 (OAS2 and OAS3 respectively), ribonuclease L (RNASEL) and interleukin 6 family cytokine (LIF). The 'antigen presentation pathway' molecules associated with TMEFF2 overexpression included beta-2-microglobulin (B2M), major histocompatibility complexes class A, B, C, E and F (HLA-A, HLA-B, HLA-C, HLA-E and HLA-F respectively). Also, in this pathway are NLR family CARD domain containing 5 (NLRC5), protein isomerase family A member 3 (PDIA3), proteasome subunit beta 8 and 9 (PSMB8 and PSMB9), transporter 1, ATP binding cassette subfamily B member (TAP1), TAP2 and TAB binding protein (TAPBP). It was evident from the IPA Canonical Pathways analysis that a strong interferon response was induced upon TMEFF2 overexpression in MCF-7 cells. Moreover, according to both IPA and WikiPathway, both type I and type II interferon were activated. In addition, as IFNL2 and IFNL3 exhibited a 34.3 and 41.9 fold increase in their expression levels respectively, type III interferon response also appeared to be activated upon TMEFF2 overexpression in MCF-7 cells. As I mentioned earlier, the WikiPathways tool is integrated in the Transcriptome Analysis Console 4.0. The interferon

alpha/beta signalling and type II interferon signalling from WikiPahtways analysis are shown in supplementary figures S.7 and S.8 respectively. The IPA software has several tools for analysis and presentation. One of these tools is 'grow', whereby the user can add molecules to a pathway, which are linked upstream or downstream to the chosen molecules. IPA gives the option to add the linked molecules from either generalized literature search or from the list of given genes. For the purposes of depiction only, I used the grow tool in IPA in the Interferon signalling canonical pathway and used the differentially expressed genes in MCF-7 cells as the list of molecules that IPA could use for growing the network further (figure 5.19). Importantly, the connections shown in figure 5.19, downstream of the first-tier interferon induced genes are there only for the purpose of expressing that both type I and type II interferon pathways are activated upon TMEFF2 overexpression. The connections downstream from interferon induced first-tier genes do not necessarily occur as depicted and their expression could be induced through alternative regulators. For example, the upregulation of IFNL1 could be the result of transcriptional activity of IRF1, IRF7 or IRF9, but in figure 5.19, IFNL1 is only shown to be downstream of IRF1.

Table 5.6: **Molecules involved in 'Interferon Signalling'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

Interferon Signalling[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
IFI6	interferon alpha inducible protein 6	↑ 24.3	7.11E-08	6.11E-05	↑ Up	other
IFI35	interferon induced protein 35	↑ 2.45	2.50E-06	2.00E-04	↑ Up	other
IFIT1	interferon induced protein with tetratricopeptide repeats 1	↑ 6.93	3.99E-05	8.00E-04	↑ Up	other
IFIT3	interferon induced protein with tetratricopeptide repeats 3	↑ 7.53	6.00E-04	4.50E-03	↑ Up	other
IFITM1	interferon induced transmembrane protein 1	↑ 8.52	2.00E-04	2.20E-03	↑ Up	transmembrane receptor
IFITM2	interferon induced transmembrane protein 2	↑ 3.22	6.16E-07	1.00E-04	↑ Up	other
IFITM3	interferon induced transmembrane protein 3	↑ 2.94	2.10E-05	6.00E-04	↑ Up	other
IFNB1	interferon beta 1	↑ 42.5	8.12E-05	1.30E-03	↑ Up	cytokine
IRF1	interferon regulatory factor 1	↑ 10.64	3.98E-06	2.00E-04	↑ Up	transcription regulator
IRF9	interferon regulatory factor 9	↑ 2.72	9.00E-04	6.40E-03		transcription regulator
ISG15	ISG15 ubiquitin-like modifier	↑ 2.71	1.00E-04	1.80E-03	↑ Up	other
OAS1	2'-5'-oligoadenylate synthetase 1	↑ 4.98	3.67E-06	2.00E-04	↑ Up	enzyme
PSMB8	proteasome subunit beta 8	↑ 3.71	3.67E-06	2.00E-04	↑ Up	peptidase
SOCS1	suppressor of cytokine signalling 1	↑ 2.16	8.62E-05	1.30E-03	↓Down	other
STAT2	signal transducer and activator of transcription 2	↑ 5.1	6.80E-05	1.10E-03	↑ Up	transcription regulator
TAP1	transporter 1, ATP binding cassette subfamily B member	↑ 11.97	5.93E-06	3.00E-04	↑ Up	transporter
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from Figure 5.18

Table 5.7: **Molecules involved in 'Activation of IRF by Cytosolic Pattern Recognition Receptors'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

Activation of IRF by Cytosolic Patter Recognition Receptors[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
DDX58	DExDH-box helicase 58	↑ 5.64	4.00E-04	3.50E-03	↑ Up	enzyme
IFIH1	interferon induced with helicase C domain 1	↑ 8.65	7.61E-06	3.00E-04	↑ Up	enzyme
IFIT2	interferon induced protein with tetratricopeptide repeats 2	↑ 32.79	6.47E-07	1.00E-04	↑ Up	other
IFNA4	interferon alpha 4	↑ 2.01	4.78E-05	9.00E-04	↑ Up	cytokine
IFNA7	interferon alpha 7	↑ 2.09	1.11E-06	1.00E-04	↑ Up	cytokine
IFNA8	interferon alpha 8	↑ 2.95	1.80E-03	1.01E-02	↑ Up	cytokine
IFNB1	interferon beta 1	↑ 42.5	8.12E-05	1.30E-03	↑ Up	cytokine
IL6	interleukin 6	↑ 2.76	1.10E-03	7.20E-03	↑ Up	cytokine
IRF9	interferon regulatory factor 9	↑ 2.72	9.00E-04	6.40E-03	↑ Up	transcription regulator
ISG15	ISG15 ubiquitin-like modifier	↑ 2.71	1.00E-04	1.80E-03	↓ Down	other
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	↑ 2.45	1.20E-03	7.70E-03	↑ Up	transcription regulator
NFKB2	nuclear factor kappa B subunit 2	↑ 4.21	2.00E-04	2.10E-03	↑Up	transcription regulator
NFKBIA	NFKB inhibitor alpha	↑ 8.48	1.38E-06	1.00E-04	↓ Down	transcription regulator
NFKBIE	NFKB inhibitor epsilon	↑ 2.99	3.34E-05	7.00E-04	↓ Down	transcription regulator
STAT2	signal transducer and activator of transcription 2	↑ 5.1	6.80E-05	1.10E-03	↑ Up	transcription regulator
TBK1	TANK binding kinase 1	↑ 3.14	1.41E-07	6.71E-05	↑ Up	kinase
TNF	tumour necrosis factor	↑ 7.29	1.04E-05	4.00E-04	↑ Up	cytokine
ZBP1	Z-DNA binding protein 1	↑ 7.97	7.49E-06	3.00E-04	↑ Up	other
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from Figure 5.18

Table 5.8: **Molecules involved in 'Role of Pattern Recognition Receptors in Recognition of Bacteria and Virus'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

Role of Pattern Recognition Receptors in Recognition of Bacteria and Virus[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
C3	complement C3	↑ 5.26	3.60E-07	9.01E-05	↑ Up	peptidase
C5	complement C5	↓ -3.06	4.00E-04	3.90E-03	↑ Up	cytokine
CASP1	caspase 1	↑ 2.07	1.00E-03	6.60E-03	↑ Up	peptidase
CCL5	C-C motif chemokine ligand 5	↑ 22.89	6.51E-07	1.00E-04	↑ Up	cytokine
CLEC7A	C-type lectin domain containing 7A	↑ 2.95	1.09E-06	1.00E-04		transmembrane receptor
CXCL8	C-X-C motif chemokine ligand 8	↑ 7.29	6.95E-06	3.00E-04		cytokine
DDX58	DExDH-box helicase 58	↑ 5.64	4.00E-04	3.50E-03	↑ Up	enzyme
IFIH1	interferon induced with helicase C domain 1	↑ 8.65	7.61E-06	3.00E-04	↑ Up	enzyme
IFNA4	interferon alpha 4	↑ 2.01	4.78E-05	9.00E-04	↑ Up	cytokine
IFNA8	interferon alpha 8	↑ 2.95	1.80E-03	1.01E-02	↑ Up	cytokine
IFNB1	interferon beta 1	↑ 42.5	8.12E-05	1.30E-03	↑ Up	cytokine

IL6	interleukin 6	↑ 2.76	1.10E-03	7.20E-03	↑ Up	cytokine
IRS2	insulin receptor substrate 2	↑ 2.57	1.98E-06	2.00E-04	↑ Up	enzyme
LIF	LIF, interleukin 6 family cytokine	↑ 2.08	2.64E-06	2.00E-04		cytokine
MYD88	myeloid differentiation primary response 88	↑ 2.98	7.65E-05	1.20E-03	↑ Up	other
NFKB2	nuclear factor kappa B subunit 2	↑ 4.21	2.00E-04	2.10E-03	↑ Up	transcription regulator
OAS1	2'-5'-oligoadenylate synthetase 1	↑ 4.98	3.67E-06	2.00E-04		enzyme
OAS2	2'-5'-oligoadenylate synthetase 2	↑ 29.3	3.66E-07	9.05E-05		enzyme
OAS3	2'-5'-oligoadenylate synthetase 3	↑ 4.8	3.95E-06	2.00E-04		enzyme
RIPK2	receptor interacting serine/threonine kinase 2	↑ 2.73	6.00E-04	4.80E-03	↑ Up	kinase
RNASEL	ribonuclease L	↑ 2.83	5.63E-05	1.00E-03		enzyme
TLR3	toll like receptor 3	↑ 8.75	9.74E-05	1.40E-03	↑ Up	transmembrane receptor
TNF	tumour necrosis factor	↑ 7.29	1.04E-05	4.00E-04	↑ Up	cytokine
TNFSF10	TNF superfamily member 10	↑ 25.49	6.09E-05	1.10E-03		cytokine
TNFSF15	TNF superfamily member 15	↑ 4.86	1.54E-05	5.00E-04		cytokine
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from Figure 5.18

Table 5.9: **Molecules involved in 'Role of RIG1-like Receptors in Antiviral Innate Immunity'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

Role of RIG1-like Receptors in Antiviral Innate Immunity[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
CASP10	caspase 10	↑ 3.65	2.01E-05	5.00E-04	↑ Up	peptidase
DDX58	DExDH-box helicase 58	↑ 5.64	4.00E-04	3.50E-03	↑ Up	enzyme
IFIH1	interferon induced with helicase C domain 1	↑ 8.65	7.61E-06	3.00E-04	↑ Up	enzyme
IFNA4	interferon alpha 4	↑ 2.01	4.78E-05	9.00E-04	↑ Up	cytokine
IFNA7	interferon alpha 7	↑ 2.09	1.11E-06	1.00E-04	↑ Up	cytokine
IFNA8	interferon alpha 8	↑ 2.95	1.80E-03	1.01E-02	↑ Up	cytokine
IFNB1	interferon beta 1	↑ 42.5	8.12E-05	1.30E-03	↑ Up	cytokine
NFKB2	nuclear factor kappa B subunit 2	↑ 4.21	2.00E-04	2.10E-03	↑ Up	transcription regulator
NFKBIA	NFKB inhibitor alpha	↑ 8.48	1.38E-06	1.00E-04		transcription regulator
NFKBIE	NFKB inhibitor epsilon	↑ 2.99	3.34E-05	7.00E-04		transcription regulator
TBK1	TANK binding kinase 1	↑ 3.14	1.41E-07	6.71E-05	↑ Up	kinase
TRIM25	tripartite motif containing 25	↑ 2.33	5.00E-04	4.20E-03	↑ Up	transcription regulator
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from figure 5.18

Table 5.10: **Molecules involved in 'Role of Hypercytokinemia / hyperchemokinemia in Pathogenesis of Influenza'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control

Role of Hypercytokinemia / hyperchemokinemia in Pathogenesis of Influenza[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
CCL2	C-C motif chemokine ligand 2	↑ 7.59	7.59E-05	1.20E-03		cytokine

CCL5	C-C motif chemokine ligand 5	↑ 22.89	6.51E-07	1.00E-04		cytokine
CXCL8	C-X-C motif chemokine ligand 8	↑ 7.29	6.95E-06	3.00E-04		cytokine
CXCL10	C-X-C motif chemokine ligand 10	↑ 235.49	8.33E-07	1.00E-04		cytokine
IFNA4	interferon alpha 4	↑ 2.01	4.78E-05	9.00E-04		cytokine
IFNA7	interferon alpha 7	↑ 2.09	1.11E-06	1.00E-04		cytokine
IFNA8	interferon alpha 8	↑ 2.95	1.80E-03	1.01E-02		cytokine
IFNB1	interferon beta 1	↑ 42.5	8.12E-05	1.30E-03		cytokine
IFNL1	interferon lambda 1	↑ 22.92	8.45E-06	3.00E-04		cytokine
IL6	interleukin 6	↑ 2.76	1.10E-03	7.20E-03		cytokine
TNF	tumour necrosis factor	↑ 7.29	1.04E-05	4.00E-04		cytokine
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from figure 5.18

Table 5.11: **Molecules involved in 'Antigen Presentation Pathway'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

Antigen Presentation Pathway[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
B2M	beta-2-microglobulin	↑ 7	2.38E-06	2.00E-04		transmembrane receptor
HLA-A	major histocompatibility complex, class I, A	↑ 2.99	1.29E-05	4.00E-04		other
HLA-B	major histocompatibility complex, class I, B	↑ 5.13	4.75E-05	9.00E-04		transmembrane receptor
HLA-C	major histocompatibility complex, class I, C	↑ 4.73	2.24E-05	6.00E-04		other
HLA-E	major histocompatibility complex, class I, E	↑ 8.03	4.94E-07	1.00E-04		transmembrane receptor
HLA-F	major histocompatibility complex, class I, F	↑ 32.09	4.82E-07	9.93E-05		transmembrane receptor
NLRC5	NLR family CARD domain containing 5	↑ 2.09	1.66E-07	6.98E-05		transcription regulator
PDIA3	protein disulfide isomerase family A member 3	↑ 2.22	3.85E-06	2.00E-04		peptidase
PSMB8	proteasome subunit beta 8	↑ 3.71	3.67E-06	2.00E-04		peptidase
PSMB9	proteasome subunit beta 9	↑ 17.52	4.73E-05	9.00E-04		peptidase
TAP1	transporter 1, ATP binding cassette subfamily B member	↑ 11.97	5.93E-06	3.00E-04		transporter
TAP2	transporter 2, ATP binding cassette subfamily B member	↑ 4.93	3.00E-04	3.30E-03		transporter
TAPBP	TAP binding protein	↑ 3.93	6.13E-05	1.10E-03		transporter
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from figure 5.18

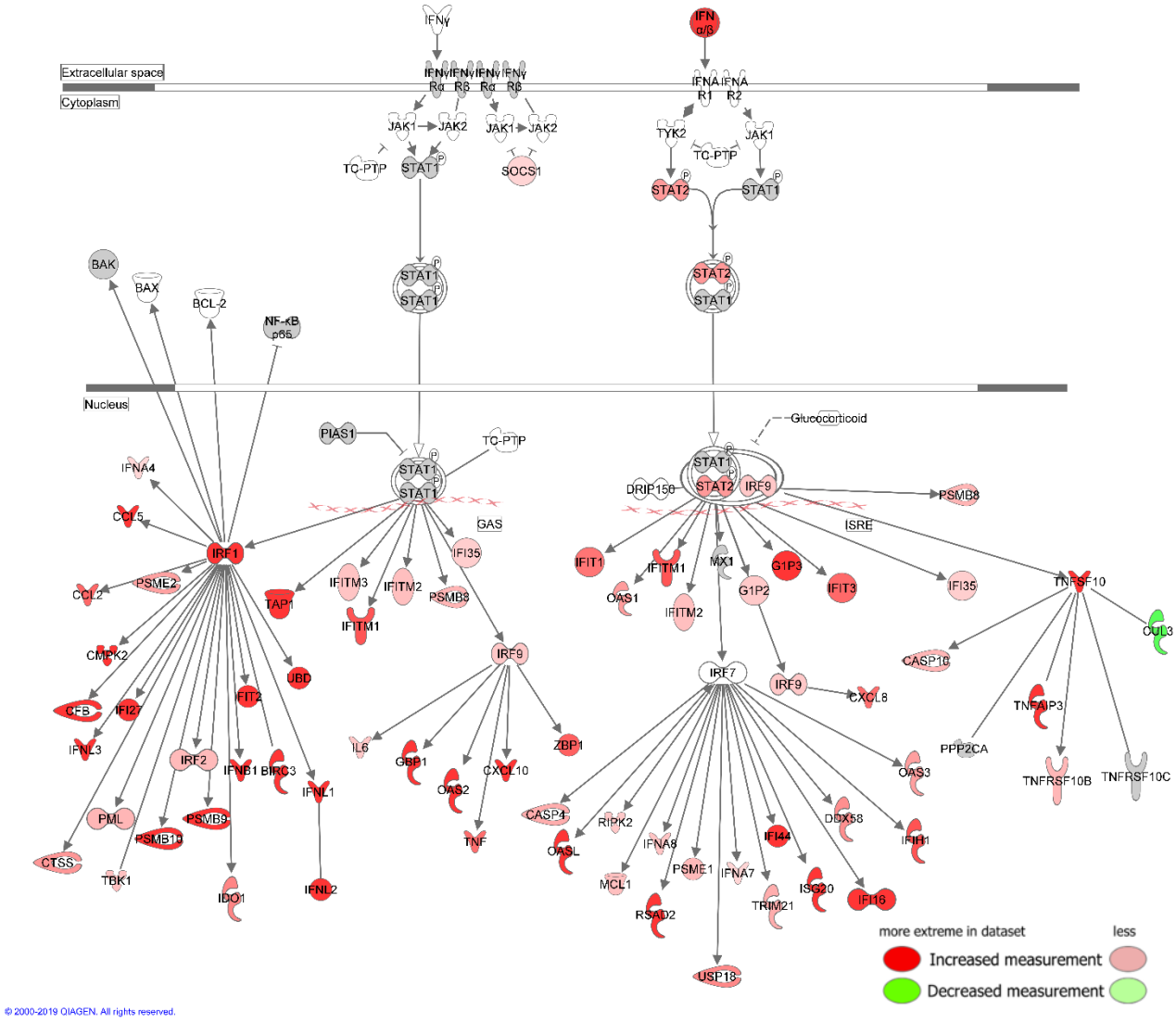


Figure 5.19: **Interferon Signalling canonical pathway outlook upon TMEFF2 overexpression in MCF-7 cells.** IPA interferon signalling only shows first-tier molecules downstream of STAT1-STAT1 dimer and STAT1-STAT2-IRF9 complex. Second-tier molecules were added through the 'grow' tool in IPA using the differentially expressed genes between MCF-7 cells transiently transfected with TMEFF2 *vs* MCF-7 cells transiently transfected with luciferase 'control'. DGE Analysis filter: -2 ≤ fold change ≥ 2, p ≤ 0.005 and FDR ≤ 0.05. NB: The downstream connections from interferon induced first-tier genes do not necessarily occur as depicted and can be activated/inhibited through alternative upstream factors. Expression changes are true depiction of events. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and image source: Ingenuity Pathway Analysis - IPA (QIAGEN®).

When it comes to apoptosis, the analysis results predominantly pointed towards the extrinsic apoptosis upon TMEFF2 overexpression. Figure 5.20 depicts the 'apoptosis' canonical pathway from

the WikiPathways analysis. The apoptosis related canonical pathways from IPA analysis included the 'death receptor signalling', 'TNF2 signalling' and the 'apoptosis pathway'. The related molecules involved in these pathways are depicted in tables 5.12, 5.13 and 5.14. Briefly, molecules involved in the 'death receptor signalling' upon TMEFF2 overexpression can be categorized into (i) tumour necrosis factor family and related molecules including TNF, TNFSF10 or TRAIL and its receptor, TNF receptor family member 10b (TNFRSF10B), (ii) Poly(ADP-ribose) polymerase family members, PARP1, PARP6, PARP8, PARP9, PARP12 and PARP14, (iii) NF-κB related, NFKBIA and NFKBIE and NF-κB subunit 2 (NFKB2), (vi) baculoviral IAP repeat containing, BIRC2 and BIRC3, and lastly (v) caspases 7 and 10 (CASP7, CASP10) and calpain 9 (CAPN9). The molecules related to the 'TNFR2 signalling' mostly overlapped those involved in the death receptor pathway except for AP-1 transcription factor (JUN), TNFAIP3 and intriguingly, the X-linked inhibitor of apoptosis (XIAP). The differentially expressed molecules that were related to the 'apoptosis signalling' and which did NOT overlap with those involved in either 'TNFR2 signalling' or 'death receptor signalling' included APAF1, MCL1 and the proapoptotic BLC2 family member BCL2L11 (also known as BIM). Interestingly, HRAS related molecules were also found to be differentially expressed upon TMEFF2 overexpression. These included HRAS itself, which was downregulated, whereas the RAS related dexamethasone induced 1 (RASD1), RAS related 2 (RASD2) and another RAS family member RAP2B were upregulated upon TMEFF2 overexpression. Moreover, when the expression fold change values of the death receptor ligands are taken into account, TRAIL increased in expression by 25.5-fold whereas the fold upregulation of TNF was three times less (7.3-fold). Furthermore, upon TMEFF2 overexpression, the molecules which are directly related to TRAIL, i.e., BIRC3 and TNFAIP3, also exhibited a substantial upregulation of more than 39- and 37-fold respectively.

Table 5.12: **Molecules involved in 'Apoptosis Signalling'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

Apoptosis Signalling[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Expr Fold	p-value	FDR		Type(s)
APAF1	apoptotic peptidase activating factor 1	↑ 2.31	2.90E-03	1.40E-02	↑ Up	other
BCL2L11	BCL2 like 11	↑ 4.85	4.00E-04	3.50E-03	↑ Up	other
BIRC2	baculoviral IAP repeat containing 2	↑ 3.75	1.90E-03	1.03E-02	↓ Down	enzyme
BIRC3	baculoviral IAP repeat containing 3	↑ 39.56	5.76E-05	1.00E-03	↓ Down	enzyme
CAPN9	calpain 9	↑ 2.77	3.00E-04	3.00E-03	↑ Up	peptidase
CASP7	caspase 7	↑ 3.65	4.77E-07	9.90E-05	↑ Up	peptidase
CASP10	caspase 10	↑ 3.65	2.01E-05	5.00E-04	↑ Up	peptidase
HRAS	HRas proto-oncogene, GTPase	↓ -2.38	5.00E-04	3.90E-03	↓ Down	enzyme
MCL1	MCL1, BCL2 family apoptosis regulator	↑ 3.2	2.46E-05	6.00E-04	↓ Down	transporter
NFKB2	nuclear factor kappa B subunit 2	↑ 4.21	2.00E-04	2.10E-03	↓ Down	transcription regulator
NFKBIA	NFKB inhibitor alpha	↑ 8.48	1.38E-06	1.00E-04	↑Up	transcription regulator
NFKBIE	NFKB inhibitor epsilon	↑ 2.99	3.34E-05	7.00E-04	↑Up	transcription regulator

PARP1	Poly (ADP-ribose) polymerase 1	↓ -2.03	1.50E-03	8.70E-03	↓ Down	enzyme
RAP2B	RAP2B, member of RAS oncogene family	↑ 3.49	2.00E-04	2.40E-03	↓ Down	enzyme
RASD1	ras related dexamethasone induced 1	↑ 3.45	1.62E-08	5.77E-05		enzyme
RPS6KA1	ribosomal protein S6 kinase A1	↑ 2.56	3.64E-05	8.00E-04	↓ Down	kinase
RRAS2	RAS related 2	↑ 2.01	3.98E-05	8.00E-04	↓ Down	enzyme
TNF	tumour necrosis factor	↑ 7.29	1.04E-05	4.00E-04	↑ Up	cytokine
XIAP	X-linked inhibitor of apoptosis	↑ 2	5.00E-04	4.00E-03	Down	enzyme
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from figure 5.18

Table 5.13: **Molecules involved in 'Death Receptor Signalling'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

Death Receptor Signalling[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
ACTG2	actin, gamma 2, smooth muscle, enteric	↑ 2.11	1.68E-07	6.99E-05	↑ Up	other
APAF1	apoptotic peptidase activating factor 1	↑ 2.31	2.90E-03	1.40E-02	↑ Up	other
BIRC2	baculoviral IAP repeat containing 2	↑ 3.75	1.90E-03	1.03E-02	Down	enzyme
BIRC3	baculoviral IAP repeat containing 3	↑ 39.56	5.76E-05	1.00E-03	Down	enzyme
CASP7	caspase 7	↑ 3.65	4.77E-07	9.90E-05	↑ Up	peptidase
CASP10	caspase 10	↑ 3.65	2.01E-05	5.00E-04	↑ Up	peptidase
NFKB2	nuclear factor kappa B subunit 2	↑ 4.21	2.00E-04	2.10E-03	↑ Up	transcription regulator
NFKBIA	NFKB inhibitor alpha	↑ 8.48	1.38E-06	1.00E-04	↓ Down	transcription regulator
NFKBIE	NFKB inhibitor epsilon	↑ 2.99	3.34E-05	7.00E-04	↓ Down	transcription regulator
PARP1	poly(ADP-ribose) polymerase 1	↓ -2.03	1.50E-03	8.70E-03	↑ Up	enzyme
PARP6	poly(ADP-ribose) polymerase family member 6	↑2.93	2.00E-04	2.20E-03	↑ Up	other
PARP8	poly(ADP-ribose) polymerase family member 8	↑ 9.28	9.05E-05	1.40E-03	↑ Up	other
PARP9	poly(ADP-ribose) polymerase family member 9	↑ 5.39	4.00E-04	3.30E-03	↑ Up	enzyme
PARP12	poly(ADP-ribose) polymerase family member 12	↑ 8.42	5.42E-05	1.00E-03	↑ Up	other
PARP14	poly(ADP-ribose) polymerase family member 14	↑ 4.8	1.00E-03	6.60E-03	↑ Up	enzyme
TBK1	TANK binding kinase 1	↑ 3.14	1.41E-07	6.71E-05	↑ Up	kinase
TNF	tumour necrosis factor	↑ 7.29	1.04E-05	4.00E-04	↑ Up	cytokine
TNFRSF10B	TNF receptor superfamily member 10b	↑ 3.62	3.42E-06	2.00E-04	↑ Up	transmembrane receptor
TNFSF10	TNF superfamily member 10	↑ 25.49	6.09E-05	1.10E-03	↑ Up	cytokine
TNFSF15	TNF superfamily member 15	↑ 4.86	1.54E-05	5.00E-04	↑ Up	cytokine
XIAP	X-linked inhibitor of apoptosis	↑ 2	5.00E-04	4.00E-03	↓ Down	enzyme
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from figure 5.18

Table 5.14: **Molecules involved in 'Apoptosis Signalling'** IPA canonical pathway among differentially expressed genes between MCF-7 transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase control.

TNFR2 Signalling[a]						
		Observed			Expected	
Symbol	Entrez Gene Name	Fold Change	p-value	FDR		Type(s)
BIRC2	baculoviral IAP repeat containing 2	↑ 3.75	1.90E-03	1.03E-02	↓ Down	enzyme
BIRC3	baculoviral IAP repeat containing 3	↑ 39.56	5.76E-05	1.00E-03	↓ Down	enzyme
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	↑ 2.45	1.20E-03	7.70E-03	↑ Up	transcription regulator
NFKB2	nuclear factor kappa B subunit 2	↑ 4.21	2.00E-04	2.10E-03	↑ Up	transcription regulator
NFKBIA	NFKB inhibitor alpha	↑ 8.48	1.38E-06	1.00E-04	↑ Up	transcription regulator
NFKBIE	NFKB inhibitor epsilon	↑ 2.99	3.34E-05	7.00E-04	↑ Up	transcription regulator
TBK1	TANK binding kinase 1	↑ 3.14	1.41E-07	6.71E-05	↑ Up	kinase
TNF	tumour necrosis factor	↑ 7.29	1.04E-05	4.00E-04		cytokine
TNFAIP3	TNF alpha induced protein 3	↑ 37.99	2.55E-06	2.00E-04		enzyme
XIAP	X-linked inhibitor of apoptosis	↑ 2	5.00E-04	4.00E-03	↓ Down	enzyme
© 2000-2019 QIAGEN. All rights reserved.						

[a]Data extracted from figure 5.18

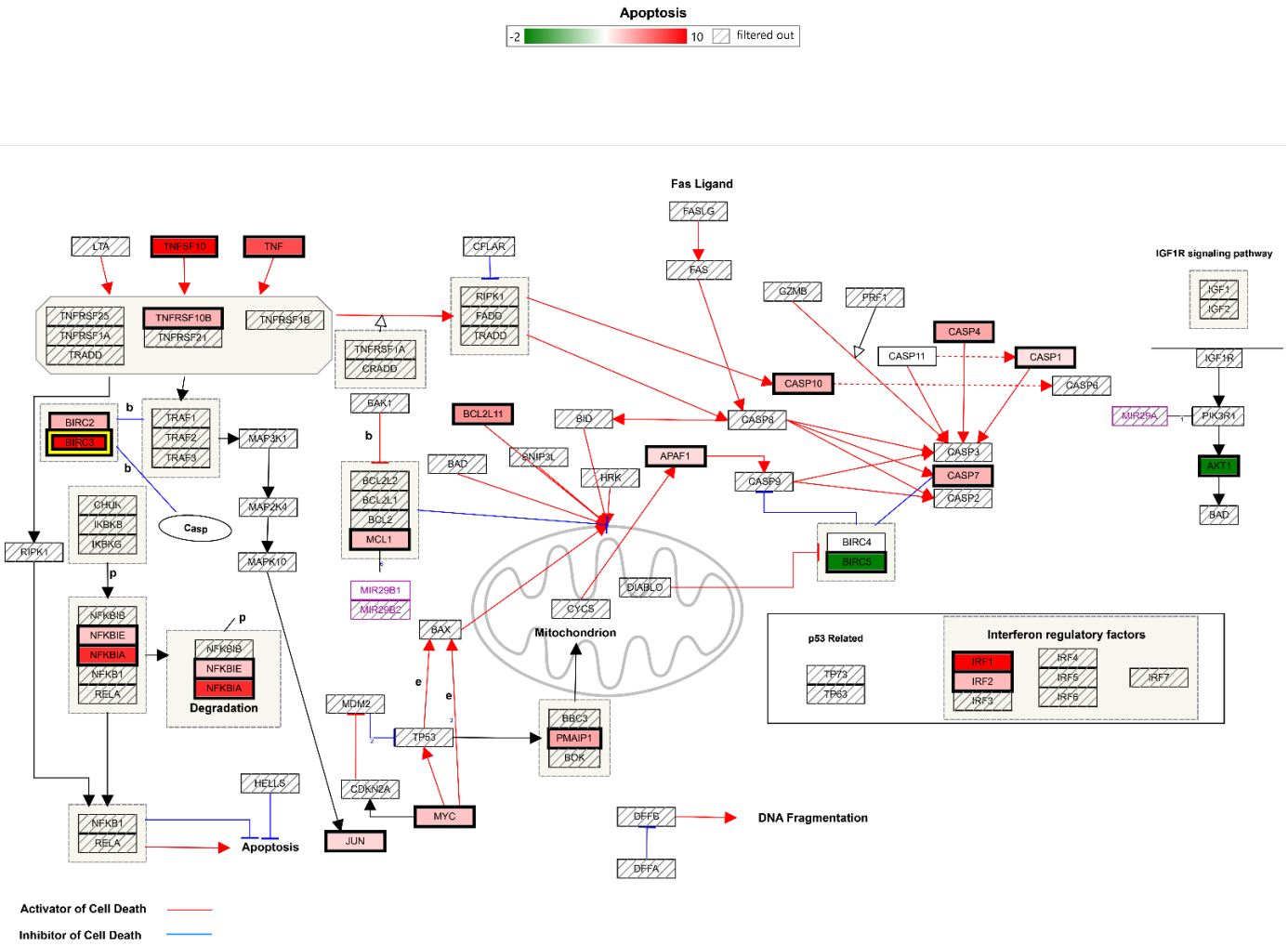


Figure 5.20: **Apoptotic pathway (WikiPathway) upon TMEFF2 overexpression in MCF-7 cells.** Apoptosis canonical pathway (WikiPathway) with overlapping differentially expressed genes between MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with Luciferase 'control'. Filter applied: Fold change ≥ 2 to ≤ -2, p ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1. with integrated WikiPathways (Source).

5.2.7.2 Upstream Regulator Analysis

Along with the Canonical Pathways analysis IPA also performed the Upstream Regulator Analysis. As mentioned earlier, the inability of transcriptomics to assay any change other than an increase or decrease in the mRNA levels of genes poses a limitation. The IPA Upstream Regulator Analysis utilizes the previously known regulatory effects in order to predict the upstream changes which could have caused the given (observed) changes in the transcriptome. The Upstream Regulator Analysis showed a picture similar to that of the canonical pathways with a few interesting exceptions. The top 40 upstream regulators based on Benjamini-Hochberg (B-H) adjusted p value are shown in table 5.15. Again, the interferon related genes along with the death receptor ligands dominated the list of upstream regulators. Additionally, interferon lambda 1 (IFNL1) came out to be the second most significant upstream regulator in this analysis. This again implied that the TMEFF2 overexpression induced not only type I and type II but also type III interferon response in MCF-7 cells. Other interferon related regulators included the transcription factors IRF7, IRF1, IRF3, IRF5 and IRF9 (in order of significance), toll-like receptors TLR3, TLR9, TLR4 and TLR7 and among the kinases, which were predicted to be activated, included the eukaryotic translation initiation factor 2 alpha kinase 2 (EIF2AK2), the interleukin 1 receptor associated kinase 4 (IRAK 1/4), the TANK binding kinase 1 (TBK1) and inhibitor of nuclear factor kappa B kinase subunit epsilon (IKBKE). The former three are the activators of NF-kB while the latter is its potent inhibitor. Noteworthy, as the gene set provided to IPA was dominated by interferon related genes, there was a considerable overlap among the downstream targets of the predicted upstream regulators. Among the few upstream regulators which were predicted to be inhibited are the mitogen-activated protein kinase 1 (MAPK1), otherwise known as ERK2. Also included in the list of inhibited upstream regulators were the transcription factors NK2 homeobox 3 (NKX2-3), CCR4-NOT transcription complex subunit 7 (CNOT7) and tripartite motif containing 24 (TRIM24). Lastly, both TNF and TRAIL were predicted to be activated. TNF had a higher activation z-score whereas TRAIL exhibited a higher expression fold change. The schematic diagrams of upstream regulation by IRF7, TNF and TRAIL are shown in figures 5.21, 5.22 and 5.23 respectively.

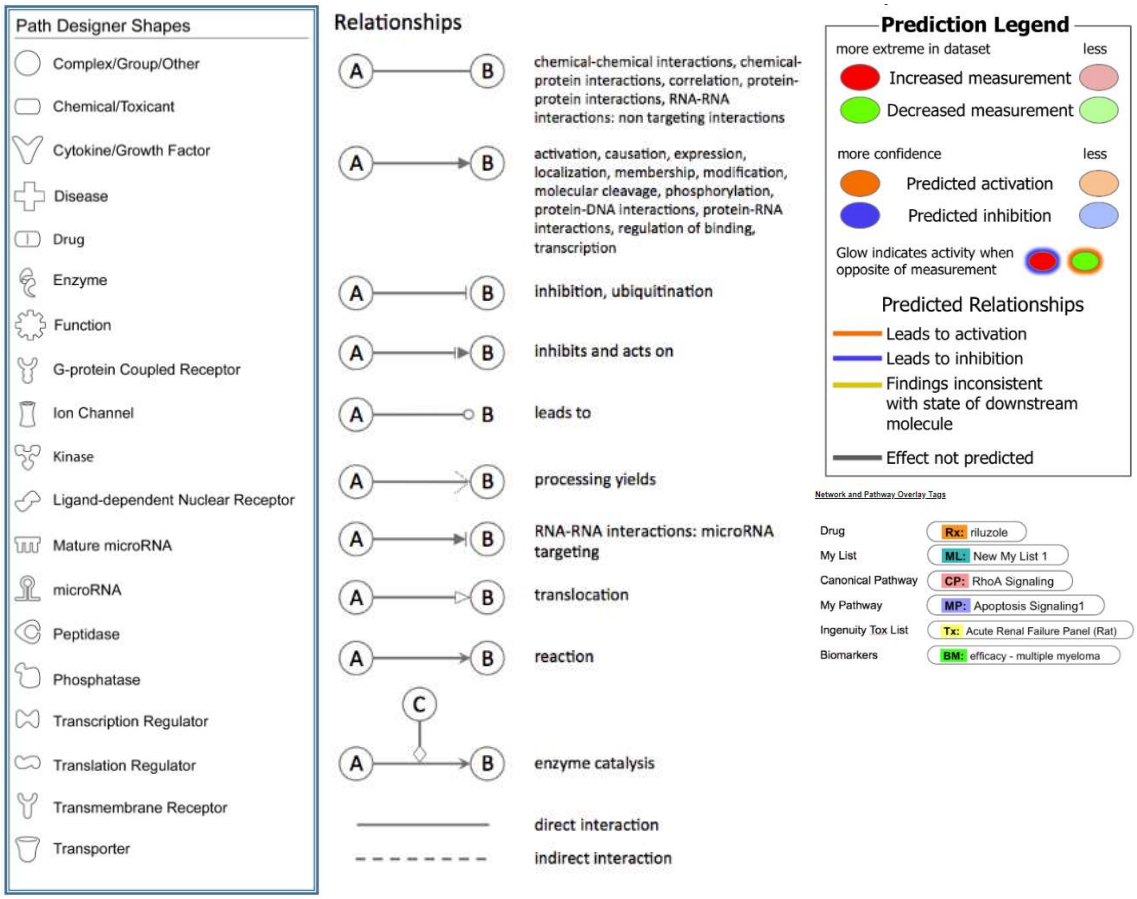
Table 5.15: **IPA Upstream Regulator Analysis for the gene expression changes observed upon TMEFF2 overexpression in MCF-7 cells.** List of the 40 most significant upstream regulators according to B-H corrected p value. Results were generated by running the IPA Upstream Regulator Analysis using the differentially expressed gene set between MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'. DGE Analysis filter; -2 ≤ fold change ≥ 2, p ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and source: Ingenuity Pathway Analysis - IPA (QIAGEN ®).

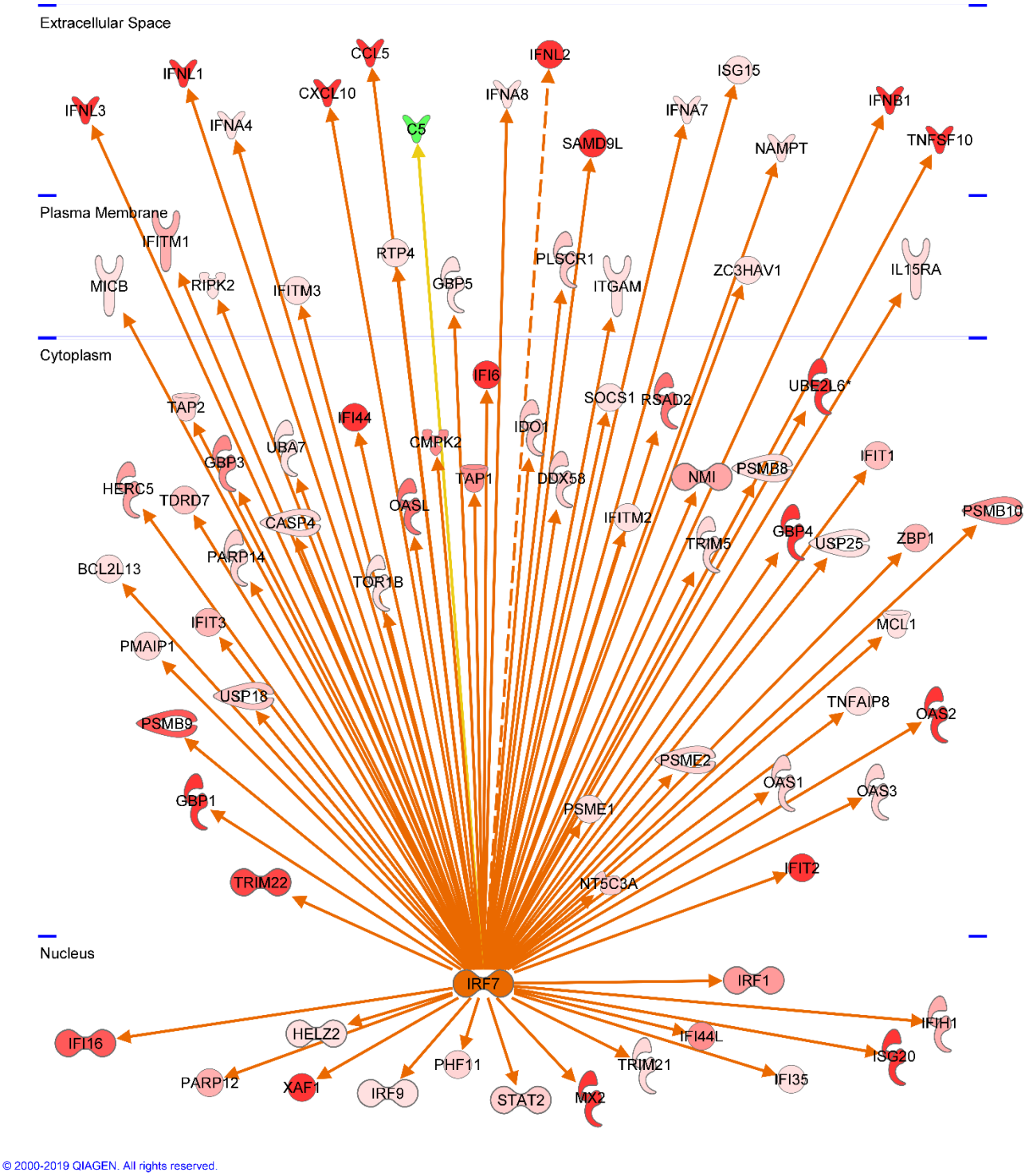
Upstream Regulator	Fold Change	Molecule Type	Predicted Activation State	Activation z-score	p-value of overlap	B-H corrected p-value
IRF7		transcription regulator	**Activated**	**8.582**	1.34×10^{-52}	8.62×10^{-49}
IFNL1	↑ 22.92	cytokine	**Activated**	**7.626**	1.91×10^{-50}	6.08×10^{-47}
IFNA2		cytokine	**Activated**	**8.82**	5.38×10^{-47}	8.63×10^{-44}
STAT1	↑ 3.64	transcription regulator	**Activated**	**7.909**	1.41×10^{-40}	1.51×10^{-37}
IFNG		cytokine	**Activated**	**10.793**	3.42×10^{-40}	3.13×10^{-37}
MAPK1	↓ -1.69	kinase	**Inhibited**	**-5.53**	2.1×10^{-37}	1.68×10^{-34}
IRF1	↑ 10.64	transcription regulator	**Activated**	**6.824**	9.88×10^{-36}	7.05×10^{-33}
NKX2-3		transcription regulator	**Inhibited**	**-6.801**	3.63×10^{-34}	2.33×10^{-31}
IRF3		transcription regulator	**Activated**	**6.904**	1.15×10^{-32}	5.67×10^{-30}
TNF	↑ 7.29	cytokine	**Activated**	**9.983**	2.01×10^{-32}	9.21×10^{-30}
PRL		cytokine	**Activated**	**6.048**	4.05×10^{-27}	1.73×10^{-24}
IFNB1	↑ 42.5	cytokine	**Activated**	**6.035**	5.31×10^{-26}	2.13×10^{-23}
EIF2AK2		kinase	**Activated**	**4.862**	6.27×10^{-26}	2.37×10^{-23}
IFNAR2		transmembrane receptor	**Activated**	**3.207**	2.32×10^{-25}	8.28×10^{-23}
IFNAR1		transmembrane receptor	**Activated**	**4.085**	1.03×10^{-23}	3.14×10^{-21}
IL1RN	↑ 1.97	cytokine	**Inhibited**	**-6.423**	2.33×10^{-23}	6.8×10^{-21}
DDX58	↑ 5.64	enzyme	**Activated**	**4.768**	3.27×10^{-23}	9.12×10^{-21}
TLR3	↑ 8.75	transmembrane receptor	**Activated**	**5.969**	8.92×10^{-23}	2.39×10^{-20}
CNOT7	↓ -1.78	transcription regulator	**Inhibited**	**-2.433**	1.86×10^{-21}	4.59×10^{-19}
TRIM24	↓ -1.57	transcription regulator	**Inhibited**	**-6.125**	9.13×10^{-21}	2.09×10^{-18}
STAT3	↑ 1.93	transcription regulator	**Activated**	**2.231**	2.24×10^{-20}	4.96×10^{-18}
IRF5		transcription regulator	**Activated**	**4.72**	5.26×10^{-20}	1.09×10^{-17}
IL1B		cytokine	**Activated**	**8.241**	1.79×10^{-19}	3.48×10^{-17}
CD40LG		cytokine	**Activated**	**4.166**	2.47×10^{-18}	4.54×10^{-16}
STAT2	↑ 5.1	transcription regulator	**Activated**	**3.76**	2.65×10^{-18}	4.72×10^{-16}
TBK1	↑ 3.14	kinase	**Activated**	**5.18**	4.14×10^{-16}	6.99×10^{-14}
NF-κB (complex)		Complex	**Activated**	**7.569**	8.7×10^{-16}	1.4×10^{-13}
TGM2		enzyme	**Activated**	**6.254**	1.3×10^{-15}	2.08×10^{-13}
TNFSF10	↑ 25.49	cytokine	**Activated**	**3.202**	2.36×10^{-15}	3.52×10^{-13}
OSM		cytokine	**Activated**	**5.898**	2.98×10^{-15}	4.35×10^{-13}
SPI1	↓ -1.83	transcription regulator	**Activated**	**4.236**	8.36×10^{-15}	1.14×10^{-12}
RELA	↑1.73	transcription regulator	**Activated**	**5.216**	9.99×10^{-15}	1.34×10^{-12}
IRAK4		kinase	**Activated**	**2.631**	1.19×10^{-14}	1.55×10^{-12}
NUPR1	↑ 10.08	transcription regulator	**Activated**	**5.33**	1.73×10^{-14}	2.18×10^{-12}
TLR9		transmembrane receptor	**Activated**	**5.211**	2.76×10^{-14}	3.28×10^{-12}

TLR4		transmembrane receptor	**Activated**	**5.438**	1.13×10^{-13}	1.21×10^{-11}
IRF9	↑ 2.72	transcription regulator	**Activated**	**2.201**	3.32×10^{-13}	3.44×10^{-11}
IKBKE	↓ -1.77	kinase	**Activated**	**3.453**	3.89×10^{-13}	3.96×10^{-11}
TLR7		transmembrane receptor	**Activated**	**5.822**	4.56×10^{-13}	4.58×10^{-11}
IFNA1/IFNA13		cytokine	**Activated**	**4.243**	4.9×10^{-13}	4.84×10^{-11}

© 2000-2019 QIAGEN. All rights reserved

Ingenuity Pathways Analysis (IPA) Legend





© 2000-2019 QIAGEN. All rights reserved.

Figure 5.21: (further to table 5.15) **Network view of IRF7 as the predicted upstream regulator for the transcriptomic changes observed upon TMEFF2 overexpression in MCF-7 cells**. IPA Upstream Regulator Analysis was run on differentially expressed gene set between MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'. DGE Analysis filter; Fold change ≥ 2 to -2 ≤, p ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and source: Ingenuity Pathway Analysis - IPA (QIAGEN ®).

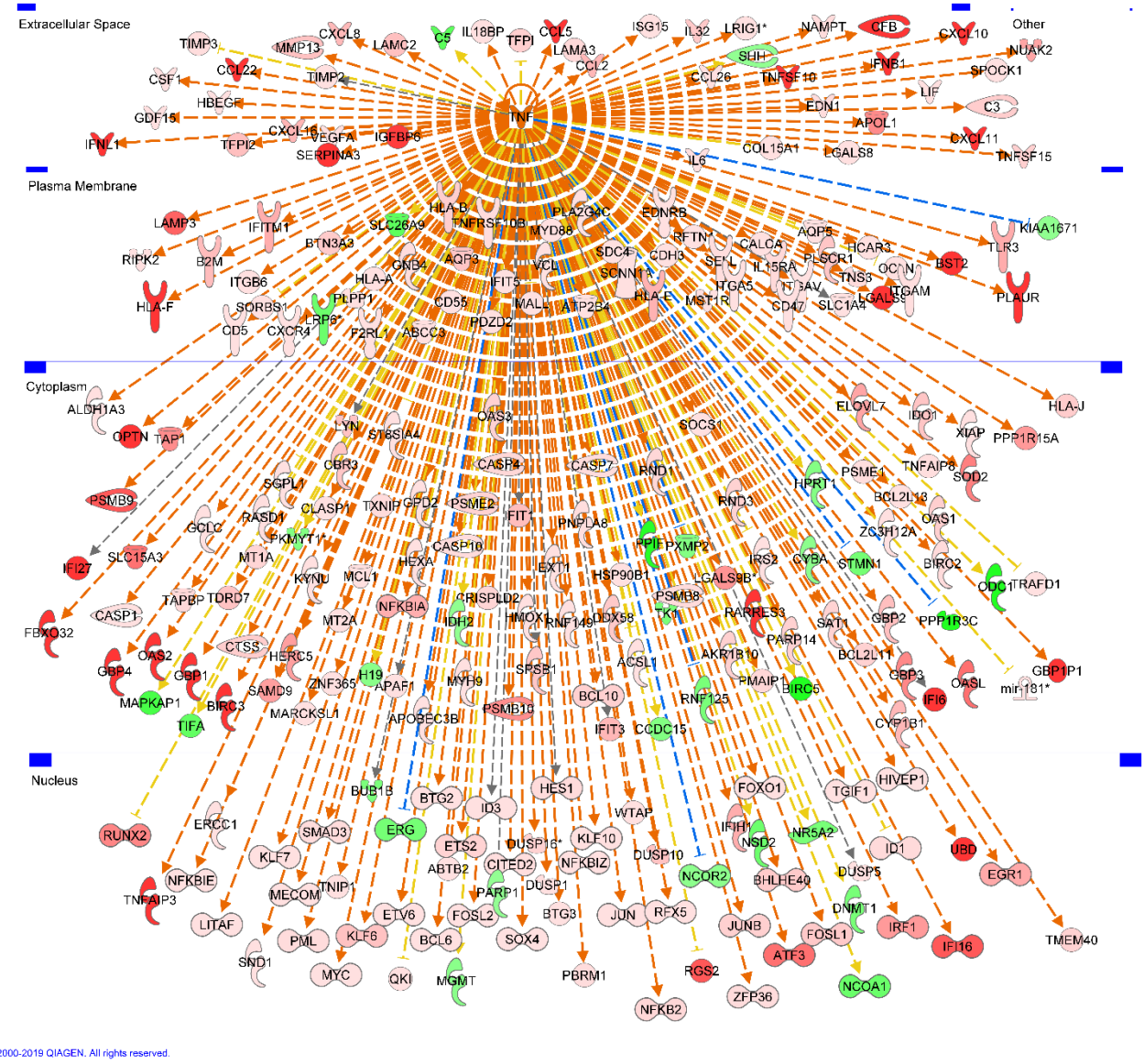


Figure 5.22: (further to table 5.15) **Network view of TNF as the predicted upstream regulator for the transcriptomic changes observed upon TMEFF2 overexpression in MCF-7 cells.** IPA Upstream Regulator Analysis was run on differentially expressed gene set between MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'. DGE Analysis filter; Fold change ≥ 2 to -2 ≤, p ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and source: Ingenuity Pathway Analysis – IPA (QIAGEN ®).

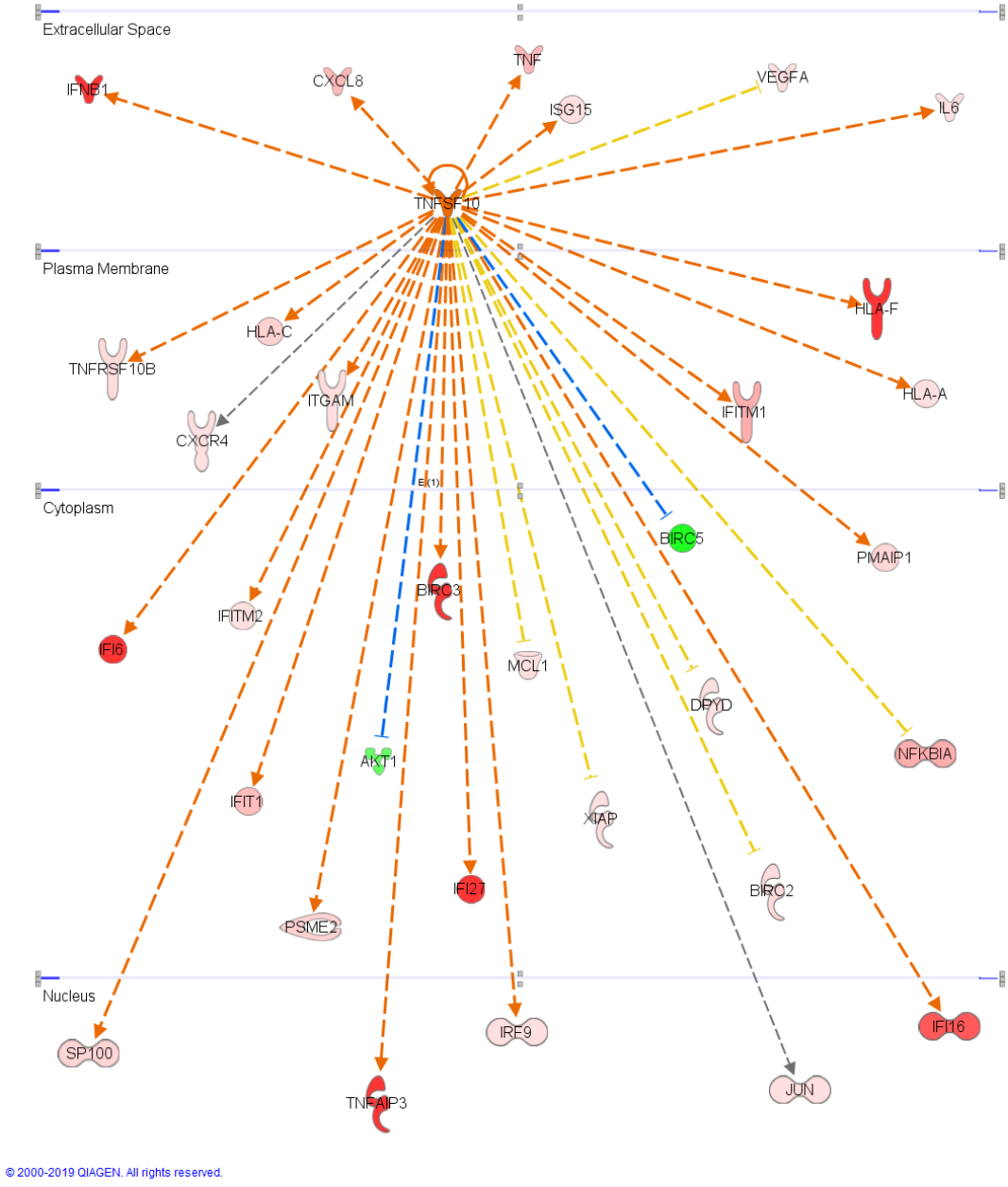


Figure 5.23: (further to table 5.15) **Network view of TRAIL as the predicted upstream regulator for the transcriptomic changes observed upon TMEFF2 overexpression in MCF-7 cells.** IPA Upstream Regulator Analysis was run on differentially expressed gene set between MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'. DGE Analysis filter; Fold change ≥ 2 to -2 ≤, p ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and image source: Ingenuity Pathway Analysis – IPA (QIAGEN ®).

5.2.7.3 Downstream Effects Analysis

The IPA also performed the Downstream Effects Analysis. This type of analysis is aimed at identifying the functions which are expected to increase or decrease as a result of the observed changes in the gene expression. The results of downstream analysis are depicted in figure 5.24. It was not surprising to observe that the 'organismal injury and abnormalities,' 'cancer 'and 'cell death and survival' were the top 3 functions which were expected to be modulated as the result of differentially expressed genes upon TMEFF2 overexpression. Moreover, given the strong interferon response, it was equally unsurprising to observe a decrease in the 'infectious diseases'. The Downstream Effects Analysis is primarily useful for investigating the effectors for which the downstream functions are not known. In my case, I already knew that the overexpression of TMEFF2 causes apoptosis in MCF-7 cells and I was more interested in investigating how TMEFF2 causes its previously known downstream effect rather than the other way around. Nevertheless, the results of the downstream analysis provided further evidence in support of the previous findings.

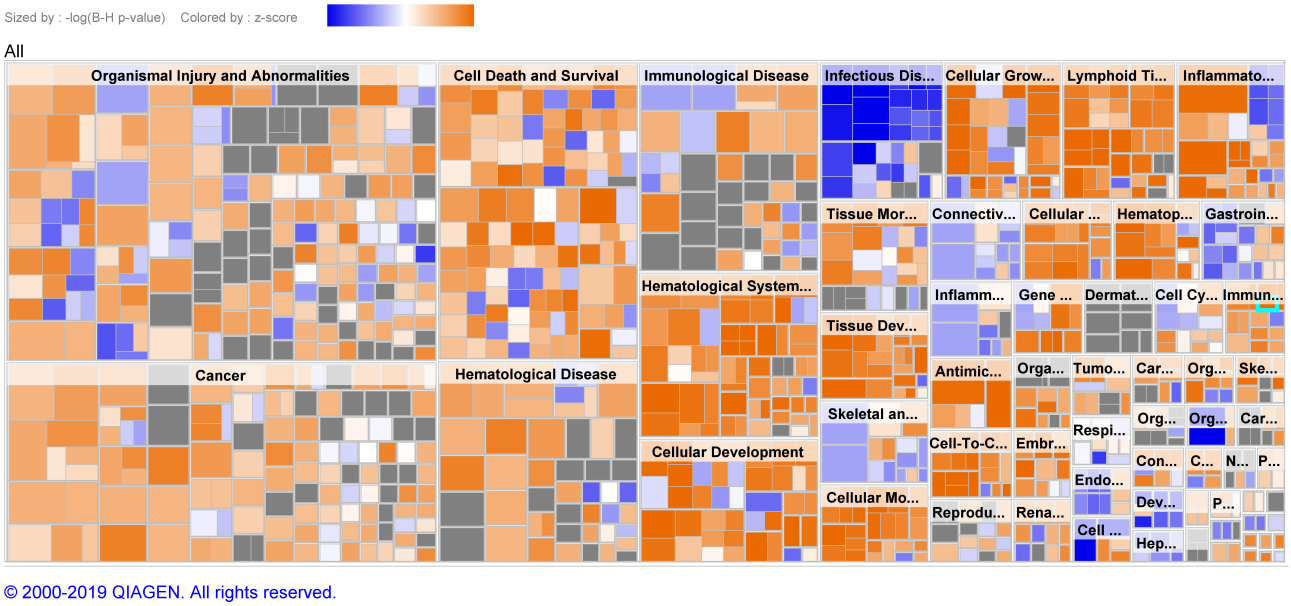


Figure 5.24: **IPA Downstream Effect Analysis – Diseases & Biological Functions Analysis – for the gene expression changes observed upon TMEFF2 overexpression in MCF-7 cells.** The boxes are sized according the Benjamini-Hochberg adjusted p values of overlap and the coloured according to the z-score. Analysis was run on the differentially expressed gene set between MCF-7 cells transfected with TMEFF2 *vs* MCF-7 cells transfected with luciferase 'control'. DGE Analysis filter; Fold change ≥ 2 to -2 ≤, p ≤ 0.005 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and image source: Ingenuity Pathway Analysis – IPA (QIAGEN ®).

5.2.7.4 Verification of differentially expressed genes

The results of differential gene expression analysis and the subsequent pathway analysis indicated a strong interferon response accompanied with the induction of apoptosis. It is important to note that a multitude of differentially expressed genes were closely related to each other. This fact in itself is a strong evidence for the validity of the data. Nevertheless, experimental validation through qRT-PCR was necessary. The differential gene expression analysis produced a large number of genes and choosing a small subset was difficult. Importantly, however advanced the algorithms for the pathway analysis become, it is just a tool for the projection of probable events. The IPA Canonical Pathways, Upstream Regulators Analysis and Downstream Effects Analysis, all compared the given transcriptomic changes with previous findings from a variety of systems. However, what was observed in other cell lines and/or *in vivo* models could not necessarily be applicable to MCF-7 cells. Providentially, findings from a previous study where MCF-7 cells were treated with TRAIL and subjected to differential gene expression analysis in 2008 were available. In this study by Kuman-Sinha *et al.* [48], a strong interferon response was induced in MCF-7 upon treatment with TRAIL. As TRAIL was among the top most upregulated genes upon TMEFF2 overexpression, I chose a subset of 5 genes closely related to TRAIL in the IPA Upstream Regulator Analysis (figure 5.23) for validation through qRT-PCR. In addition, the second most upregulated gene, CXCL10, was also included in the list of genes tested for validation. The results are depicted in figure 5.25.

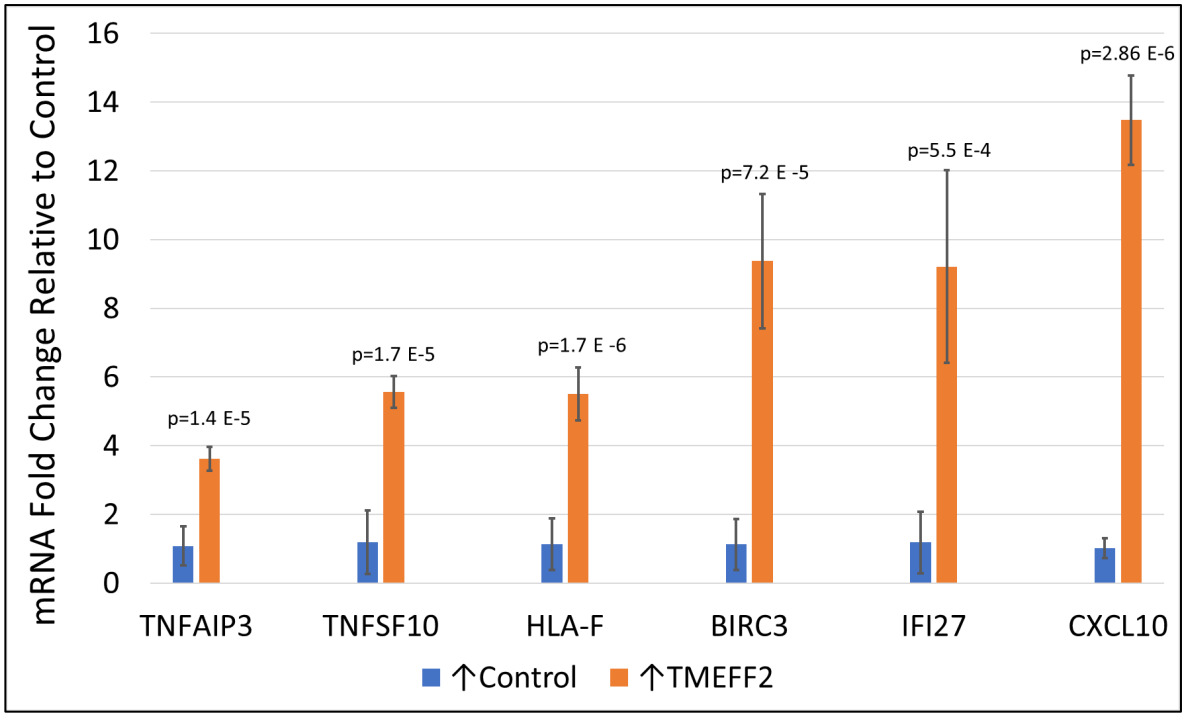


Figure 5.25: **qRT-PCR results for the verification of microarray data.** Relative mRNA fold change for each gene was calculated according to ΔΔCt method using L19 as the housekeeping gene. MCF-7 cells were transfected in 24-well plates with TMEFF2 and Luciferase 'control'. Cells were lysed 26.5-hours post-transfection. Cells lysates from 12 wells of a 24-well plate were pooled and RNA extracted. Process was repeated in triplicate, i.e., three 'control' and three 'TMEFF2' samples each were extracted from 12 wells of a 24-well plate. Bars represent the average from 2 PCR-reactions for three independent pools each; n = 6. Error bars represent +/- standard deviation. p values were calculated using two-samples t-test assuming unequal variances.

Table 5.16: Comparison of qRT-PCR and microarray data for the selective differentially expressed genes between MCF-7 cells transfected with TMEFF2 and MCF-7 cells transfected with Luciferase 'control'

Gene Symbol	qRT-PCR		Microarray	
	Average Fold Change	p-Value	Average Fold Change	p-Value
CXCL10	13.48	2.88×10^{-06}	235.49	8.33×10^{-07}
BIRC3	9.38	7.46×10^{-05}	39.56	5.76×10^{-05}
IFI27	9.22	5.48×10^{-04}	36.49	1.00×10^{-04}
HLA-F	5.51	1.67×10^{-06}	32.09	4.82×10^{-07}
TNFSF10	5.57	1.73×10^{-05}	25.49	6.09×10^{-05}
TNFAIP3	3.62	1.38×10^{-05}	37.99	2.55×10^{-06}

Relative mRNA fold change for qRT-PCR was calculated according to ΔΔCt method using L19 as the housekeeping gene. Expression fold change for the microarray data was generated by Transcriptome Analysis Console 4.0.1 (Affymetrix®) through dividing the expression value of test sample by that of the control sample. p value for the qRT-PCR was calculated using two-sample t-test assuming unequal variances while the p value for the microarray data was calculated by TAC 4.0.1 using one-way ANOVA with Empirical Bayes Statistics adjustment. For sample preparation, lysates from 12 wells of the 24-well plate were pooled and RNA extracted. Process was repeated in triplicate; three 'control' and three 'TMEFF2' samples each were extracted from 12 wells of a 24-well plate. For the qRT-PCR data Average Fold Change represents readings from two PCR-reactions for three independent pools each for control and test samples, n = 6, whereas for the microarray data, the fold change is calculated for 3 independent pools each, n = 3.

The qRT-PCR results revealed that although all 6 genes assayed followed the similar trend in terms of expression, the degree of upregulation observed was considerably less than those observed in microarray analysis (table 5.16). As the two techniques used to measure the gene expression are fundamentally different, one could not expect a close match between the data produced by them. Nonetheless, qRT-PCR data confirm the upregulation of this small subset of genes upon experimental overexpression of TMEFF2. Taken together with the fact that the gene set produced by differential gene expression analysis comprises a plethora of genes which overlap highly related functional groups, the validity of data produced through the microarray analysis can be accepted with adequate confidence.

5.2.8 Attempts of functional validation

In order to ascertain if TMEFF2 overexpression was causing apoptosis in MCF-7 cells through induction of interferon and/or TRAIL, I aimed to knockdown some of the crucial effectors in each pathway and check if the apoptotic effects of TMEFF2 were attenuated. There were two challenges in this approach. The first one was the choice of targets to knockdown, i.e., the induction of both the interferon signalling pathway and the TRAIL signalling occurs through a multitude of effectors and knocking down of all the effectors for each pathway was not reasonably possible. The second problem was the methodology of knocking down the target genes. Using siRNA would yield a higher degree of knockdown, but co-transfecting siRNA with plasmid DNA hardly produces the desired results (personal experience). In case the stable transfection or viral transduction with shRNA was to be employed, the cytotoxic antibiotic selection process causes the cells to become resistant to apoptosis (discussed later). With these problems in focus, I decided to use the shRNA plasmids with a fluorescent marker. In this approach, TMEFF2 was co-transfected (transiently) with the relevant

shRNA plasmid and assayed for apoptosis using the TMRE/DRAQ7™ staining measured through flow cytometry and importantly, only the cells expressing the fluorescent marker were gated for the subsequent detection of TMRE/DRAQ7™ fluorescence. I selected STAT1, FADD and TRAIL receptors A and B (separately) for the shRNA mediated transient knockdown. The reason behind choosing STAT1 was that its self-dimerization in type I interferon pathway and its dimerization with STAT2 in gamma interferon pathway are (in general) crucial for the induction of the respective interferon stimulated genes (ISGs). FADD was chosen because hypothetically, its knockdown would block the death receptor driven apoptosis. Lastly, knocking down the TRAIL receptors would cause a reduction in the TRAIL-induced apoptosis. The results of the duplicate experiments (depicted in figure 5.26) were deemed unreliable due to the fact that the pGIPZ plasmid which harbours the relevant shRNA sequences has a special variant of GFP (Turbo-GFP™) which is specifically designed for extensive early protein expression. Moreover, the marker lies under a CMV promoter. The extensive overexpression of this GFP variant was toxic for MCF-7 cells and despite co-transfecting TMEFF2 with the relevant shRNA plasmid in a 2:1 ratio (TMEFF2: pGIPZ), the background cell death in the control transfections was very high.

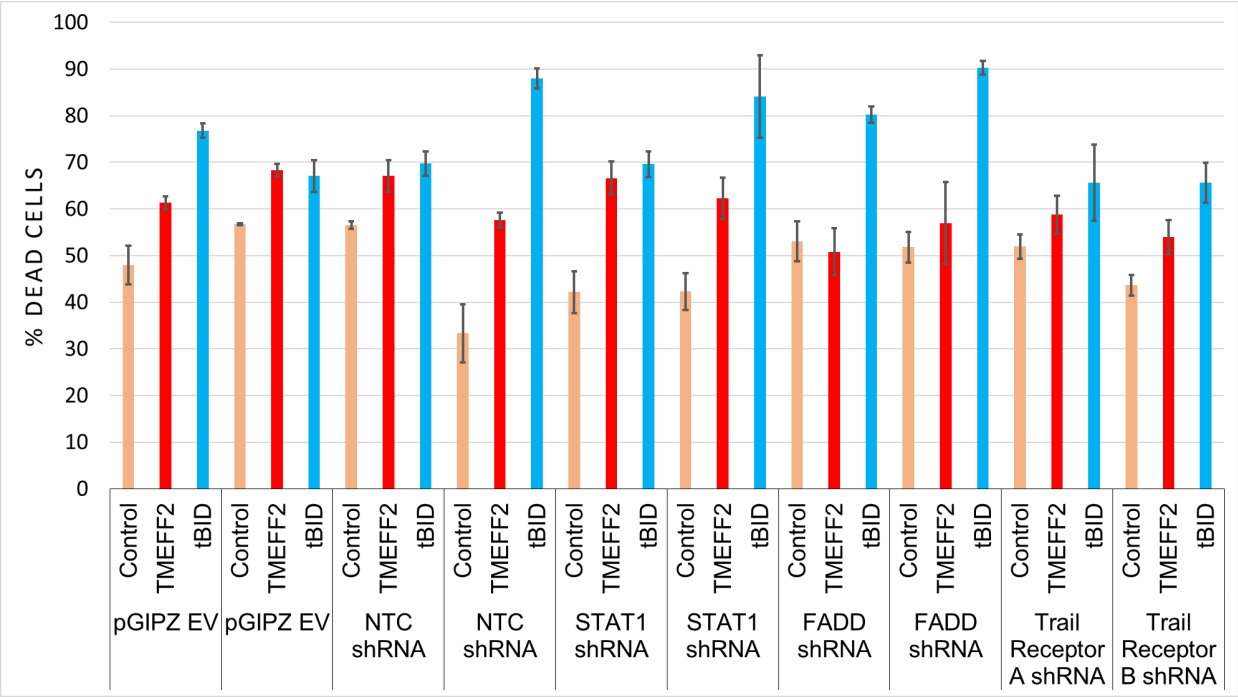


Figure 5.26: **TMEFF2-induced apoptosis upon knockdown of STAT1, FADD, TNFRSF10A and TNFRSF10B using TMEFF2 + shRNA co-transfection.** EV: empty vector, NTC: non-target control. TMEFF2

was co-transfected with the relevant empty vector or shRNA into MCF-7 cells in a 2:1 ratio (TMEFF2: shRNA) in a 24-well plate format (see material and methods). Luciferase was used as negative 'control' and tBID as positive control. Cells were harvested 72 hours post-transfection. Cell death was measured using TMRE/DRAQ7™ staining through flow cytometry. Only green fluorescent cells (GFP positive) were gated for further measurement of TMRE and DRAQ7™ fluorescence. Each bar represents average from 3 independently transfected wells of a 24-well plate. Error bars represent +/- standard deviation. The repeated groups represent duplicate experiments.

After the unsuccessful attempt of overexpressing TMEFF2 while concomitantly knocking down its putative apoptotic effectors (through co-transfection), I developed shRNA stable pools of MCF-7 cells with pGIPZ empty vector, STAT1, FADD and TRAIL knockdowns. The antibiotic was removed from the media 10 days before TMEFF2 transfection. Results are shown in figure 5.27. A minor reduction in apoptosis was observed in MCF-7 cells stably transfected with STAT1 and TRAIL shRNA. Contrary to the expectations, the apoptosis in FADD shRNA stable cells increased upon overexpression of both TMEFF2 and tBID (separately). Due to time constraints, I was unable to repeat the experiment or validate the knockdowns.

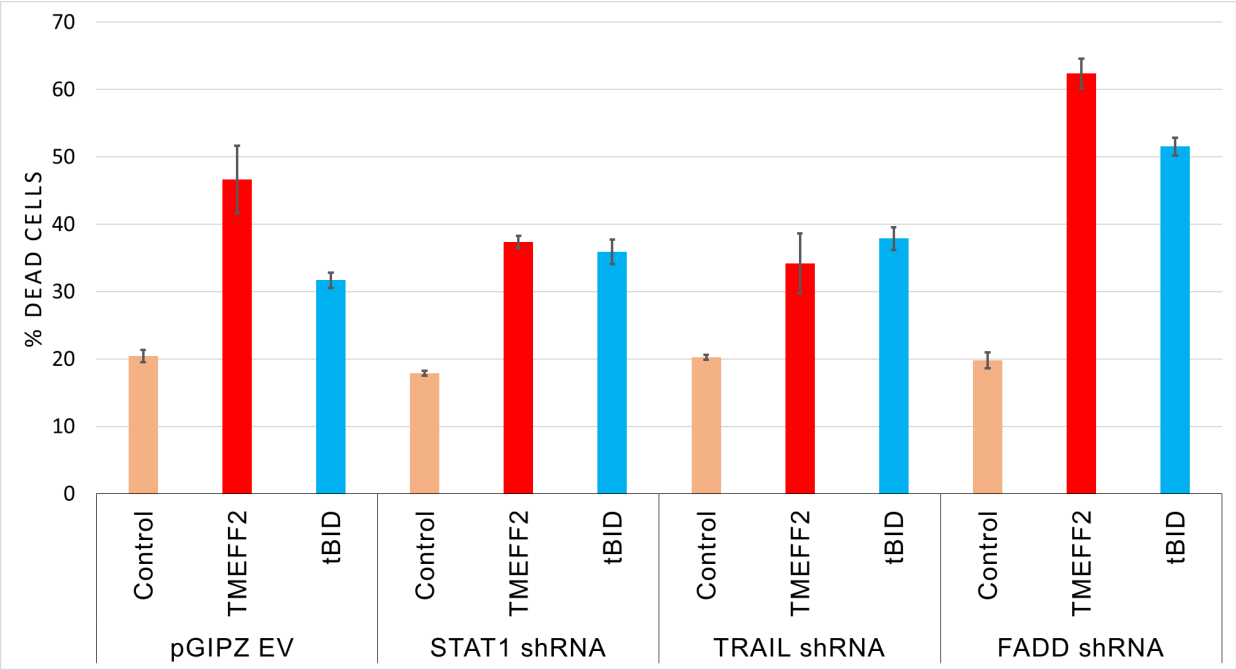


Figure 5.27: **TMEFF2-induced apoptosis upon stable knockdowns of STAT1, TRAIL and FADD.** EV: Cell death in STAT1 shRNA stable, TRAIL shRNA stable and FADD shRNA stable MCF-7 cells and MCF-7 cells stably transfected with pGIPZ1 empty vector (EV) upon TMEFF2 overexpression, detected using TMRE and

DRAQ7™ staining and flow cytometry. Stably transfected pools were selected by treating the MCF-7 cells with 0.5 µg/ml puromycin for ~1 week and maintained in 0.2 µg/ml puromycin thereafter. Puromycin was removed 10 days before transfection. Cells were transiently transfected with TMEFF2, luciferase (negative 'control') and tBID (positive control), harvested 48 hours post-transfection, stained with TMRE/DRAQ7™ and subjected to flow cytometry. Only GFP positive cells were gated in for measuring TMRE and DRAQ7™ fluorescence. Bars represent average from three different wells of a 24-well plate each with independently transfected population of cells. Error-bars represent the +/- standard deviation.

5.3 Discussion

A very strong interferon response along with a stalk upregulation of TRAIL were observed in MCF-7 cells upon TMEFF2 overexpression. As I have already mentioned, the up- and downregulation of a plethora of genes related to the same processes are themselves indicative of the validity of the data. Although the Ingenuity Pathway analysis employs an elaborate causal network approach where previously reported findings (peer reviewed) are used to relate the given changes in the transcriptome, even the most advanced algorithms used in pathways analyses only present the hypothetical projections as the connections observed in one system would not necessarily be transferable into the others. In this regard, two reported studies can greatly aid in a preliminary interpretation of the results of gene expression profiling and subsequent pathways analysis upon TMEFF2 overexpression in MCF-7 cells. The first study is by Elahi *et al.* [192] who observed that the overexpression of HPP1 in HCT116 colorectal cells resulted in activation of STAT1 signalling. The second study is by Kumar-Sinha *et al.* [48] where treatment of MCF-7 cells with TRAIL resulted in a stalk upregulation of genes involved in the interferon signalling pathway. The latter study is worthy of deep examination because it is highly relevant to the results presented in this chapter. In this study, firstly, Kumar-Sinha *et al.* [48] treated MCF-7 cells with TRAIL and studied the changes in gene expression profiles in MCF-7 cells at various post-treatment timepoints, i.e., 4, 8, 16 and 24 hours. Moreover, the TRAIL treated cells were concomitantly treated with zVAD-FMK to prevent apoptosis and cycloheximide in order to study only the primary response upon treatment with TRAIL. The genes upregulated at 16 hours post-treatment (authors call it intermediate primary response cluster) included TRNFAIP3, BIRC3, SERPINA3, MX2, IL6 and a host of cytokines and inflammatory molecules. The genes which were upregulated at 24 hours post-treatment included IFIT1, IFITM1, IFI16, IFI6, HLA-A, HLA-C, HLA-F, PSME2 and TLR3. The treatment of cells with interferons α, β or γ resulted in upregulation of a similar set of genes. Importantly, upon treating the cells with both TRAIL and interferons α, β or γ, the authors [48] state, '*treatment with IFN-α, -β, or -γ plus TRAIL for 24 h led to*

upregulation of two distinct clusters, apoptosis and NFκB pathway-related genes and IFN pathway related genes' (p. 581). The results obtained from the gene expression profiling and the consequent pathways analyses of TMEFF2 overexpressing MCF-7 cells were highly similar (if not the same) to what Kumar-Sinha quoted. Please note that all the major cytokines which were upregulated upon overexpression of TMEFF2 in MCF-7 cells are under transcriptional regulation of NF-κB. Moreover, Kumar-Sinha *et al.* [48] also reported that the concomitant treatment of MCF-7 cells with IFNβ and TRAIL resulted in a much higher apoptosis induction than the treatment of cells with TRAIL or IFNβ separately. According to Kumar-Sinha *et al.* [48], the upregulation of interferon pathways by TRAIL primed them for TRAIL-induced apoptosis. Kumar-Sinha *et al.* [48] employed immunoblotting for caspase 7 and PARP-1 to assay apoptosis; the results of their experiment (TRAIL + IFNβ) appear very similar to the results presented in figure 5.4. Furthermore, the strong apoptosis induced by the combined treatment with TRAIL and IFNβ was fully abolished upon treatment of cells with zVAD-FMK. Similar results were reported in another study where IFNγ and TNFα sensitised the TRAIL resistant SW620 colon carcinoma cells to TRAIL-induced apoptosis [208]. When considered in the context of results of TMEFF2 overexpression in MCF-7 cells, the potent apoptosis induction, a strong interferon response and the substantial and significant upregulation of TRAIL could be indicative of a concomitant TRAIL and interferon induction in MCF-7 cells as the mode of TMEFF2-induced apoptosis. Taking together these results and arguments, I chose the TRAIL-related genes for qRT-PCR validation of the results obtained from the microarray analysis. Furthermore, Kumar-Sinha *et al.* also reported that *TRAIL* itself was upregulated upon both treatment with TRAIL and treatment with interferons α, β or γ (data for only 24 hours post-treatment was presented for treatment with interferons). The question of the TMEFF2 overexpression inducing the upregulation of TRAIL through first activating the interferons pathway or the other way around, amounts to the chicken and egg paradox. Nevertheless, one way or the other, the end result would be a combination of TRAIL overexpression and interferon induction. It would have been highly beneficial if I could have performed a comparison analysis between the two datasets, i.e., one obtained from gene expression analysis upon TMEFF2 overexpression in MCF-7 cells and the other one generated by Kumar-Sinha *et al.* following the treatment of MCF-7 cells with TRAIL, interferons α, β or γ. Despite the most thorough search through both Gene Expression Omnibus (GEO) from NCBI and ArrayExpress from EBI, I could not find the gene expression analysis data from Kumar-Sinha *et al.* [48], nor did they respond to my request for the dataset.

A counterargument to the concomitant induction of TRAIL and/or interferons being the primary mode of TMEFF2-induced apoptosis in MCF-7 cells could be that Kumar-Sinha *et al.* [48] also reported that the treatment of MCF-7 cells with TNFα or Fas ligand mimicking antibody (Fas Ab) resulted in upregulation of genes related to the interferon pathway. However, in my study, upon TMEFF2 overexpression in MCF-7 cells, TNFα was upregulated by only 7.29-fold, whereas the expression levels of TRAIL mRNA increased 25.49-fold. In addition, TRAIL was among the 30 most upregulated genes upon TMEFF2 overexpression in MCF-7 cells (table 5.5). On the other hand, TNFα showed a 3 times higher activation z-score in the IPA Upstream Regulator Analysis. One of the reasons behind a higher z-score for TNFα could be that a larger number of downstream targets are known for TNFα as compared to TRAIL. Therefore, the question remains if the primary mode of TMEFF2-induced apoptosis in MCF-7 cells happens to be the induction of interferons, upregulation of TNFα, upregulation of TRAIL or a combination of all. It would be a daunting task to dissect the exact ignition factor. As I mentioned in the introduction, in my view, the very advantage of anticancer genes is their ability to impart a multitude of downstream effects in a heterogeneous population of cancer cells. Regardless of the causative factor, the overexpression of TMEFF2 in MCF-7 cells resulted in the induction of all three types (type I, type II, type II) of interferon signalling along with the upregulation of TNFα and TRAIL.

Another counterargument to TRAIL + interferon concomitant induction as the mode of TMEFF2-induced apoptosis could be that Kumar-Sinha *et al.* [48] only observed the upregulation of *TRAIL* in the late cluster (24 hours post-treatment) of upregulated genes upon treatment of MCF-7 cells with recombinant TRAIL. I harvested cells at 26.5 hours post-transfection; a close observation of the time lapse microscopy conducted after co-transfecting MCF-7 cells with TMEFF2 and GFP (figures 5.5 and 5.6 and table 5.2) revealed that the expression of GFP was detected as early as 12-14 hours post-transfection. If TRAIL was upregulated at 12- 14 hours post-transfection and cells were harvested at 26.5 hours post-transfection, the cells would have experienced the treatment with secreted TRAIL for merely 14 hours – i.e., unless there is a significant difference between the activity-dynamics of TRAIL upon upregulation *vs* treatment with recombinant TRAIL, the *TRAIL* should not have been observed among the most upregulated genes upon TMEFF2 overexpression at 26.5 hours post-transfection.

A question which should be addressed, however, is how do interferons induce apoptosis? Moreover, how relevant is interferon therapy with respect to cancer treatment? To address the latter first,

despite getting substantial attention in the late 1970s to early 1980s as a potential 'miracle cure', interferon therapy failed to gain the support as the anticancer therapeutic agent of choice [209]. There could be four main reasons behind this lack of success: (i) interferon therapy hardly resulted in a cure in the clinical trials, (ii) expectations were too high, i.e., no considerable side effects were expected but the interferon therapy resulted in the typical flu-like symptoms and occasionally some worse side effects were reported (e.g., acute respiratory arrest was observed in one patient [209] which was enough to heighten the scepticism), (iii) interferons were used as immunomodulatory agents when little was known about anticancer immunology and (iv) as Jenny Bryan [209] puts it, according to Dr Priestman, '*The demise of interferon was partly due to a long-term sense of disillusion in the immune system approach, but also to the arrival of a new generation of biological agents, such as the tyrosine kinase drugs and antiangiogenesis agents, which are much more exciting, and for which the underlying biological mechanisms are much better established*' (p. 637). In addition, not all cells and cancer types are sensitive to interferon-induced apoptosis [210]. In this regard, the matter appears to be similar to that of TRAIL (explained in detail in the introduction). Also, to the best of my knowledge, a combination of various interferons has not been used in the clinical trials, nor could I find a reference where TRAIL and interferons were used in combinational therapy in any clinical settings, though there are several reports available for the onco-suppressive and anti-tumoral effects of interferons in animal models [211]. However, a Phase I-II Study of IFNγ combination treatment with paclitaxel, Trastuzumab and Pertuzumab in patients with HER-2 positive breast cancer is underway [212].

Being inflammatory cytokines themselves and potent inducers of cytokines in malignant cells, the role of interferons and the interferon stimulated genes (ISGs) in tumour microenvironment is noteworthy (reviewed elaborately by Cheon *et al.*, Ref. [213]). Inflammation can act as a double-edged sword when it comes to tumour progression: the inflammatory response itself is anti-tumoral, but when inflammation becomes chronic, it can aid the tumour growth and progression by inducing immuno-tolerance, supplying the growth promoting factors in the microenvironment, enhancing the cell-death-and-renewal cycles and also by adding to the size of tumour through the recruitment of immune cells [1, 211]. Research into the field of inflammation and tumour microenvironment has gained significant attention recently. Considerable effort is being put into inhibition of the tumour immuno-tolerance and enhancement of the tumour immunogenicity. With regards to tumoral inflammation, by and large, the role of interferons is onco-suppressive as they can induce apoptosis,

induce growth arrest, inhibit angiogenesis and augment the activation of dendritic cells and natural killer (NK) cells [213]. On the other hand, chronic exposure to low levels of interferons and the resultant partial downstream signalling (including incomplete phosphorylation of STAT1) can give rise to a small subset of ISGs which are termed as the interferon-related damage-resistant signature (IRDS). The presence of IRDS in tumours is indicative of reduced survivability and worsening outcomes of radio- and chemotherapy. Compared to the vast list of ISGs, the ones related to IRDS constitute only a small subset of approximately 30-odd ISGs. Moreover, even in the presence of IRDS, the provision of the high dose of exogenous interferons can reconstitute their anti-tumoral effects [213]. The interferon inducing toll-like receptors (TLRs) and the stimulator of interferon genes (STING) are gaining popularity as the targets in cancer immunotherapy [211, 214].

As a plethora of inflammatory cytokines and chemoattractants are upregulated upon overexpression of TMEFF2, there is a strong chance that the TMEFF2-induced apoptosis in MCF-7 cells is immunogenic. This could be an interesting prospective line of investigation. Testing for the activation of immune cells in culture after treating them with conditioned media from TMEFF2-transfected MCF-7 cells could reveal the potential immunogenicity of TMEFF2-induced apoptosis. A noteworthy cytokine which was upregulated upon TMEFF2 overexpression in MCF-7 cells is the C-C motif chemokine 22 (CCL22) – it was the most upregulated gene. Canonically speaking, CCL22 is considered as the main chemoattractant for the regulatory T-cells (Tregs), which in turn are responsible for immunosuppression within the tumour microenvironment. The presence of a Tregs-specific forkhead box transcription factor, FOXP3, is indicative of poor survival and cancer immuno-evasion [1]. However, according to a study by Freir *et al.* [215], Tregs (along with CCL22) play an opposite role in the case of breast cancer – in a cohort of 80 breast cancer patients, there was no correlation found between the stages of cancer and the presence of Tregs in tumours. In fact, the presence of FOXP3-positive Tregs had a negative correlation with the nodal metastasis of breast cancer [215]. Studies like this one are indicative of the vastly complex nature of tumour-immunology.

As to how interferons induce apoptosis, the answers are more complex and interesting than anticipated. All three types of interferons are capable of inducing apoptosis in a multitude of malignant cells [210]. Moreover, the interferon-induced apoptosis exhibits classic morphological features of apoptosis: detachments, rounding, membrane blebbing, DNA fragmentation as well as loss of mitochondrial potential and cytochrome *c* release [210, 216]. Furthermore, the interferon-

induced apoptosis is usually independent of p53 status and BCL2 activation [216]. Please note that upon TMEFF2 overexpression in MCF-7 cells, caspase-dependent and p53-independent apoptosis was observed which also involved the loss of mitochondrial membrane potential (figures 5.3 to 5.7). The results presented in figure 5.3c (c.iv) are particularly interesting, as upon inhibition of caspase activity by zVAD-FMK, cells appear to be 'stuck' in early apoptosis with extensive loss of mitochondrial membrane potentials. Continuing the discussion regarding the mechanisms behind interferon-induced apoptosis, in addition to the well-known apoptotic genes like TRAIL, TNFα, FAS and caspases, several other genes in the list of interferon stimulated genes (ISGs) are capable of inducing apoptosis on their own. These include (but are not limited to) x-linked inhibitor of apoptosis associated factor-1 (XAF1), OAS1, OAS2, OAS3, IFI27, IFITM2, IFIT2 and very interestingly, CXCL10. Irrespective of its role as a cytokine, CXCL10 has also been reported to induce apoptosis in several cell lines in tissue culture [217, 218, 219, 220]. Interestingly, Sistigu *et al.* [221] reported that anthracyclines induce a rapid and strong type I interferon response in malignant cells, and the tumours lacking interferon α receptors failed to respond to anthracyclines unless CXCL10 was artificially supplied. Conversely, the study also reported that the lack of interferon-related-signature in tumours was indicative of a poor response to anthracyclines in breast cancer. Please note that CXCL10 was the second most upregulated gene upon overexpression of TMEFF2 in MCF-7 cells. Besides, although JAK/STAT pathway occupies the central role when it comes to interferons signalling (reviewed in detail in Ref [222]), several additional factors can also play very unconventional and interesting roles with regards to interferon-induced apoptosis. For example, Thyrell *et al.* [223] reported that IFNα induced caspase-dependent apoptosis in U266 lymphocytes by activation of BAX, BAK and the subsequent loss of mitochondrial membrane potential. Interestingly, this IFNα-induced apoptosis was abolished upon inhibition of PI3K/AKT pathway (upstream of mitochondria). Furthermore, Panaretakis *et al.* [224] demonstrated that IFNα can induce apoptosis in the immortalized human epidermal keratinocytes (RHEK-1) without affecting the JAK/STAT signalling and the subsequent downstream transcription of ISGs. Also, in this study, IFNα-induced apoptosis was shown to be dependent on the ERK1/2 and JNK signalling, downstream of PI3K/AKT signalling pathway. Possibly, the constitutionally active PI3K/AKT pathway in MCF-7 cells [125] could be aiding in the interferon + TRAIL induced apoptosis upon TMEFF2 overexpression. In an earlier publication, Thyrell *et al.* [225] had reported that the IFNα-induced apoptosis in U266 and RHEK-1 cells was independent of FAS-receptor signalling. As to how interferons are unable to induce apoptosis in some instances, Chawla-Sarkar *et al.* [216] propose the antiapoptotic factors such as XIAP and cFLIP to play a central role in this regard, whereas Kotredes

and Gamero [210] suggest an important role of NF-κB. However, NF-κB was among the activated regulators in the IPA upstream regulators analysis of TMEFF2 overexpressing MCF-7 cells. The reason behind NF-κB being flagged as the activated upstream regulator could be that the majority of cytokines in the list of ISGs also occur to be under the positive transcriptional control of NF-κB complex. Nonetheless, even if the NF-κB complex was actively playing the anti-apoptotic role, the effect of TMEFF2 overexpression must have successfully overcome these anti-apoptotic effects.

Still, the results of gene expression profiling upon TMEFF2 overexpression in MCF-7 raise some important questions. Firstly, after observing a very strong interferon response in MCF-7 cells upon TMEFF2 overexpression, one could argue that the interferon response in MCF-7 cells could have been triggered by the introduction of large amounts of double stranded DNA into the cells. There are two strong arguments against this possibility: firstly, the control cells were also transfected with double stranded DNA. Moreover, in order to avoid detection of non-specific signal due to cell stress induced by excessive translation of an ectopic gene, instead of using an empty vector, the control batches of MCF-7 cells were transfected with the plasmid harbouring luciferase. Secondly, through the use of microarrays for gene expression profiling, Elahi *et al.* [192] had also demonstrated that the overexpression of TMEFF2 in HCT116 colorectal cells resulted in the activation of STAT1-interferon pathway. Although they used empty vector as control plasmid, they generated stable clones for their microarray study which should have abolished the interferon induction in response to the presence of a foreign double stranded DNA in the cytoplasm. Elahi *et al.* [192] state, '*18 of the 42 genes that were upregulated greater than 3-fold are known to be interferon-inducible genes*' (p. 1571). Additionally, they also noted that the number of genes upregulated was considerably higher than those downregulated (8 ≥ 3-fold downregulated genes compared to 42 genes ≥ 3-fold upregulated ones). In a later study by the same group, Hernandez *et al.* [194] republished the same differentially expressed genes set but, in addition, the study reported that HPP1 upregulation in HCT116 cells activated both type I and type II interferon signalling pathways. Additionally, Hernandez *et al.* [194] also reported that HPP1 overexpression marginally sensitized HCT116 cells to IFNα-induced growth arrest but not to interferon induced apoptosis (HCT116 cells are otherwise insensitive to interferons). Neither Elahi *et al.* [192] nor Hernandez *et al.* [194] report any gene ontology analysis and they did not report observing TRAIL or TNFα in the list of differentially expressed genes in HPP1 overexpressing HCT116 cells either. Elahi *et al.* did not make their dataset available on any microarray databases. Moreover, the depth of transcriptomic analysis (both in terms of number of transcripts and fold

change observed) that Elahi *et al.* reported was at least 20 times less than the results of the gene expression profiling of TMEFF2 overexpressing MCF-7 cells.

The discrepancies between the results obtained from the microarray transcriptomic analysis and qRT-PCR of the selective subset of genes were surprising. Although the upregulation of all six genes tested were validated through qRT-PCR, the degree of upregulation was lower by several orders of magnitude. There could be several explanations for this apparent discrepancy. Firstly, the cDNA solution used for the qRT-PCR was diluted 5 times in order to avoid non-specific primer binding and also to conserve the resources which I believe have costed in lowering the sensitivity of the qRT-PCR assay. Secondly, the methods used to measure the signals and calculate the fold change for qRT-PCR and microarray analysis were fundamentally different. Relative mRNA fold change for qRT-PCR was calculated according to ΔΔCt method using L19 as housekeeping gene whereas the expression fold change observed in microarrays was calculated after subjecting the signal intensities of multiple probe sets to rigorous averaging algorithms (Tukey's Bi-weight average algorithm, using multiple housekeeping genes through transcriptomic analysis console TAC 4.0 software). Conventionally, qRT-PCR would be expected to have a higher sensitivity and lower specificity. However, I would argue that the ability to measure the exon-level expression profiles along with a much more sophisticated cDNA generation increases the sensitivity of this modern microarray analysis many fold compared to qRT-PCR. For example, for the qRT-PCR of BIRC3, I designed a pair of primers which amplified only one exon junction, whereas 10 exons of BIRC3 mRNA could hybridize to more than 30 probe-sets on the Affymetrix® GeneChip™ Clariom™ D microarray (supplementary figure S.13). Additionally, a highly sophisticated cDNA preparation method was used for the microarray analysis (Ovation® Pico WTA System V2 from NuGEN Technologies Inc.) which involved a multistep process of first generating the tagged cDNA using an optimized mix of random and oligo dT primers, followed by fragmentation of mRNAs and generation of dsDNA using the tag specific primers, etc. On the other hand, for qRT-PCR, a relatively simple isothermal reverse transcription was employed. Notably, for the optimization of qRT-PCR sensitivity in the experiments related to the forkhead box proteins (chapter 4), interventions in the reverse transcription reaction were able to deliver the desired results, i.e., increasing the sensitivity of the assay and achieving the optimum PCR reaction efficiencies. Again, the change in expression of a large number of genes related to the same processes is itself indicative of the validity of the data. Nevertheless, qRT-PCR of a larger subset of genes could indeed provide a higher degree of validity to the dataset obtained through microarray analysis.

Besides the findings detailed previously, Kumar-Sinha *et al.* [48] also observed that when MCF-7 cells stably transfected with the dominant negative FADD were treated with TRAIL, TNFα, or Fas Ab, the induction of the interferon signalling was abolished. I also imagined FADD to be the common factor in the death receptor signalling and that is why I included it in the attempted functional validations presented in figures 5.26 and 5.27. I have already explained the reasons behind choosing the transient co-transfection with TMEFF2 and the shRNAs against these targets for the attempted functional validation. Briefly, when the dependence of the TMEFF2-induced apoptosis on p53 was being assayed, I employed antibiotic selection for generating the p53 shRNA stable MCF-7 cells (figure 5.6). Upon TMEFF2 overexpression in wild type MCF-7 cells, p53-knockdown MCF-7 cells, and the MCF-7 cells stably transfected with the empty vector, the TMEFF2-induced apoptosis was reduced even in the cells stably transfected with the pGIPZ empty vector (compared to wild type MCF-7 cells). Moreover, the positive control, tBID, exhibited some surprising results where the tBID-induced apoptosis was reduced upon the knockdown of p53. The latter could be attributed to the non-specific effects of the p53 shRNA; however, for the reduction in TMEFF2-induced apoptosis in the MCF-7 cells stably transfected with empty vector, the reason could be that when treated with puromycin, a sub-population of cells which were relatively resistant to apoptosis was selected. This premise could be aided by a close observation of the three populations of MCF-7 cells transiently transfected with luciferase, i.e., control protein (figure 5.6 a). One could notice that even the background cell death (i.e., induced by the transfection process) was reduced in the MCF-7 cells stably transfected with either the pGIPZ empty vector or p53 shRNA. Nonetheless, no significant reduction in TMEFF2-induced apoptosis was observed between MCF-7 cells stably transfected with p53 shRNA and those which were stably transfected with pGIPZ empty vector. Still, owing to the possible limitations of this experimental technique, I used the potent p53 inhibitor, Pifithrin-alpha (PFT- α), in order to retest the dependence of TMEFF2-induced apoptosis on the activity of p53; the results of this experiment were self-evident (figure 5.7). Taken together, these results support the premise that the TMEFF2-induced apoptosis in MCF-7 cells is not dependent on the activity of p53.

Owing to the difficulties faced during stable transfection with p53 shRNA and pGIPZ empty vector using antibiotic selection (i.e., emergence of apoptosis resistant cells), a different approach was devised for knocking down the likely contributing factors in TMEFF2-induced apoptosis which were identified through gene expression analysis. As pGIPZ vector, which harbours the shRNAs, also holds the Turbo-GFP™ expression cassette, I co-transfected TMEFF2 with the relevant shRNA plasmid and

assayed for apoptosis using TMRE/DRAQ7™ staining using flow cytometry and gated in only the green fluorescent cells. The emission spectra of TMRE, DRAQ7™ and GFP were significantly separated and no crossover was detected. But the experiment was unable to yield any reliable results because the excessive early expression of Turbo-GFP proved to be toxic for MCF-7 cells and extensive background cell death was detected (figure 5.26). I had to resort back to antibiotic selection, but instead of removing the antibiotic 24 hours before the start of the experiment (transfecting the cells with TMEFF2 and controls), this time it was removed 10 days before the transfection. It is possible that the removal of selection pressure could have caused the cells to lose the genome-integrated shRNAs – as the knockdowns could not be confirmed, one can only postulate this possibility. Furthermore, the process of stable transfection using antibiotic selection is primarily dependent on chance, i.e., there is no control over how plasmid DNA is cleaved and integrated into the host genome [226] or, to be precise, there is no guarantee of integration of sequences for shRNA into the host genome (even if the selection is successful and the fluorescent markers are expressed). The loss of shRNA sequence after a few generations, despite the stable integration of selection and florescence markers is a common occurrence even when the selection pressure is kept applied. As the antibiotic was removed from the media 10 days before the transfection with TMEFF2, there is a strong chance that the sub-population of cells which did not have the shRNA sequence would have flourished. Also, by convention, several shRNA sequences are normally tried in gene knockdown attempts because, firstly, the specificity of shRNA sequences for the target genes vary in efficiency and, secondly, a multitude of variables can affect the action and stability of shRNA transcripts. The choice of the right sequence is largely achieved through the empirical approach. Due to the time constraint, I could not try several shRNA sequences, nor could I verify any knockdowns through qRT-PCR. Nevertheless, even if the right shRNA sequences were used and the desired genes were knocked down successfully, the functional validation of the likely targets in the multitude of signals induced by TMEFF2 would still have posed a significant challenge. In the case of TRAIL, for example, even if a small number of cells in a heterogeneous population in tissue culture were to have insufficient knockdown of *TRAIL*, the secreted ligand would be able to diffuse across the media and induce apoptosis in the rest of the cells. I also considered adding the TRAIL neutralizing antibody to the media of MCF-7 cells transfected with TMEFF2. First, a large amount of antibody would be needed and, secondly, there was still a chance of juxtracrine intramembranous signalling by TRAIL, even before the antibody could neutralize it. In the case that a strong knockdown of *STAT1*, *STAT2* or their respective JAKs could be achieved, as I have explained previously, the TMEFF2-induced (potentially) interferon-dependent apoptosis could

be achieved through alternative pathways. In the case of a successful *FADD* knockdown, the vast list of apoptotic ISGs could still induce apoptosis independent of death receptors. Regarding the option of using small molecular inhibitors of the likely targets of TMEFF2 in MCF-7 cells, to the best of my knowledge, no small molecule inhibitors are available for any of the interferons, FADD or STATs. There are, however, small molecular inhibitors available for JAKs [222]. The conceivable problems with using the small molecule inhibitors for JAKs would be a widespread inhibition of the whole of JAK/STAT downstream signalling including STAT3, STAT4, STAT5, etc. The JAK/STAT pathway plays several physiological roles [222] and inhibition of the entire JAK/STAT signalling would possibly lead to extensive non-specific effects thus obscuring the possible targets of TMEFF2 within the JAK/STAT pathway. Having said that, as Kumar-Sinha *et al.* [48] reported, the abolishment of interferon response in MCF-7 cells stably transfected with the dominant negative FADD upon treatment with TRAIL, TNFα, or Fas Ab, using the co-transfection of TMEFF2 and the dominant negative FADD followed by apoptosis assay, could potentially prove to be a successful strategy for checking if TRAIL and/or TNFα have a significant role in TMEFF2-induced apoptosis. In case even a minor reduction in TMEFF2-induced apoptosis in MCF-7 cells overexpressing the dominant negative FADD could be observed, this could also be a means to indicate if TMEFF2 overexpression induces extrinsic apoptosis in MCF-7 cells.

Lastly, as the literature review suggested that promoter methylation of TMEFF2 is the primary source of its silencing, checking the effects of 5-Aza-2'-deoxycytidine (DNA methyltransferase inhibitor) treatment on the expression levels of TMEFF2 in MCF-7 cells could be an interesting prospective experiment. Interestingly, at least two studies (Ref [227] and [228]) have reported that the treatment of MCF-7 cells with 5-Aza-2'-deoxycytidine results in a substantial decrease in cell proliferation.

Summarising the discussion, TMEFF2-induced apoptosis could have several causative sequences of events. The sequences of events could include (but are not limited to) (i) TMEFF2 → TRAIL → interferons → TRAIL + interferons + ISGs → apoptosis, (ii) TMEFF2 → interferons → TRAIL → TRAIL + interferons + ISGs → apoptosis, (iii) TMEFF2 → interferons → ISGs → apoptosis, or (iv) TMEFF2 → TNFα → interferons → TRAIL + TNFα + interferons + ISGs → apoptosis. Irrespective of the sequence of events, the end result is the upregulation of a plethora of apoptotic genes including the death receptors as well as their ligands. Again, in my view, the true strength of anticancer genes lies in their ability to impart a multitude of cell death-inducing effects. In the rapidly evolving heterogeneous

population of cancer cells in tumours, if a subpopulation of cells could have developed evasive strategies against death receptor signalling for example, these cells could still be sensitive to any of several ISGs which are capable of inducing apoptosis independently.

6 FBLN5 as an Anticancer Gene

As I have already explained in the beginning of the previous chapter, I had decided to investigate the functioning of FBLN5 in CV-1 and COS-7 cells. Briefly, there were three reasons behind this decision: (i) the cMYC overexpressing CV-1 cells that Qize Ding had used as the transformed cell model had not been reported to have the ability of anchorage independent growth, (ii) if HEK293T cells were used as a model for transformed cells, the apparent genetic differences between HEK293T cells and CV-1 cells would have posed challenges while trying to decipher the differences in function(s) of the anticancer gene between the two systems, and (iii) a transformed version of CV-1 cells was already available in the form of the well-established COS-7 cells, a cell line developed by transforming the CV-1 cells using the SV40 origin defective DNA [128]. In this chapter, I shall be discussing the functioning of FBLN5 in CV-1 and COS-7 cells. Moreover, I shall also report the experiments which led to the decision of studying the functioning of TMEFF2 in MCF-7 cells and that of FBLN5 in CV-1 and COS-7 cells.

6.1 FBLN5, a Brief Overview

Fibulin-5 (FBLN5), also known as DANCE, EVEC, UP50, ARMD3, 'embryonic vascular EGF-like repeat-containing protein' and 'developmental arteries and neural crest EGF-like' [130, 131] is a 488 amino acid long secreted glycoprotein. FBLN5 belongs to a small but versatile family of secreted extracellular matrix (ECM) proteins known as fibulins. The name was derived from the Latin word *fibula* which means buckle, because the first fibulin was originally thought to connect β1 integrins with the cytoskeleton [229]. Fibulin family proteins are predominantly involved in the formation and stabilization of microfibrils, elastic fibres, loose connective tissue and basement membranes. In addition, fibulins also have roles in cell-ECM interaction, integrin signalling, neoplastic transformation and modulation of cell proliferation [229]. Similar to the typical ECM proteins, fibulins harbour a signal peptide and have modular protein structures. Additionally, fibulins contain a characteristic C-terminal fibulin domain. Fibulin-1 and 2 are the large multi-modular ECM proteins containing N-terminal anaphylatoxin modules, cysteine rich domains and long tandem repeats of calcium binding EGF (cbEGF)-like domains (8 repeats in fibulin-1 and 11 repeats in fibulin-2). Fibulin 3, 4 and 5 are relatively smaller and simpler. Each contains six tandem repeats of cbEGF-like modules followed by the C-terminal fibulin domain. The cbEGF-like domain has a characteristic disulphide bridging (in the

pattern Cys 1–3, Cys 2–4 and Cys 5–6) which imparts mechanical properties. Together with calcium binding, the tandem arrays of cbEGF constitute rigid rod-like structures. The first N-terminal cbEGF-like domains in fibulins 3, 4 and 5 are interrupted with a proline rich insert, the length of which varies between these three proteins [100, 229].

The distinguishing feature of fibulin-5 is the presence of an integrin binding RGD motif. Through this Arg-Gly-Asp tripeptide, fibulin-5 is capable of binding αvβ3, αvβ5, and α9β1 integrins. Moreover, fibulin-5 binds numerous ECM proteins including tropoelastin, elastin microfibril interface-located protein-1 (ELMILN-1), apolipoprotein A, and lysyl oxidase-like 1 (LOXL1) [100]. Nakamura *et al.* [230] suggested the bridging of integrins with elastic fibres to be the primary role of fibulin-5. The physiological functions of fibulin-5 remain largely elusive. *Fibln5*$^{-/-}$ mice exhibit widespread abnormalities in their elastic fibre system. Although these mice survive till adulthood, they experience severe emphysema, tortuous aorta and *cutis laxa* (loose skin) [230]. Moreover, mutations in *FBLN5* have been implicated in human *cutis laxa* and age-related molecular degeneration (AMD) [231, 232]. Fibulin-5 is expressed throughout the murine and human tissues and is dramatically upregulated in the vascular endothelial and smooth muscle cells during tissue repair [233]. FBLN5 is positively regulated by TGF-β [233, 234] and negatively regulated by VEGF [234], Myc [235], and IL-1β [236]. Just as was the case with TMEFF2, both oncogenic and onco-suppressive roles of FBLN5 have been reported, but the phenomenon is somewhat clearer in the case of FBLN5 – the oncogenic *vs* onco-suppressive roles of FBLN5 seem to be context specific and appear to depend on the type of cancer. FBLN5 is downregulated in numerous malignancies, especially in the cancers of kidney [233], breast [233], ovary [233, 237], bladder [238] , liver [239, 240], lung [241] and colon [233]. The opposite has been reported in the case of gastric [242] and nasopharyngeal [243] cancers.

FBLN5 enhances proliferation, DNA synthesis and motility in fibroblasts but acts oppositely in epithelial and endothelial cells [233, 234]. Could such context-specific qualities of FBLN5 be attributed to its regulation by TGF-β? Scheimann *et al.*, [233] demonstrated that FBLN5 enhances the proliferation and DNA synthesis in 3T3-L1 cells independent of TGF-β stimulation. This effect was observed both when FBLN5 was overexpressed and when cells were stimulated with the conditioned media containing secreted FBLN5. In addition, the study demonstrated that FBLN5 activates p38 MAPK and ERK 1/2 in 3T3-L1 cells independently without TGF-β stimulation. When similar experiments were conducted in mink lung Mv1Lu epithelial cells, treatment with FBLN5 conditioned-media and overexpression of FBLN5 resulted in a dramatic reduction in cell proliferation, DNA

synthesis and transcription of cyclin A. Again, these effects were independent of TGF-β stimulation. On the other hand, the study also reported that the FBLN5-stimulated MAPK-dependent AP-1 activity in Mv1Lu cells was only observed upon simultaneous treatment of the cells with TGF-β. The same research group also demonstrated later [244] that FBLN5 indirectly enhances apoptosis in 3T3-L1 cells through upregulating another protein, fibromodulin. Fibromodulin in turn stabilizes IκBα and inhibits NF-κB. When it comes to endothelial cells, Albig and Schiemann [234] demonstrated that the overexpression of FBLN5, or the stimulation through recombinant FBLN5, decreased the cell proliferation, invasion and the angiogenic sprouting of endothelial cells. In addition, they also reported that FBLN5 antagonizes VEGF signalling in endothelial cells and elevates the antiangiogenic factor, thrombospondin-1.

Tang *et al.* [239] demonstrated that the addition of recombinant FBLN5 to the cell culture of hepatocellular carcinoma cells (HCC) resulted in a marked inhibition of cell migration and invasion. Also, this effect was attributed to the FBLN5-dependent downregulation of matrix metalloprotease-7 (MMP-7). Mohamedi *et al.* [245] reported that fibulin-5 overexpression results in downregulation of Ki-67 in MCF-7, T47D and MDA-MB-231 breast cancer cells. In this study, FBLN5 overexpression also inhibited proliferation and the invasion capacity of these breast cancer cells. The latter was attributed to FBLN5-dependent reduction in the nuclear translocation of β-catenin. Similarly, Chen *et al.* [241] observed that fibulin-5 inhibited β-catenin signalling both *in vivo* and *in vitro* models of metastatic lung cancer, leading to downregulation of MMP-7 and cMYC. Hoe *et al.* [237] showed a reduction in the invasion and migration capacity of SKOV3 ovarian cancer cells upon FBLN5 overexpression. In addition, the study also reported that FBLN5 overexpression results in G2/M arrest and downregulation of matrix-metalloproteases 3, 9 and 13 in SKOV3 cells. Likewise, Møller *et al.* [246] reported that fibulin-5 prevents liver and lung metastasis of fibroblasts through inhibition of metalloprotease 9.

Another body of evidence regarding the functions of FBLN5 suggests a link between FBLN5 and oxygen metabolism. As I mentioned earlier, the *Fibln5*$^{-/-}$ mice exhibit extensive abnormalities in their elastic fibres. However, when the integrin binding RGD motif of Fbln5 is mutated in mice, the animals remain normal but exhibit increased levels of ROS in tissues with abundant fibronectin [247]. Wang *et al.* [248] used a mouse model of pancreatic ductal adenocarcinoma (PDAC) and crossed them with the ones having a mutated Fbln5 RGD domain. The loss of Fbln5-integrin interaction increased ROS production, subsequently leading to a decrease in tumour size and an increase in survivability, in

other words, the interaction of Fbln5 with integrins reduces the integrin-induced ROS production. The same research group also reported in a later study [249] that hypoxia triggers the upregulation of Fbln5 in PDAC both *in vivo* and *in vitro*. Also, this upregulation was alleviated upon inhibition of TGF-β receptors, PI3K or AKT. Similarly, Gaudall *et al.* [250] had demonstrated that FBLN5 mRNA and protein are upregulated in the bovine aortic endothelial cells in response to hypoxia and this upregulation was impeded upon the inhibition of the PI3K/AKT/mTOR signalling. Interestingly, the study also reported that the hypoxia-induced upregulation of FBLN5 ensues through the action of hypoxia-induced transcription factor HIF-1α. The investigators suggested that *FBLN5* promoter harbours hypoxia-response elements (HRE) and demonstrated that under hypoxic conditions the binding of HIF-1α to the putative HRE of FBLN5 was enhanced. Taken together, these reports are suggestive of FBLN5 being a stress-related adaptive protein.

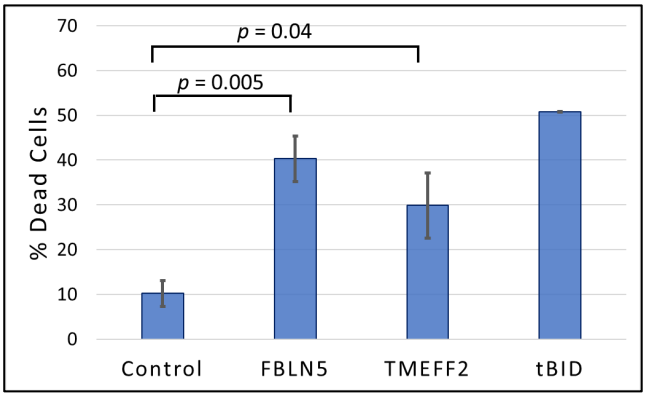
In short, FBLN5 is an ECM protein which is (i) positively regulated by TGF-β and HIF-1α, (ii) involved in the development and stabilization of ECM fibrils, especially elastin, (iii) downregulated in numerous malignancies and upregulated in a few, (iii) involved in integrin signalling, (iv) downregulated by Myc, VEGF and IL-1β, (v) it reduces the ECM driven ROS production and (vi) has antiangiogenic properties. Lastly, the pro- *vs* antiproliferative roles of FBLN5 are not fully elucidated.

6.2 Results

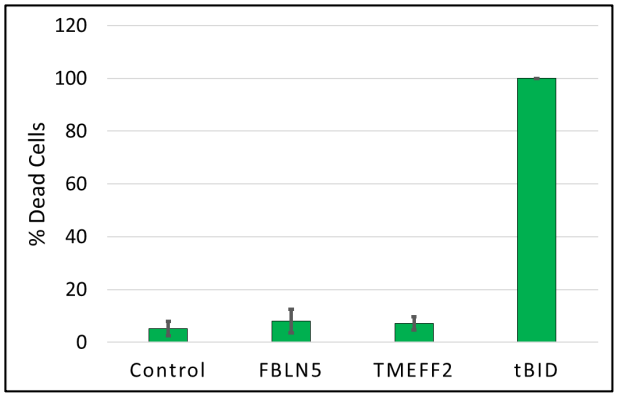
6.2.1 FBLN5 and TMEFF2 caused cell death in COS-7 but not in CV-1 cells

Following the decision to use COS-7 cells as the model for the transformed cells, I aimed to test if TEMFF2 and FBLN5 could induce cell death in these cells. The primary challenge in this regard was assaying the cell death in COS-7 cells. Ideally, a relatively straightforward and less time-consuming assay was required. The reported cell death assays for COS-7 cells included testing the DNA fragmentation through electrophoresis, trypan blue staining and manually counting the dead cells [251], co-transfection with luciferase reporter plasmid followed by luciferase activity assay, Hoechst dye nuclear staining and manually counting dead cells [252, 253], immunoblotting of apoptotic factors [254], TUNEL staining and manually counting the dead cells [255]. In the few reports where flow cytometry was used to assay cell death in COS-7 cells, the cells were treated with low molecular weight compounds rather than transfected with nucleic acid [256]. I tried several methods to assay cell death in COS-7 cells using flow cytometry without success; these included $Dioc_6$/PI staining, TMRE

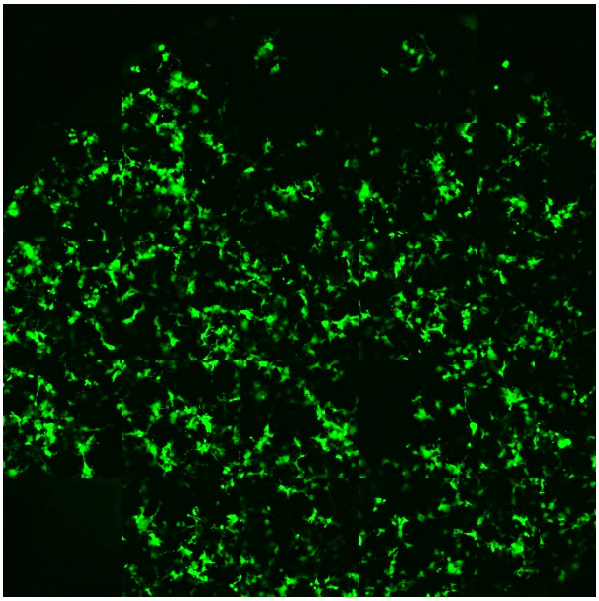
/DRAQ7™ staining and Annexin V-FITC/7AAD staining. I also tried MTT assay to assess cell viability but no reduction in cell viability was observed upon transfection of COS-7 cells with the positive control (tBID). Potentially, the problem was an exceptionally high division rate of COS-7 cells. Although transfection efficiency was optimum, the surviving untransfected cells divided at a very high rate, making the measurement of the signal from the remaining dead cells very difficult. Subsequently, I co-transfected FBLN5 and TMEFF2 (separately) along with a GFP reporter plasmid in the ratio of 1:4 (1 *GFP*: test plasmid) and manually counted the dead cells based on cell morphology. Approximately 3000 COS-7 cells and 2500 CV-1 cells were counted manually among three independently transfected populations of cells. Results are shown in figure 6.1. As expected, no obvious cell death was observed in CV-1 cells upon the overexpression of TMEFF2 or FBLN5. However, both TMEFF2 and FBLN5 caused significant cell death in COS-7 cells.



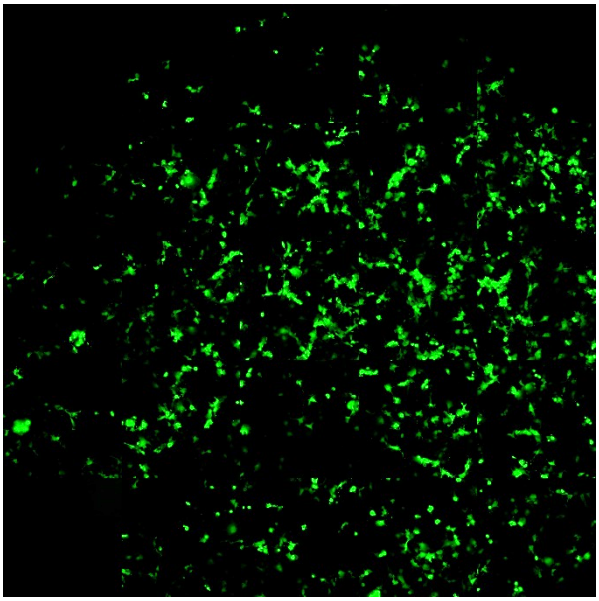
(a.) COS-7 Cells



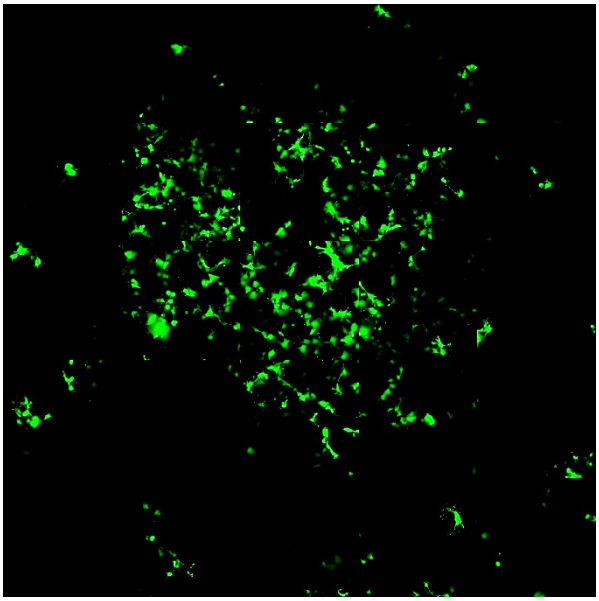
(b.) CV-1 Cells



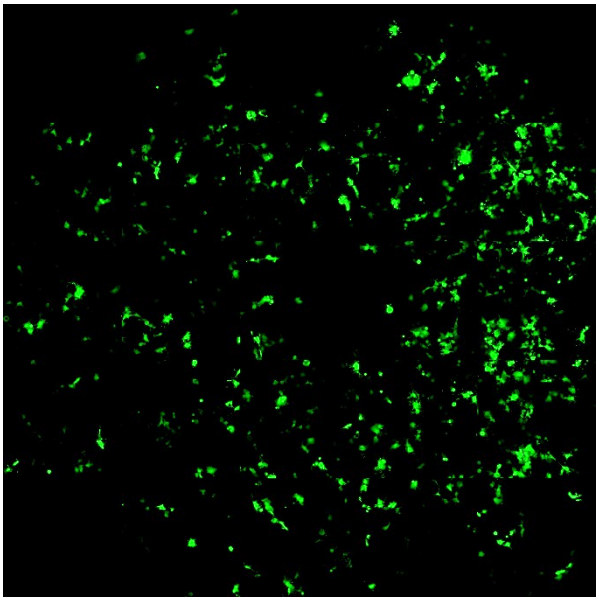
(c.i) COS-7 Control



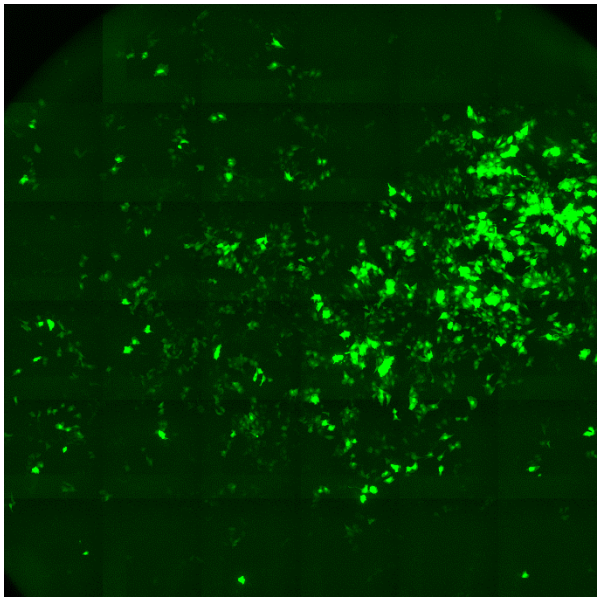
(c.ii) COS-7 FBLN5



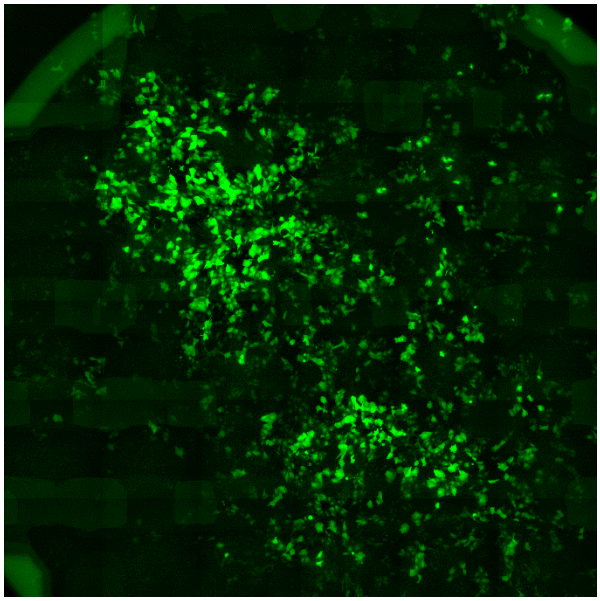
(c.iii) COS-7 TMEFF2



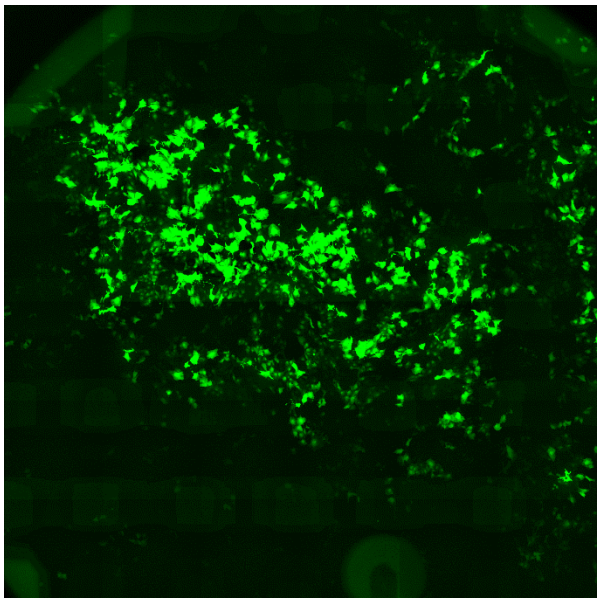
(c.iv) COS-7 tBID



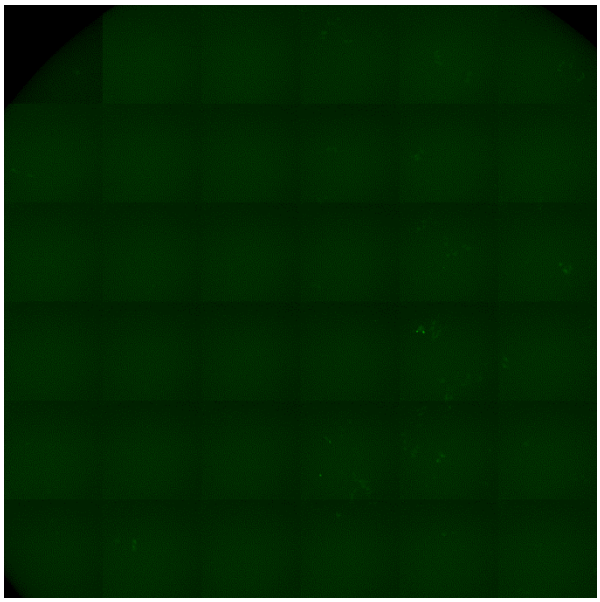
(d.i) CV-1 Control



(d.iii) CV-1 FBLN5



(d.ii) CV-1 TMEFF2



(d.iv) CV-1 tBID

***Figure 6.1*: Difference of cell death induced in CV-1 and COS-7 by overexpression of FBLN5 and TMEFF2.** CV-1 and COS-7 cell were co-transfected with GFP and either FBLN5, TMEFF2, Luciferase (negative 'control') or tBID (positive control) in 1:4 ratio (1 GFP: 4 test plasmid) in 96-well plates. 48 hours post-transfection, green fluorescent cells were imaged using the IN Cell Analyzer 2000 (GE Healthcare) with 20x objective. 36 fields in each well of a 96-well plate were imaged. ~ 3000 COS-7 cells and ~ 2500 CV-1 cells were counted manually for each condition. Three independently transfected populations from three different wells of a 96-well plate were counted for each condition. Rounded, clustered and/or cells with disfigured morphology were considered as dead. (a) The percentage of dead green fluorescent COS-7 cells co-transfected with GFP and either FBLN5, TMEFF2, Luciferase (negative 'control') or tBID (positive control) in 1:4 ratio. Bars represent

averages from 3 wells of a 96-well plate each with an independently transfected population of cells. p values were calculated using two-sample t-test assuming unequal variances. (b) The percentage of dead green fluorescent CV-1 cells co-transfected with GFP and either FBLN5, TMEFF2, Luciferase (negative 'control') or tBID (positive control) in 1:4 ratio. Bars represent averages from 3 wells of a 96-well plate each with an independently transfected population of cells. Error bars represent +/- standard deviation. NB: no green cells could be detected in either wells transfected with tBID and 100% was added manually in the chart. (c.i to c.iv) Representative 36 fields montages from one out of three wells imaged for calculating data in figure a: (c.i) COS-7 cells co-transfected with GFP and Luciferase (negative 'control') in 1:4 ratio. (c.ii) COS-7 cells co-transfected with GFP and FBLN5 in 1:4 ratio. (c.iii) COS-7 cells co-transfected with GFP and TMEFF2 in 1:4 ratio. (c.iv) COS-7 cells co-transfected with GFP and tBID (positive control) in 1:4 ratio. (d.i to d.iv) Representative 36 fields montages from one out of 3 wells imaged for calculating data in figure b: (d.i) CV-1 cells co-transfected with GFP and Luciferase (negative 'control') in 1:4 ratio. (d.ii) CV-1 cells co-transfected with GFP and FBLN5 in 1:4 ratio. (d.iii) CV-1 cells co-transfected with GFP and TMEFF2 in 1:4 ratio. (d.iv) CV-1 cells co-transfected with GFP and tBID (positive control) in 1:4 ratio. For both (c) and (d), image processing was performed using Fiji – 'batteries-included' image processing package.

Manually counting cells could be prone to error and declaring a cell as dead on the basis of morphology is subjective. However, during this experiment I noticed that when transfected with FBLN5, TMEFF2 and tBID, the total number of green fluorescent COS-7 cells was less when compared to the number of green fluorescent cells in negative control. I therefore counted the total number of green cells at various time points after transfection. Similar to the previous experiment, both the CV-1 and COS-7 cells were co-transfected with GFP and either the control or test plasmids in 1:4 ratio. Instead of imaging the live cells, the cells were fixed at 24 and 48 hours post-transfection with 4% paraformaldehyde (w/v), washed with PBS and imaged for green fluorescence. Instead of counting manually, Fiji – 'batteries-included' image processing package [83] was used to count the total number of cells. The results are presented in figure 6.2. Both FBLN5 and TMEFF2 overexpression resulted in a marked decrease in the number of green fluorescent COS-7 cells but not in CV-1 cells. Moreover, no green fluorescent cells were detected when GFP was co-transfected with tBID in CV-1 cells, i.e., tBID overexpression resulted in cell death of all the transfected cells.

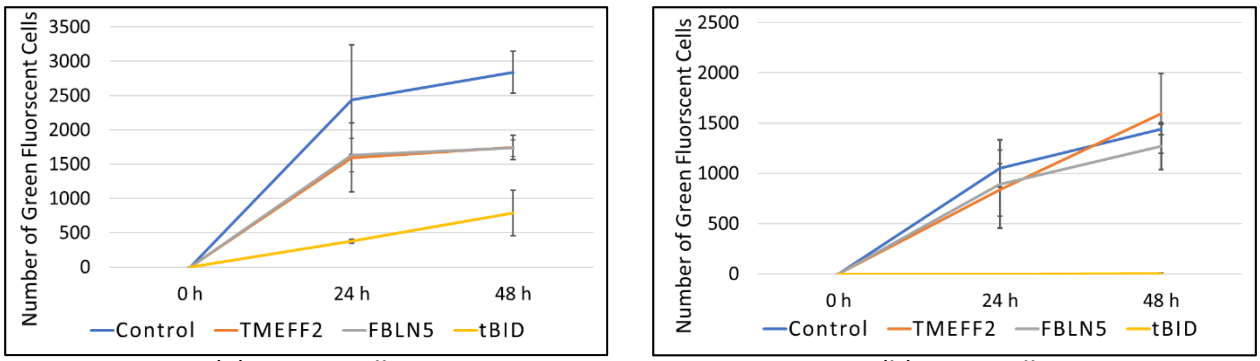


***Figure 6.2*: Effects of FBLN5 and TMEFF2 overexpression on cell viability in COS-7 and CV1 cells.** CV-1 and COS-7 cell were co-transfected with GFP and either FBLN5, TMEFF2, Luciferase (negative 'control') or tBID (positive control) in 1:4 ratio (1 GFP: 4 test plasmid) in 96-well plates. At 24 and 48 hours post-transfection, media was removed and cells were fixed with 4 % paraformaldehyde. After washing twice with PBS, green fluorescent cells were imaged using the IN Cell Analyzer 2000 (GE Healthcare) with 10x objective. 9 fields in each well of a 96-well plate were imaged. Green fluorescent cells were counted using the Fiji – 'batteries-included' image processing package. (a.) The number of total green fluorescent COS-7 cells co-transfected with GFP and either FBLN5, TMEFF2, Luciferase (negative 'control') or tBID (positive control) in 1:4 ratio. Each data point represents averages from 3 wells of a 96-well plate each with an independently transfected population of cells. (b) The number of total green fluorescent CV-1 cells co-transfected with GFP and either FBLN5, TMEFF2, Luciferase (negative 'control') or tBID (positive control) in 1:4 ratio. Each data point averages from 3 wells of a 96-well plate each with independently transfected population of cells. Error bars represent +/- standard deviation.

In order to further validate the phenomenon of selective cell death induction by TMEFF2 and FBLN5 in COS-7 cells, I employed another assay where COS-7 and CV1 cells were transfected with TMEFF2 and FBLN5 along with the controls and antibiotic (G418) was added onto the media until the colonies of untreated cells diminished. Results are shown in figure 6.3. Again, overexpression of FBLN5 or TMEFF2 decreased the survival of COS-7 cells but not of CV-1 cells and the overexpression of tBID wiped out the CV-1 cells. Moreover, the overexpression of FBLN5 decreased the COS-7 colonies dramatically. Together these results indicated a clear anticancer gene activity of FBLN5 and TMEFF2 where upon their overexpression, the transformed cells underwent cell death and/or reduction in cell viability whereas the normal cells remained largely unharmed. Conversely, as FBLN5 and TMEFF2 were in pME18SFL3 vector which does not have an antibiotic resistance cassette, for this experiment, I cloned FBLN5 and TMEFF2 into pcDNA 3.1. (details in materials and methods section).

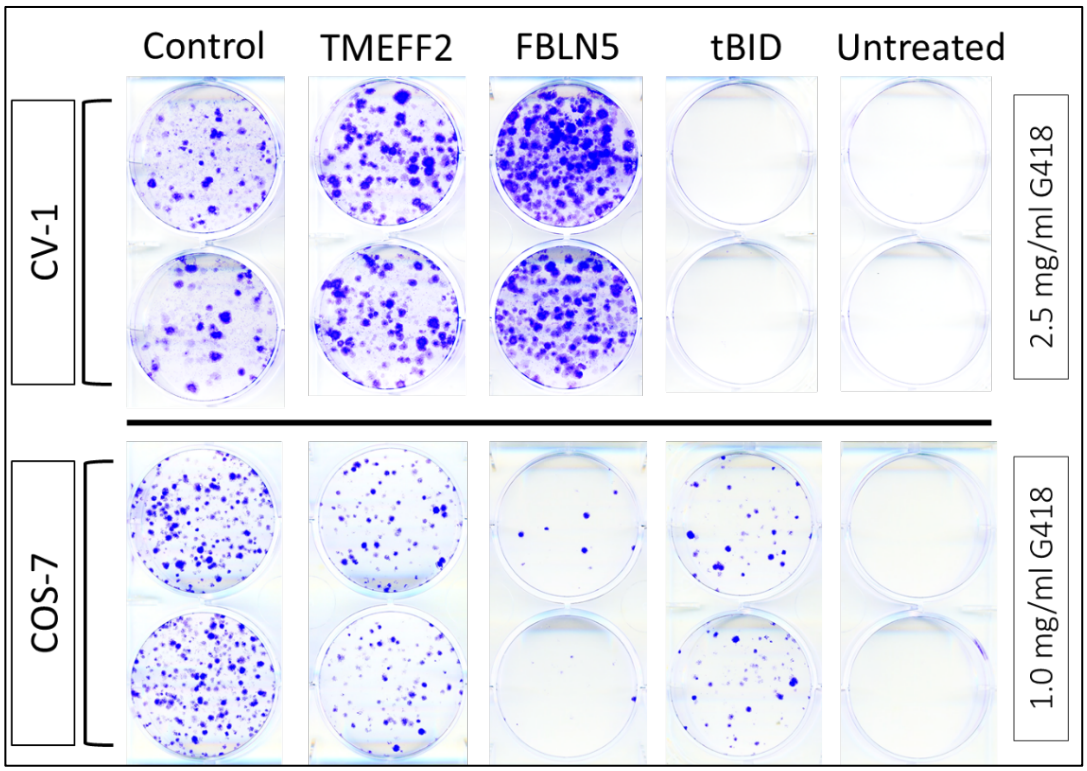


***Figure 6.3*: Modified clonogenicity assay for the effects of FBLN5 and TMEFF2 overexpression on long-term cell viability in COS-7 and CV1 cells.** For CV-1 cells: 70,000 CV-1 cells/well were seeded in 6-well plates and transfected with pcDNA 3.1 empty vector (negative 'Control'), TMEFF2 (in pcDNA3.1), FBLN5 (in pcDNA3.1) and tBID (in pcDNA 3.1, positive control). 48 hours post-transfection, media was replaced with fresh one containing 2.5 mg/ml G418 and changed every third day till there were no cells left in the untransfected ('Untreated') population (~ 2.5 weeks). Cells were fixed with 4 % PFA, stained with crystal violet (see materials and methods) and imaged using a conventional table-top scanner. Each well represents an independently transfected and processed population of cells. For COS-7 cells: 20,000 cells/well were seeded in 24-well plates and transfected with pcDNA 3.1 empty vector (negative 'Control'), TMEFF2 (in pcDNA3.1), FBLN5 (in pcDNA3.1) and tBID (in pcDNA 3.1 positive control). 24 hours post-transfection, cells were trypsinized and resuspended in 1 ml DMEM + 10% FBS (final volume). Out of 1 ml cell suspension, 10 µl were re-seeded in the corresponding wells of 6-well plates. 24 hours after re-seeding, media was replaced with fresh one containing 1.0 mg/ml G418 and changed every third day till there were no cells left in the untransfected ('Untreated') population (~ 10 days). Cells were fixed with 4 % paraformaldehyde, stained with crystal violet (see materials and methods) and imaged using a conventional table-top scanner. Each well represents an independently transfected and processed population of cells.

6.2.2 Comparative levels of TMEFF2 and FBLN5 overexpression between CV-1 and COS-7 cells

As I have mentioned during the introduction, the degree of overexpression of the putative anticancer gene could be the deciding factor in selective cell death of the transformed or cancer cells. As the transformed or cancer cells proliferate at a much higher rate and are relatively easier to transfect, the transfected genes could be overexpressed to a higher degree compared to the normal cells. I therefore conducted a time course experiment where following the transfection with FBLN5 or TMEFF2, CV-1 and COS-7 cells were harvested at certain time points and subjected to immunoblotting and qRT-PCR to check the protein and mRNA levels of the overexpressed genes respectively. The results are shown in figure 6.4 for TMEFF2 and figure 6.5 for FBLN5. The degree of TMEFF2 overexpression was vastly different between COS-7 and CV-1 cells where both protein and mRNA levels of TMEFF2 increased marginally in CV-1 cells. Moreover, the overexpression did not increase over time. As for FBLN5, again there was a difference in the degree of overexpression which was not surprising. However, this difference was much less than that observed in the case of TMEFF2. CV-1 cells not only overexpressed FBLN5 upon its transfection successfully, but the overexpression also increased proportionally over time. NB: these experiments were running in parallel with those presented in chapter 5 and it was at this point that I had decided to investigate the effects of TEMFF2 in MCF-7 cells and those of FBLN5 in CV-1 and COS-7 cells.

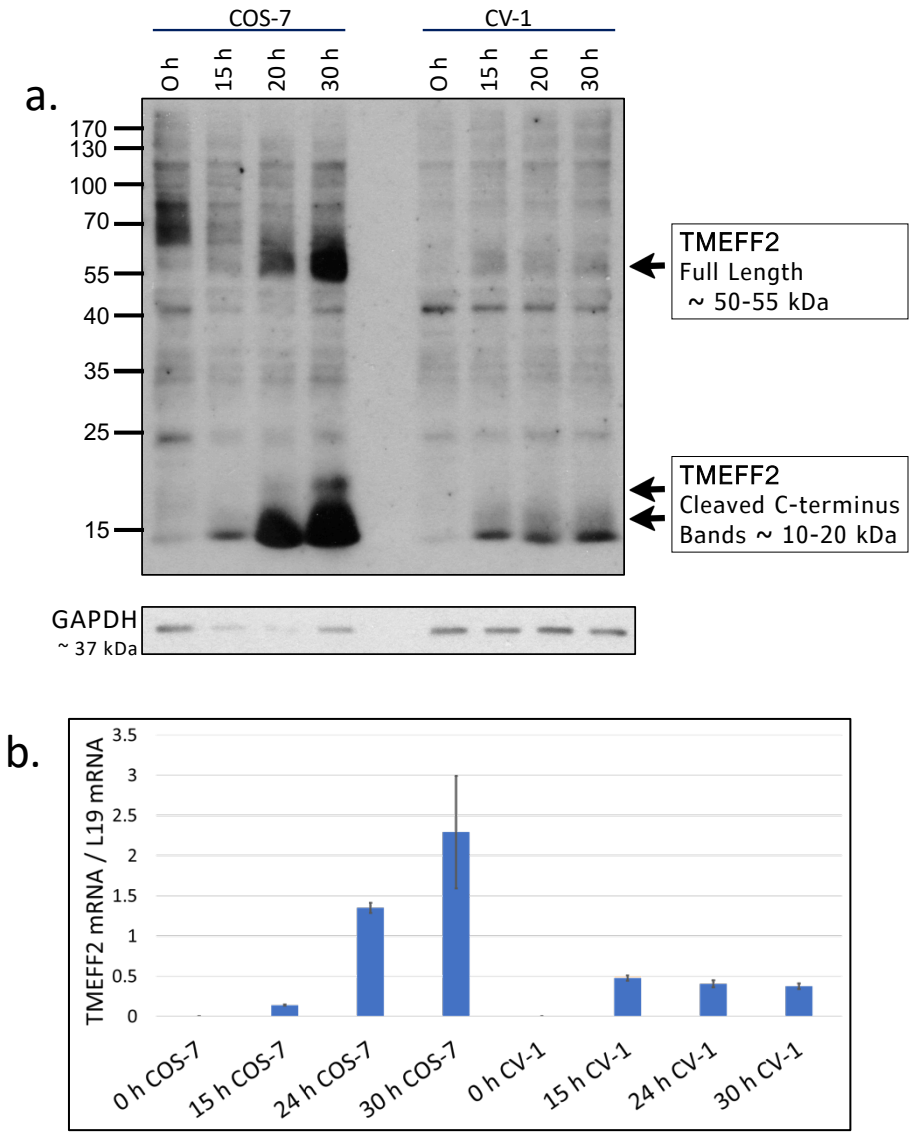


***Figure 6.4*: Comparison of TMEFF2 overexpression between COS-7 and CV-1 cells. (a.)** Immunoblot showing time course of overexpressed TMEFF2 in CV-1 and COS-7 cells. COS-7 and CV-1 cells were transfected with TMEFF2 in 10 cm dishes. Cells were harvested 15, 20 and 30 hours post-transfection, protein extracted and subjected to immunoblotting for TMEFF2. Wild type COS-7 and CV-1 cells were used for 0 hour time point. Proteins were separated using 10% SDS-PAGE under reducing conditions and electroblotted onto nitrocellulose membrane. Immunoblotting procedure is described in detail in the materials and methods section. **(b.)** qRT-PCR representing time course of overexpressed TMEFF2 mRNA in CV1 and COS-7 cells. L19 was used as housekeeping gene for calculating the relative TMEFF2 mRNA expression levels using the 'relative standard curve' method. Cells were transfected in 24-well plates. RNA for each sample represented was extracted after pooling cells form 12 wells of a 24-well plate, each with an independently transfected population of cells. Wild type COS-7 and CV-1 cells were used for 0 hour time point. Bars represent the average from 3 PCR-reactions. Error bars represent +/- standard deviation.

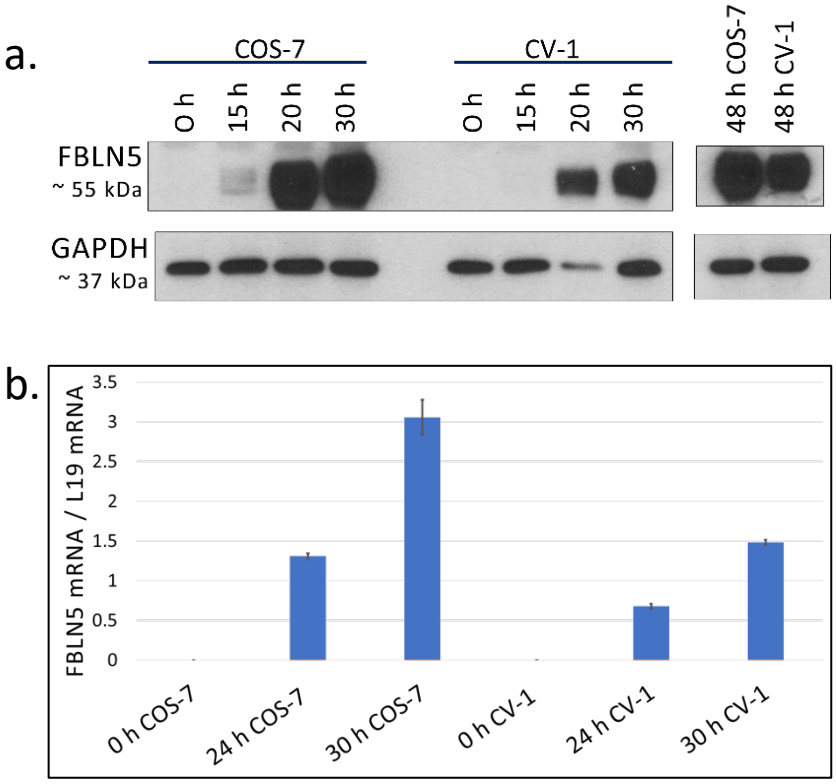
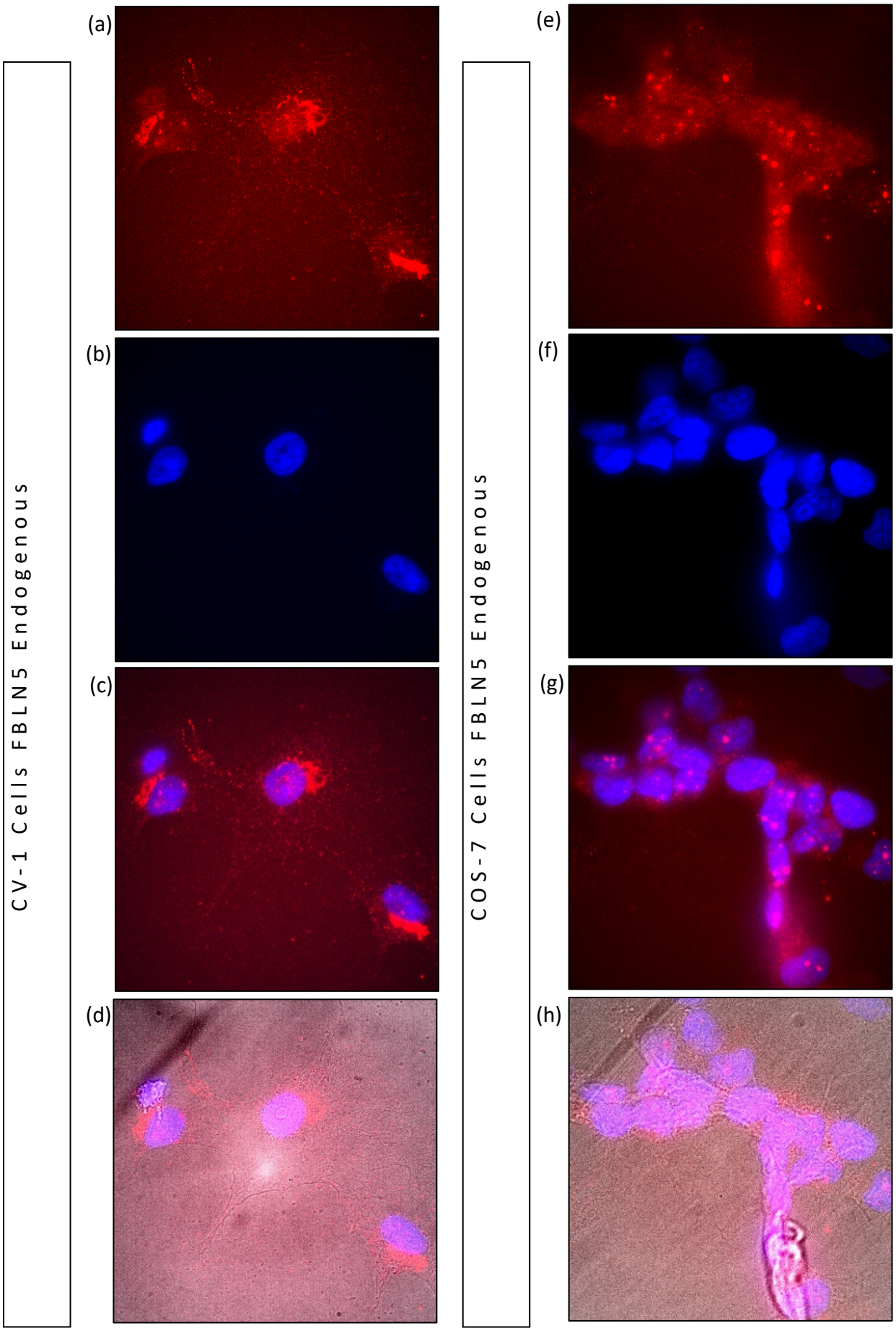


***Figure 6.5*: Comparison of FBLN5 overexpression between COS-7 and CV-1 cells. (a.)** Immunoblot showing time course of overexpressed FBLN5 in CV-1 and COS-7 cells. COS-7 and CV-1 cells were transfected with FBLN5 in 10 cm dishes. Cells were harvested 15, 20, 30 and 48 hours post-transfection, protein was extracted and subjected to immunoblotting for FBLN5. Wild type COS-7 and CV-1 cells were used for 0 hour time point. Proteins were separated using 10% SDS-PAGE under reducing conditions and blotted onto nitrocellulose membrane. Immunoblotting procedure is described in detail in materials and methods section. **(b.)** qRT-PCR representing time course of overexpressed FBLN5 mRNA in CV1 and COS-7 cells. L19 was used as housekeeping gene for calculating the relative FBLN5 mRNA expression levels using the 'relative standard curve' method. Cells were transfected in 24-well plates. RNA for each sample represented was extracted after pooling cells form 12 wells of a 24-well plate, each with an independently transfected population of cells. Wild type COS-7 and CV-1 cells were used for 0 hour time point. Bars represent the average from 3 PCR-reactions. Error bars represent +/- standard deviation.

6.2.3 Intracellular localization of FBLN5 in CV-1 and COS-7 cells

One of the factors in imparting cancer-specific cell death by several anticancer genes is the different intracellular localization between normal and transformed cells. I therefore assayed for the intracellular localization of FBLN5 in CV-1 and COS-7 cells using immunofluorescence. Results are

depicted in figure 6.6. The endogenous FBLN5 gave a perinuclear crown-like appearance in CV-1 cells whereas in the COS-7 cells it seems to largely localize in the nucleus and gave a punctate nuclear appearance. It is important to mention here that the endogenous levels of FBLN5 in COS-7 cells are extremely low, thus making immunofluorescence error prone due to the possible non-specific binding. When overexpressed, FBLN5 formed peculiar yet different cytoplasmic patterns in CV-1 and COS-7 cells. The cytoplasmic patterns of overexpressed FBLN5 in COS-7 cells appeared more thread-like and dispersed whereas in the CV-1 cells it appeared to form a three-dimensional meshwork throughout the cytoplasm.


CV-1 Cells FBLN5 Endogenous
COS-7 Cells FBLN5 Endogenous
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

CV-1 Cells FBLN5 Overexpressed
(i)
(j)
(k)
(l)
COS-7 Cells FBLN5 Overexpressed
(m)
(n)
(o)
(p)

***Figure 6.6*: Intracellular localization of FBLN5 in CV-1 and COS-7 cells**. CV-1 and COS-7 cells were transiently transfected with luciferase (for endogenous) and FBLN5 in ibidi® 8-well glass bottom chamber coated with poly L-lysine. COS-7 cells were fixed with 4.0% paraformaldehyde 30 hours post-transfection and CV-1 cells were fixed 48 hours post-transfection. Immunofluorescent staining was performed as described in materials and methods. Image processing was performed using Fiji – 'batteries-included' image processing package. (a to d) CV-1 cells transfected with Luciferase: (a) CV-1 cells; FBLN5 [Alexa Fluor® 647-Red], (b) CV-1 cells; DAPI, (c) CV-1 cells; FBLN5 and DAPI overlay, (d) CV-1 cells; FBLN5 and DAPI overlay with phase contrast. (e to h) COS-7 cells transfected with Luciferase: (e) COS-7 cells; FBLN5 [Alexa Fluor® 647-Red], (f) COS-7 cells; DAPI, (g) COS-7 cells; FBLN5 and DAPI overlay, (h) COS-7 cells; FBLN5 and DAPI overlay with phase contrast. (i to l) CV-1 cells transfected with FBLN5: (i) CV-1 cells; FBLN5 [Alexa Fluor® 647-Red], (j) CV-1 cells; DAPI, (k) CV-1 cells; FBLN5 and DAPI overlay, (l) CV-1 cells; FBLN5 and DAPI overlay with phase contrast. (m to p) COS-7 cells transfected with FBLN5: (m) COS-7 cells; FBLN5 [Alexa Fluor® 647-Red], (n) COS-7 cells; DAPI, (o) COS-7 cells; FBLN5 and DAPI overlay, (p) COS-7 cells; FBLN5 and DAPI overlay with phase contrast. NB: microscopy was performed by Dr Adam Cawte.

6.2.4 FBLN5 Comparative Transcriptomic Analyses

I employed the transcriptomics analysis to investigate the differential effects of FBLN5 in COS-7 and CV-1 cells. The schematic representation of the methodology employed in the transcriptomics study is elaborated in figure 6.7. Briefly, FBLN5 was transfected into both COS-7 and CV-1 cells and microarray analyses were performed where the transcriptomic changes between untransfected COS-7 cells and those transfected with FBLN5 were assayed. The same was conducted in wild type CV-1 cells and the CV-1 cells transfected with FBLN5. Moreover, a third differentially-expressed set of genes was analysed – between wild type COS-7 cells and wild type CV-1 cells, i.e., differential gene expression analysis of COS-7 cells/CV-1 cells. These three sets of differentially-expressed genes were then compared, primarily using the Ingenuity Pathway Analysis (IPA) (Qiagen®) [86, 207]. In comparison to the relatively simple transcriptomic analysis of TMEFF2 in MCF-7 cells, the difference in expression levels of the overexpressed FBLN5 between the COS-7 and CV-1 cells was an important factor when choosing the time point for harvesting the cells. Taking together the results presented in figure 6.2 and 6.5, I decided to harvest CV-1 cells 30 hours post-transfection and COS-7 cells 25 hours post-transfection. At these time points, the comparative expression levels of FBLN5 appeared relatively similar and, according to the results in figure 6.2, the cell death and/or reduction in viability induced by FBLN5 appeared to be initiating. For transfection and RNA extraction, I employed the same scheme that I had employed in the case of TMEFF2 (figure 5.11). FBLN5 transfection in both CV-1 and

COS-7 cells was performed in 24-well plates and cell lysates from 12 wells were pooled together to extract RNA for one sample. The whole process was repeated in triplicate where RNA for each test sample was extracted by pooling 12 independently transfected populations of cells (both CV-1 and COS-7). In order to keep the RNA extraction for wild type cells consistent with the transfected ones, the RNA samples for the wild type CV-1 and COS-7 cells were also extracted by pooling lysates from 12 wells of a 24-well plate. Similar to MCF-7 cells, Affymetrix® GeneChip™ Clariom™ D human arrays were employed for this experiment.

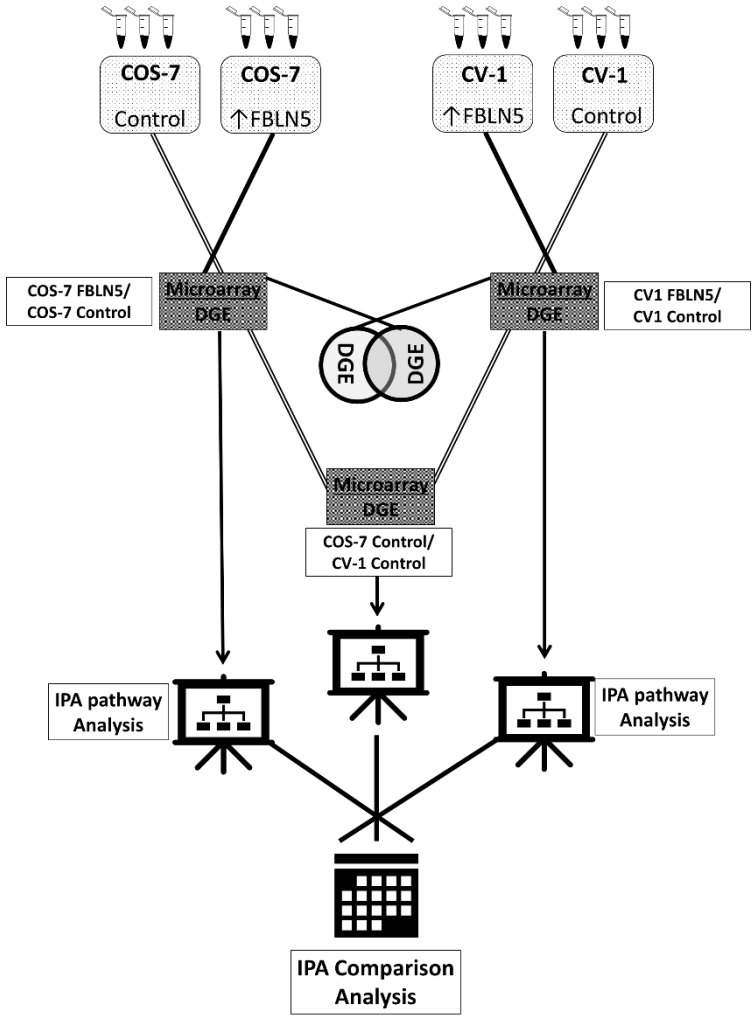


***Figure 6.7*: Schematic representation of the comparative transcriptomic analysis in CV-1 and COS-7 cells upon FBLN5 overexpression.** DGE: Differential Gene Expression. IPA: Ingenuity Pathway Analysis (Qiagen®). Wild type and FBLN5 overexpressing CV-1 and COS-7 cells were subjected to microarray hybridization and subsequent differential gene expression analysis (DGE). Three sets of differentially expressed genes were obtained: CV-1 transfected with FBLN5 / wild type CV-1, COS-7 transfected with FBLN5 / wild type COS-7 and wild type COS-7 / wild type CV-1. These three sets of differentially expressed genes were then subjected to the Ingenuity Pathway Analysis (IPA) Core Analysis individually. The three core analysis reports thus obtained were subjected to the IPA Comparison Analysis. As an additional step, the two differentially expressed sets of genes (CV-1 transfected with FBLN5 / wild type CV-1 and COS-7 transfected with FBLN5/ wild type COS-7) were also compared separately using Venn diagram generation tool 'Venny 2.1'.

6.2.4.1 Validation of overexpression

Before conducting the microarray hybridization and processing, I performed qRT-PCR on CV-1 and COS-7 cells transfected with FBLN5 in order to confirm the overexpression. The results are depicted in figure 6.8.

Similar overexpression levels for FBLN5 in COS-7 and CV-1 cells could not be achieved. I harvested COS-7 cells at 24 hours post-transfection and CV-1 cells at 30 hours post-transfection. In the previous experiments, these time points showed similar relative FBLN5 mRNA and protein expression levels. However, in this particular instance, the FBLN5 mRNA levels in COS-7 cells were approximately four times higher than those in CV-1 cells. Nevertheless, these samples were subjected to the microarray processing and the subsequent analyses.

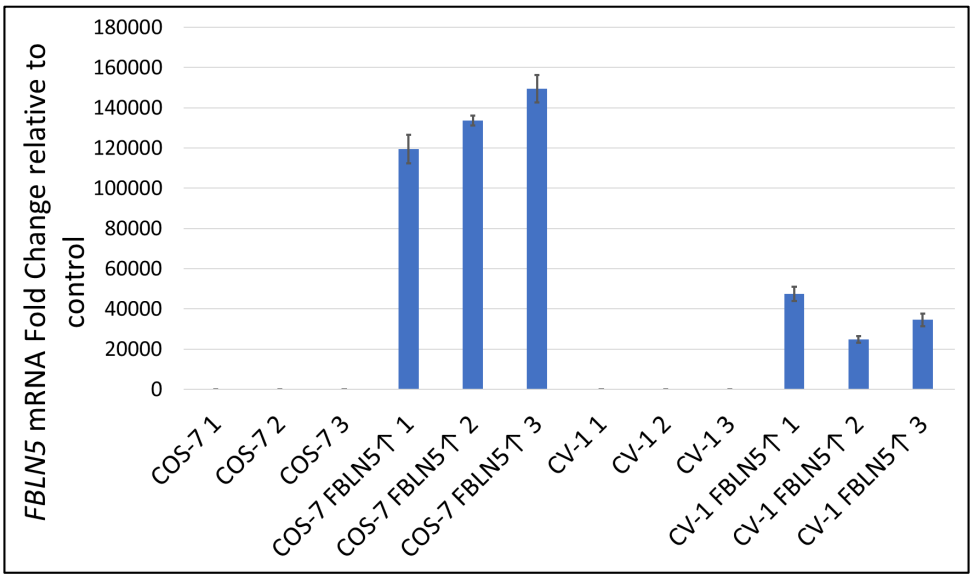


***Figure 6.8*: Confirmation of FBLN5 overexpression in triplicate samples for microarray transcriptomic analysis**. qRT-PCR results for the relative FBLN5 mRNA fold change in COS-7 and CV-1 cells transfected with FBLN5 compared to the wild type COS-7 and CV-1 cells (respectively) in triplicate. Fold change was quantified according to ΔΔCt method using L19 as the housekeeping gene and COS-7 1 as the calibrator 'control'. COS-7 cells were harvested 24 hours post-transfection while the CV-1 cells were harvested 30 hours post-transfection. RNA for each sample represented was extracted after pooling the cell lysates from 12 wells of a 24-well plate, each with an independently transfected population of cells. RNA for the wild type CV-1 and COS-7 cells was also extracted by pooling lysates from 12 wells of 24-well plates for each sample. Bars represent the average reading from 3 PCR-reactions. Error bars represent +/- standard deviation.

6.2.4.2 Differential Gene Expression Analysis of FBLN5 overexpressing COS-7 and CV-1 cells

Just as it was the case with TMEFF2 in MCF-7 cells, I employed Affymetrix® GeneChip™, Clariom™ D human exome microarrays for the transcriptomic analysis of CV-1 and COS-7 cells. Transcriptome Analysis Console (TAC) 4.0.1 (Applied Biosystems) and Ingenuity Pathway Analysis (IPA) (Qiagen®) [86, 207] were employed for the differential gene expression analysis and the pathways analysis respectively. The principle component analysis and quality control analyses are presented in supplementary figures S.9 and S.10. The results were filtered through a criteria of fold change between 2.0 and -2.0, and p value and false discovery rate (FDR) of 0.05. The results are summarized in figure 6.9 and table 6.1. Briefly, upon FBLN5 overexpression in CV-1 cells, 2858 transcripts were upregulated while 2327 transcripts were downregulated. Among the differentially expressed transcripts, 696 were for coding genes, 1696 non-coding transcripts, 268 were pseudogenes, 163 precursor microRNAs, 67 small_RNAs, 163 precursor_microRNAs, 383 novel unassigned transcripts and 1902 multiple complexes. On the contrary, when FBLN5 was overexpressed in COS-7 cells, 1972 transcripts were upregulated while 2156 transcripts downregulated. Among the differentially expressed transcripts, 519 were for coding genes, 1328 non-coding transcripts, 264 were pseudogenes, 128 precursor microRNAs, 69 small_RNAs, 163 precursor_microRNAs, 290 novel unassigned transcripts and 1522 multiple complexes. In the third set of differentially expressed genes (wild type COS-7 cell *vs* wild type CV-1 cells), 1617 transcripts were upregulated while 1292 transcripts were downregulated. Among the differentially expressed transcripts, 185 were for coding genes, 330 non-coding transcripts, 67 were pseudogenes, 15 precursor microRNAs, 17 small_RNAs, 69 novel unassigned transcripts and 2225 multiple complexes.

***Table 6.1*: Summary of differentially expressed transcripts in** (a) CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells, (b) COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells, (c) wild type COS-7 cells *vs* wild type CV-1 cells.

a.

Group	Total	Passed Filter	Up-Regulated	Down-Regulated
Multiple_Complex	29510	1902	853	1049
NonCoding	66845	1696	1129	567
Coding	18858	696	475	221
Unassigned	10001	386	256	130
Pseudogene	4340	268	43	225
Precursor_microRNA	3297	163	57	106
Small_RNA	2386	67	39	28
Ribosomal	507	7	6	1
tRNA	6	0	0	0

b.

Group	Total	Passed Filter	Up-Regulated	Down-Regulated
Multiple_Complex	29510	1522	417	1105
NonCoding	66845	1328	880	448
Coding	18858	519	345	174
Unassigned	10001	290	194	96
Pseudogene	4340	264	48	216
Precursor_microRNA	3297	128	44	84
Small_RNA	2386	69	36	33
Ribosomal	507	8	8	0
tRNA	6	0	0	0

c.

Group	Total	Passed Filter	Up-Regulated	Down-Regulated
Multiple_Complex	29510	2225	1312	913
NonCoding	66845	330	121	209
Coding	18858	185	91	94
Unassigned	10001	69	26	43
Pseudogene	4340	67	48	19
Small_RNA	2386	17	13	4
Precursor_microRNA	3297	15	6	9
Ribosomal	507	1	0	1
tRNA	6	0	0	0

Filter: -2 ≤ fold change ≥ 2, One-Way ANOVA with Empirical Bayes Statistics adjustment p value ≤ 0.05 and False Discovery Rate (FDR) calculated through Benjimini-Hochberg Step-Up FDR-controlling procedure ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome arrays. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

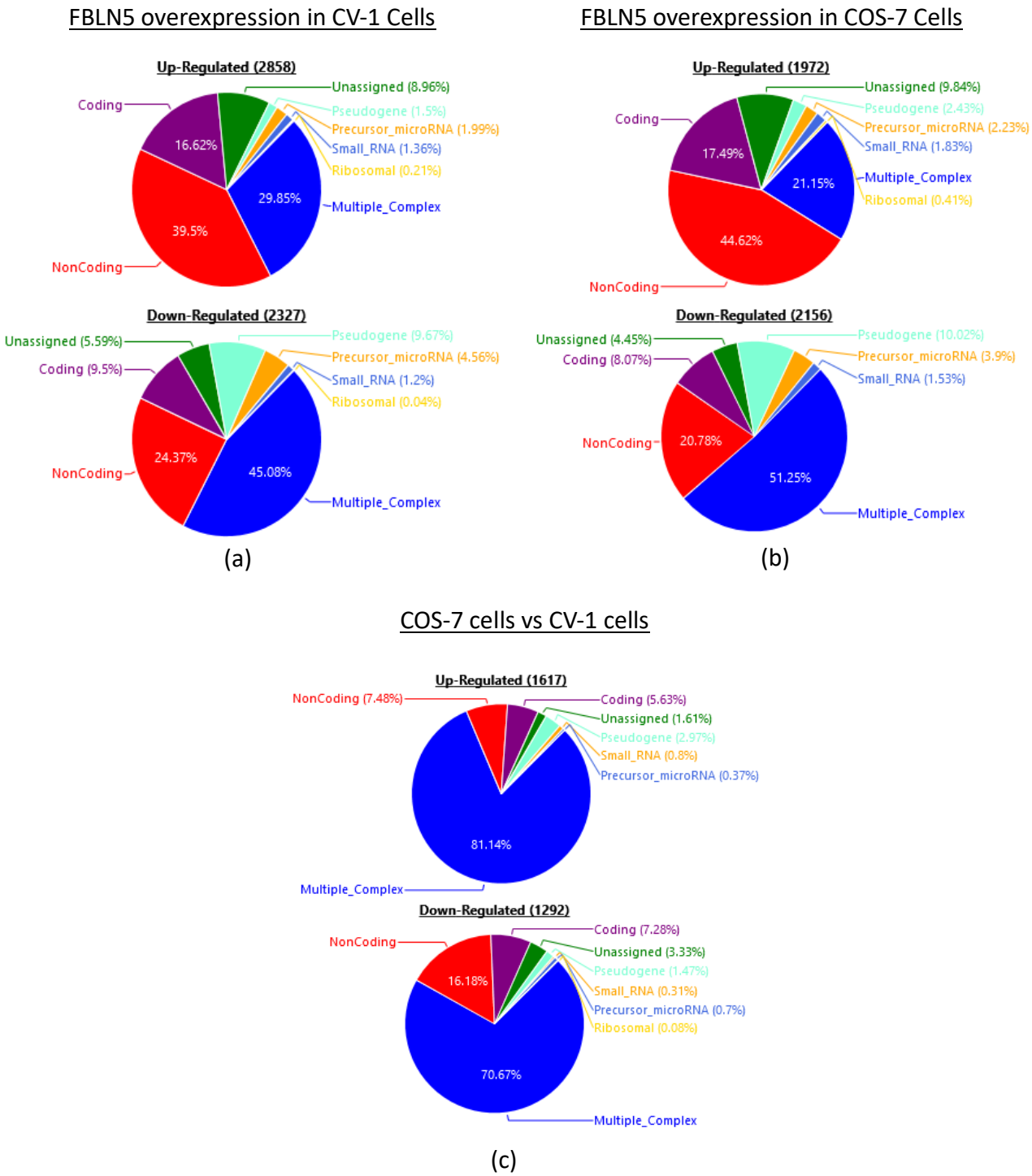


***Figure 6.9*: Summary of differentially expressed transcripts**. (a) CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells, (b) COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells, (c) wild type COS-7 cells *vs* wild type CV-1 cells. Filter: -2 ≤ fold change ≥ 2, One-Way ANOVA with Empirical Bayes Statistics adjustment p value ≤ 0.05 and False Discovery Rate (FDR) calculated through Benjimini-Hochberg Step-Up FDR-controlling procedure ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome arrays. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

Interestingly, despite an approximately 4-fold lower expression level, the FBLN5 overexpression in CV-1 cells resulted in a higher number of differentially expressed genes when compared to those in FBLN5 overexpressing COS-7 cells. Moreover, when the number of transcripts upregulated were compared with those that are downregulated, upon FBLN5 overexpression, the CV-1 cells exhibited a trend more toward upregulation, while the COS-7 cells showed a higher number of downregulated transcripts (figure 6.10). The set of differentially expressed genes between wild type COS-7 and wild type CV-1 cells was smaller than the differentially expressed genes upon FBLN5 overexpression in both CV-1 and COS-7 cells. In addition, the up- and downregulated genes appeared to have largely an even ratio. One distinguishing feature of this gene set (wild type COS-7/CV-1) was the scatter plot (figure 6.10 (f)). The genes upregulated upon SV40 transformation were spread evenly across the spectrum of genes with natively high and low expression levels in the two cell lines; the downregulated genes appeared to predominate the region with low native gene expressions. In the case of FBLN5 overexpressing CV-1 and COS-7, the scatter plots revealed upregulation mainly in the gene sets which had low native expression levels whereas the downregulation was spread evenly.

As it was hypothesized that FBLN5 would be imparting its differential effect in COS-7 cells primarily in the context of the biochemical differences between COS-7 and CV-1 cells, it was more important to understand the native transcriptomic differences between CV-1 and COS-7 cells (i.e., without the effects of the transfection process). Therefore, wild type CV-1 and COS-7 cells were used as controls for the differential gene expression analysis instead of the cells transfected with the control protein. Ideally, it would be preferable to use both conditions, i.e., wild type CV-1 and COS-7 cells, and both transfected with control protein. However, hypothetically, the effects of the transfection process would cancel out when the two differentially expressed gene sets were compared.

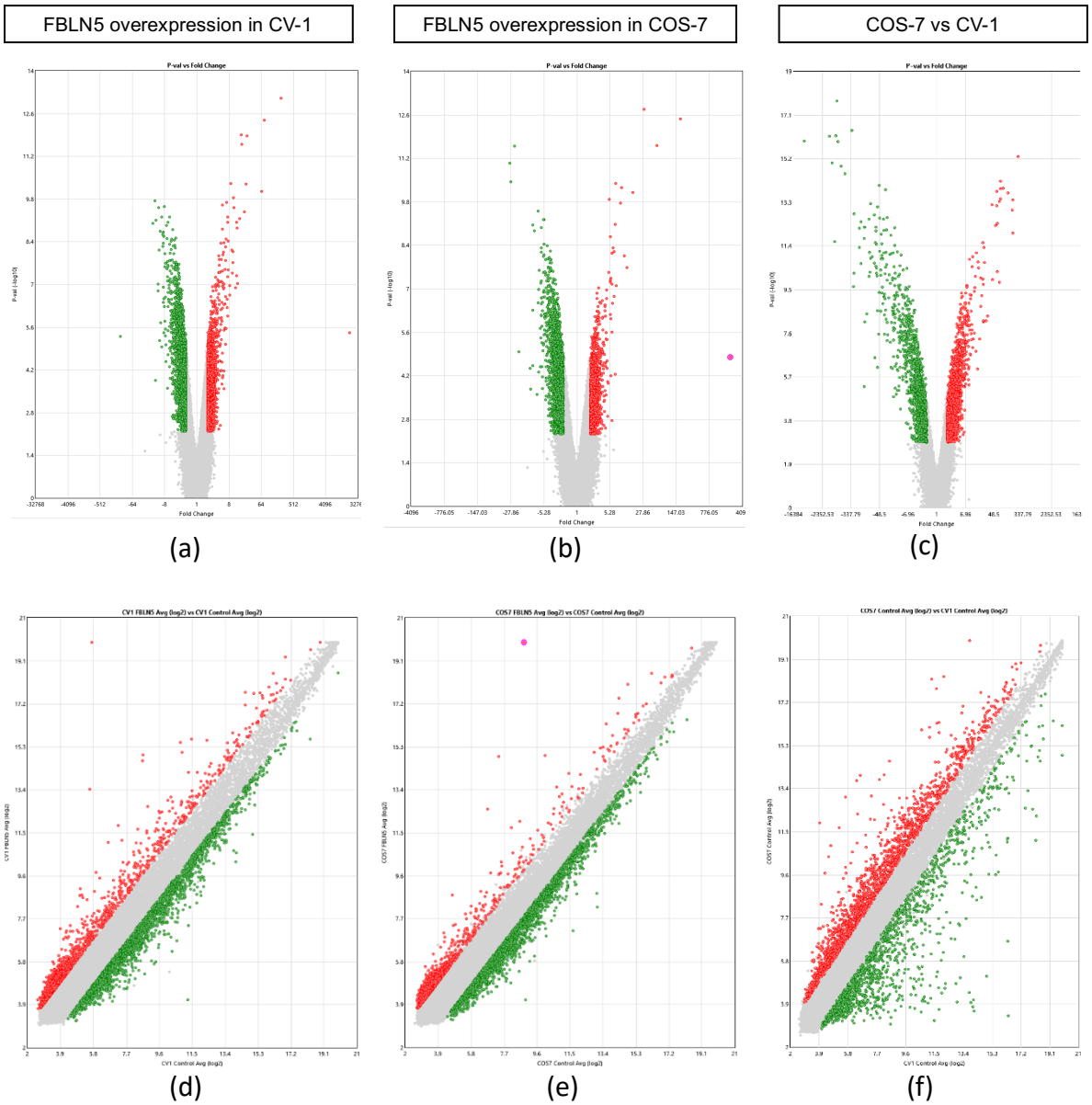


***Figure 6.10*: Volcano plots and scatter plots for the differential gene expression analysis.** (a) volcano plot for CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells, (b) volcano plot for COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells, (c) volcano plot for wild type COS-7 cells *vs* wild type CV-1 cells, (d) scatter plot for CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells, (e) scatter plot for COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells, (f) scatter plot for wild type COS-7 cells *vs* wild type CV-1 cells. Each dot presents one transcript: red = higher fold change, green = lower fold change, grey = did not pass the filter. Filter: $-2 \leq$ fold change ≥ 2, p value ≤ 0.05, FDR ≤ 0.05. For volcano plots, x-axis: fold change, left negative, right positive, y-axis: $-\log_{10}$(p value). For scatter plots, x-axis: $\log_2$ (average signal intensity) for control condition, y-axis: $\log_2$ (average signal intensity) for test condition. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

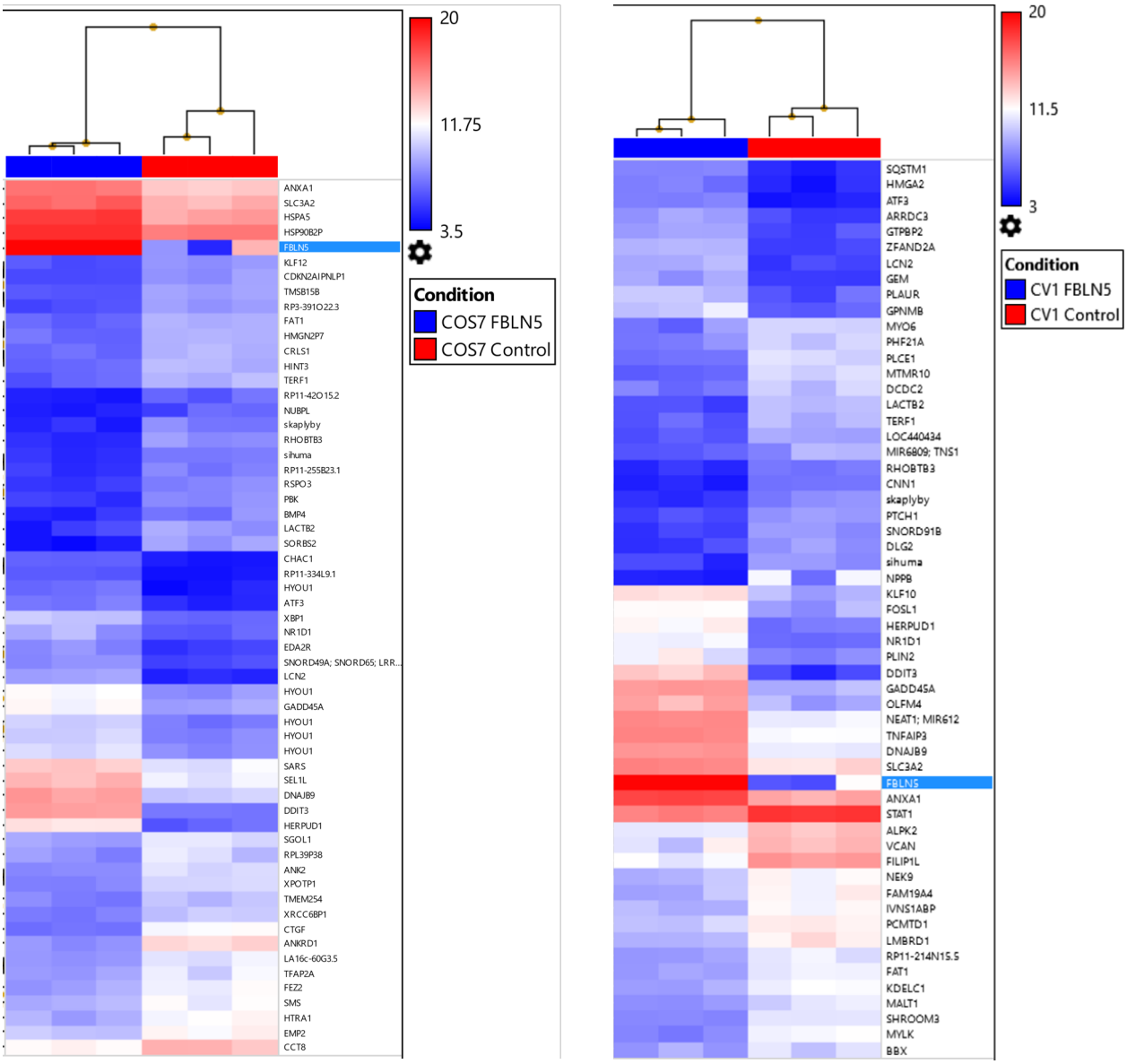


***Figure 6.11*: Heatmaps of a subset of differentially expressed genes between COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells (left) and CV-1 cells transfected with FBLN5 *vs* wild type CV- cells (right).** Left: COS-7 cells, filter applied: -5 ≤ fold change ≥ 5, p ≤ 0.05 and FDR ≤ 0.05. Only 'Coding' and 'Multiple-Complexes' included in the heatmap. Right: CV-1 cells, filter applied: -6 ≤ fold change ≥ 7.5, p ≤ 0.05, FDR ≤ 0.05. Only 'Coding' and 'Multiple-Complexes' included in the heatmap. FBLN5 is highlighted. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

6.2.4.3 Comparative transcriptomics of FBLN5 overexpressing CV-1 and COS-7

By and large, the FBLN5 overexpression resulted in up/downregulation of different sets of genes in CV-1 and COS-7 cells. Small samples of differentially expressed genes in CV-1 and COS-7 cells upon FBLN5 transfection are depicted in figure 6.10. Moreover, a Venn diagram with a comparison of the differentially expressed transcripts between CV-1 and COS-7 cells upon FBLN5 overexpression (that passed the filter) is shown in figure 6.12. Only 20.1% transcripts were common between FBLN5 overexpressing CV-1 and COS-7 cells.

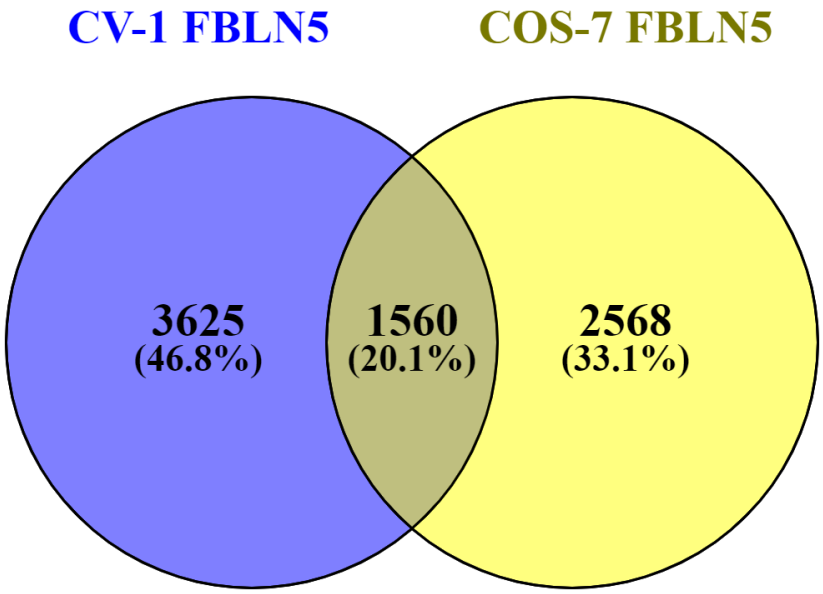


***Figure 6.12*: Comparison of differentially expressed genes between CV-1 cell (blue) and COS-7 (yellow) upon FBLN5 overexpression**. Both of the differentially expressed gene sets were generated using Transcriptome Analysis Console (TAC) 4.0.1 and fed to the online Venn diagram generation tool 'Venny 2.1' (http://bioinfogp.cnb.csic.es/tools/venny/). The differentially expressed gene sets included all transcripts which passed filter: -2 ≤ Fold change ≥ 2, p ≤ 0.05, FDR ≤ 0.05.

Figure 6.13 shows a comparative heatmap of differential expression across all four conditions, i.e., wild type COS-7 cells, COS-7 cells transfected with FBLN5, wild type CV-1 cells and CV-1 cells transfected with FBLN5. *Prima facie*, a bird's-eye view of the comparative heat map did not reveal any clusters of genes which were following the opposite trend, that is, upregulated in CV-1 cells upon FBLN5 overexpression and the opposite in the COS-7 cells. This heatmap represents a very selective gene set filtered according to the F- value. A very long time could be spent going through the list of

individual genes if I were to visually find the causative trends. Instead, I employed the Comparison Analysis using the Ingenuity Pathway Analysis (IPA) [86, 207].

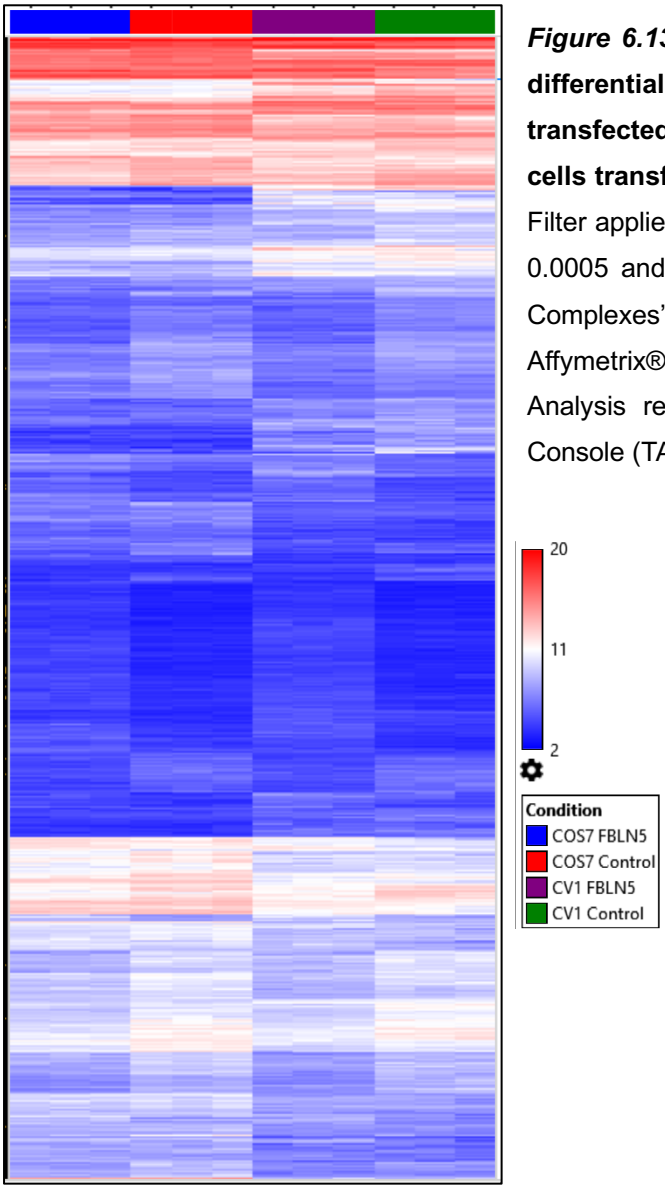


***Figure 6.13*: The heatmap representing a subset of differentially expressed genes between COS-7 cells transfected with FBLN5, wild type COS-7 cells, CV-1 cells transfected with FBLN5 and wild type CV- cells.** Filter applied: -2 < Fold change > 2, condition-F test, F < 0.0005 and FDR < 0.005. Only 'Coding' and 'Multiple-Complexes' included in the heatmap. Platform: Affymetrix® GeneChip™ Clariom™ D human array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

6.2.4.4 Comparative pathway analyses of FBLN5 overexpressing CV-1 and COS-7

The individual lists of the differentially expressed genes ('wild type COS-7 cells *vs* wild type CV-1 cells', 'COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells' and 'CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells') were provided to the Ingenuity Pathway Analysis (IPA) [86, 207]. Similar to the approach employed in the case of TMEFF2 overexpressing MCF-7 cells (figure 5.17), first, the lists of differentially expressed genes with a slightly loose filter (fold change ≥ 1.5 to -1.5 ≤, $p < 0.05$ and FDR < 0.25) were provided to IPA. Moreover, the transcripts type provided to IPA excluded 'non-

coding' transcripts. 'Multiple_complex', 'Precusor_MicroRNA',' Small_RNA', 'Pseudogenes', 'Ribosomal', 'tRNA' and 'Unassigned' transcripts were included in the list of transcripts provided to IPA along with the 'Coding' transcripts. In the case of the 'COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells', from the list of 13276 transcripts provided to IPA, the programme could successfully map 7836 transcripts while 5440 transcripts remained unmapped (IPA could not identify the transcripts). In the list of 7837 transcripts, a stringent filter was applied again with fold change ≥ 2 to -2 ≤, p ≤ 0.05 and FDR ≤ 0.05. A total of 1924 transcripts passed the filter with 1391 downregulated and 533 upregulated transcripts. Whereas for the 'CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells', from the list of 15308 transcripts provided to the IPA, the programme could successfully map 9427 transcripts while 5881 transcripts remained unmapped. In the list of 9427 transcripts, the stringent filter was applied again with fold change ≥ 2 to -2 ≤, p ≤ 0.05 and FDR ≤ 0.05. A total of 2385 transcripts passed the filter with 1341 downregulated and 1044 upregulated transcripts. Regarding the third gene set (wild type COS-7 cells *vs* wild type CV-1 cells), from the list of 5973 transcripts provided to IPA, the programme could successfully map 4936 transcripts while 1037 transcripts remained unmapped. In the list of 4936 transcripts, the stringent filter was applied again with fold change ≥ 2 to -2 ≤, p ≤ 0.05 and FDR ≤ 0.05. A total of 2398 transcripts passed the filter with 986 downregulated and 1412 upregulated transcripts. Each differentially expressed gene set was subjected to the IPA Core Analysis (individually). The three core analyses generated by IPA were then subjected to the IPA Comparison Analysis (figure 6.7) where the results of the Canonical Pathways analysis, the Upstream Regulators Analysis and the Downstream Effects Analysis were compared among the three datasets. I will only report the results of the IPA Comparison Analyses rather than individually discussing the pathway changes in each individual condition tested.

6.2.4.4.1 Canonical Pathways Comparison Analysis

The results of the IPA Canonical Pathways Comparison Analysis are shown in figure 6.14. The analysis revealed that (i) the differentially expressed genes in all three datasets overlapped a diverse range of canonical pathways which appeared mostly unrelated to each other, (ii) the SV40 transformation of CV-1 cells appeared to primarily activate pathways while FBLN5 overexpression in both CV-1 and COS-7 cells mostly resulted in the inhibition of canonical pathways, and (iii) the differentially expressed gene sets in the 'wild type COS-7 cells *vs* wild type CV-1 cells' and 'CV-1 cells transfected with FBLN5 *vs* wild type CV-1 cells' dominated the overlap with the canonical pathways. On the other hand, the differentially expressed genes in 'COS-7 cells transfected with FBLN5 *vs* wild type COS-7 cells' showed

a significant overlap with only 5 pathways. Importantly, a strict filter of $-\log_{10}$(p-value) ≥ 1.3, activation z score ≥ 2.5 and $-\log_{10}$(B-H adjusted p-value) ≥ 2 was applied in order to get a subset of the most significantly affected pathway. This meant that at least one of the conditions tested (individual pathway analyses from the differentially expressed gene set for that condition) must fulfil the criteria set above. This method was followed in order to shorten the long list of pathways involved in all three conditions. As the SV40 transformation dominated the canonical pathways overlap, it might have masked some pathways activated or inhibited by FBLN5 overexpression in CV-1 and COS-7 cells. Nonetheless, the pathways depicted in figure 6.14 can serve as a starting point for further analyses and hypothesis generation. Additionally, for the purpose of clarity, a significance threshold of $-\log_{10}$(p-value) ≥ 1.5 was applied to the comparison analysis presented in figure 6.14 and the analyses which did not show a significant overlap with the relevant canonical pathways are presented with a dot in their heatmap cell. In the interest of time, I only concentrated on two canonical pathways – the 'cyclins and cell cycle regulation' and 'oxidative phosphorylation'. In the case of the 'cyclins and cell cycle regulation' pathway, as expected, the SV40 transformation largely activated this pathway. However, it appears that FBLN5 overexpression inhibited this pathway in COS-7 cells. The comparison table of the molecules involved in the 'cyclins and cell cycle regulation' signalling pathway is listed in table 6.2. A noteworthy observation was that upon SV40 transformation, cyclins A2, B2, D3, E1, E2 and cyclin H were upregulated. The overexpression of FBLN5 in COS-7 cells resulted in the downregulation of all of these cyclins but the same was not observed in the FBLN5 overexpressing CV-1 cells. Instead, the FBLN5 overexpression in CV-1 cells resulted in the upregulation of cyclin D1 (which was unaffected upon SV40 transformation). When it comes to the 'oxidative phosphorylation', the pathway appeared to be inhibited by FBLN5 overexpression in both COS-7 and CV-1 cells while remaining unaffected during SV40 transformation of CV-1 cells. However, the inhibition of the 'oxidative phosphorylation' pathway appeared to be stronger in COS-7 cells when compared to CV-1 cells. The comparison table of genes involved in the 'oxidative phosphorylation' signalling pathway is listed in table 6.3. Although the differences in the inhibition of oxidative phosphorylation appeared minute, the inhibition of this process could have serious consequences for the rapidly dividing cells.

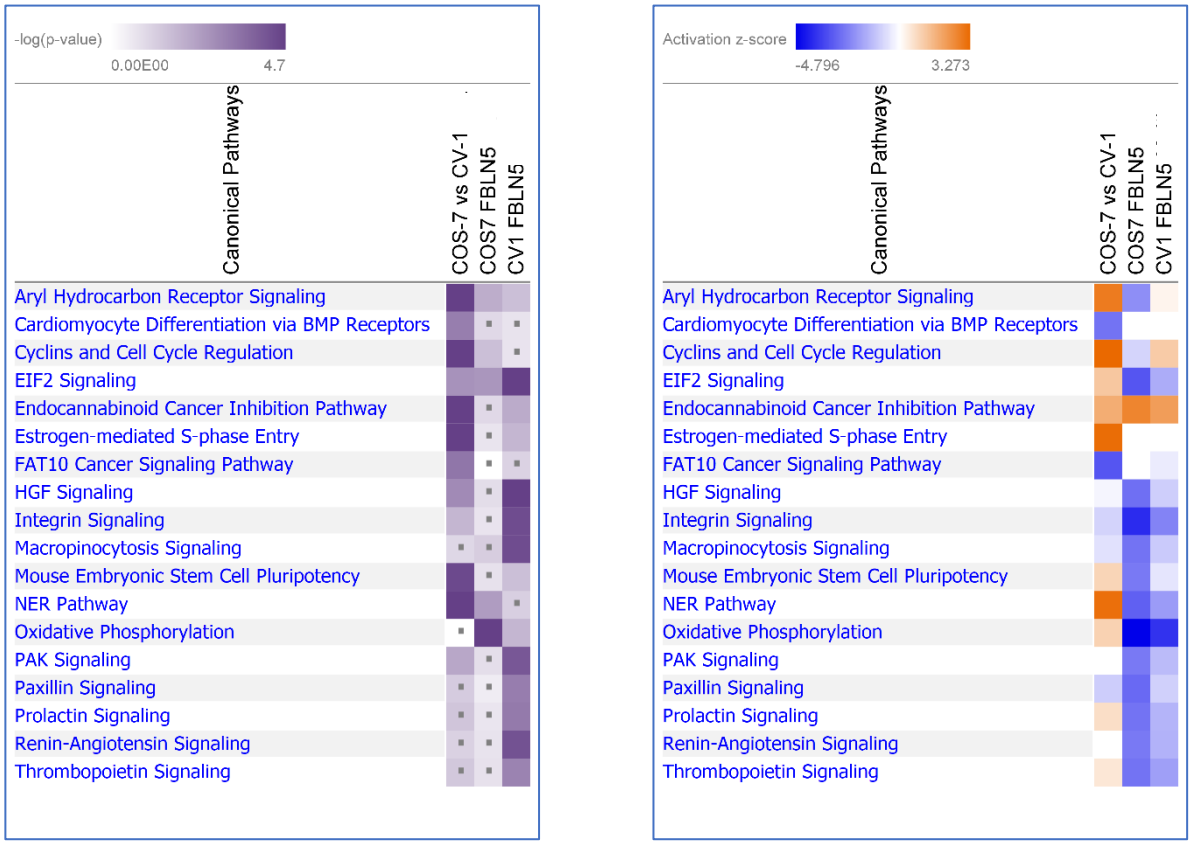


***Figure 6.14*: Canonical Pathways Comparison Analysis.** Comparison analysis results are based on the IPA Canonical Pathways analyses for differentially expressed genes between 'wild type COS-7 cells *vs* wild type CV-1 cells', 'COS-7 cells transiently transfected with FBLN5 *vs* wild type COS-7 cells' and 'CV-1 cells transiently transfected with FBLN5 *vs* wild type CV-1 cells'. These are designated as COS-7 vs CV-1, COS7 FBLN5 and CV1 FBLN5 respectively. The comparison analysis was cut off at $-\log_{10}$(p-value) ≥1.3, $-\log_{10}$(Benjimini-Hochberg p-value) ≥ 2 and activation z score ≥ 2.5. A threshold of $-\log_{10}$(p-value) ≥ 1.5 was applied to the results thus obtained. The dot represents a $-\log_{10}$(p-value) ≤ 1.5 – the results with the dot did not pass the significance threshold. DGE and pathway Analyses filter: -2 ≤ fold change ≥ 2, p ≤ 0. 05 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. DGE analyses were performed using Transcriptome Analysis Console (TAC) 4.0.1. Pathway analyses and the subsequent comparison analysis were performed using Ingenuity Pathway Analysis - IPA (QIAGEN®).

***Table 6.2*: A comparison of molecules involved in 'Cyclins and Cell cycle Regulation'** IPA canonical pathway extracted from figure 6.14.

Symbol	Entrez Gene Name	COS-7 Cells vs CV-1 Cells		FBLN5 overexpressing COS-7 cells		FBLN5 overexpressing CV-1 cells	
		Expr Fold Change	FDR	Expr Fold Change	FDR	Expr Fold Change	FDR
CCNA2	cyclin A2	↑ 2.73	8.70E-03	↓ -2.04	2.15E-02		
CCNB2	cyclin B2	↑ 4.65	2.00E-04	↓ -1.83	6.13E-02		
CCND1	cyclin D1					↑ 2.04	3.44E-02
CCND3	cyclin D3	↑ 7.65	2.85E-07	↓ -4.4	4.50E-05		
CCNE1	cyclin E1	↑ 2.51	2.10E-03	↓ -3	3.00E-03	↓ -1.84	1.16E-01
CCNE2	cyclin E2	↑ 3	2.00E-02	↓ -1.93	9.06E-02	↓ -2	3.25E-02
CCNH	cyclin H	↑ 2.29	5.80E-03	↓ -1.91	1.57E-02		
CDK1	cyclin dependent kinase 1	↑ 3.56	2.00E-04				
CDK2	cyclin dependent kinase 2	↑ 3.78	4.00E-04				
CDK4	cyclin dependent kinase 4	↑ 4.78	9.00E-04			↑ 2.29	2.06E-02
CDK6	cyclin dependent kinase 6	↓ -2.37	1.20E-03	↓ -2.62	1.80E-03		
CDK7	cyclin dependent kinase 7	↑ 5.22	4.10E-03			↑ 1.84	1.41E-01
CDKN1A	cyclin dependent kinase inhibitor 1A	↓ -8.63	6.85E-05	↑ 1.53	1.63E-01	↑ 2.92	9.00E-04
CDKN2C	cyclin dependent kinase inhibitor 2C	↑ 1.55	1.24E-01	↓ -2.2	8.70E-03	↓ -2.03	1.31E-02
E2F1	E2F transcription factor 1	↑ 2.99	6.10E-03				
E2F4	E2F transcription factor 4	↑ 2.23	4.84E-02				
E2F6	E2F transcription factor 6	↑ 8.6	1.86E-05				
HDAC1	histone deacetylase 1			↓ -2.3	1.43E-02	↓ -2.08	1.03E-01
HDAC2	histone deacetylase 2			↓ -2.26	8.20E-03	↓ -2.56	5.00E-04
HDAC3	histone deacetylase 3			↓ -2.42	3.10E-02		
HDAC5	histone deacetylase 5	↓ -3.62	4.08E-02			↓ -2.31	8.56E-02
PPP2R2B	protein phosphatase 2 regulatory subunit B beta	↓ -3.76	7.00E-04			↓ -3.28	8.00E-04
PPP2R2C	protein phosphatase 2 regulatory subunit B gamma	↑ 2.22	1.69E-02			↑ 1.83	2.23E-02
PPP2R5D	protein phosphatase 2 regulatory subunit B 'delta	↑ 2.64	1.90E-03	↓ -1.93	1.28E-02	↓ -1.6	2.07E-02
RB1	RB transcriptional corepressor 1	↓ -2.81	8.00E-04	↓ -1.51	1.55E-01		
SKP2	S-phase kinase associated protein 2	↑ 5.35	3.13E-05	↓ -4.2	4.00E-04	↓ -4.09	5.00E-04
SUV39H1	suppressor of variegation 3-9 homolog 1	↑ 2.29	7.00E-04				
TGFB2	transforming growth factor beta 2	↓ -2.8	8.08E-05	↓ -1.85	1.25E-02		
TGFB3	transforming growth factor beta 3	↓ -2.26	1.10E-03				
TP53	tumour protein p53			↓ -1.58	1.59E-01	↓ -2.43	2.60E-03
WEE1	WEE1 G2 checkpoint kinase	↑ 2.25	1.83E-01	↓ -3.02	1.01E-02		

© 2000-2019 QIAGEN. All rights reserved.

***Table 6.3*: A comparison of molecules involved in 'Oxidative Phosphorylation'** IPA canonical pathway extracted from figure 6.14.

Symbol	Entrez Gene Name	COS-7 cells vs CV-1 cells		FBLN5 overexpressing COS-7 cells		FBLN5 overexpressing CV-1 cells	
		Expr Fold Change	FDR	Expr Fold Change	FDR	Expr Fold Change	FDR
ATP5MF	ATP synthase membrane subunit f			↓ -2.66	4.20E-03	↓ -1.79	8.80E-03
ATP5F1E	ATP synthase F1 subunit epsilon			↓-3.09	2.32E-02		
ATP5PD	ATP synthase peripheral stalk subunit d			↓-3.94	4.50E-05	↓ -3.44	4.86E-05
ATP5PF	ATP synthase peripheral stalk subunit F6			↓-3.01	1.60E-03	↓ -3.89	7.00E-04
ATP5PO	ATP synthase peripheral stalk subunit OSCP			↓ -2.22	1.50E-02	↓ -1.85	2.36E-02
ATP5S	ATP synthase, H+ transporting, mitochondrial Fo complex subunit s (factor B)	↑2.08	2.00E-02				
ATPAF1	ATP synthase mitochondrial F1 complex assembly factor 1			↓ -2.09	1.76E-02		
COX17	cytochrome c oxidase copper chaperone COX17			↓ -2.26	2.00E-03		
COX6A1	cytochrome c oxidase subunit 6A1			↓ -3.11	3.00E-04	↓ -3.25	1.00E-04
COX6B1	cytochrome c oxidase subunit 6B1			↓ -2.04	4.00E-03	↓ -1.92	1.24E-02
COX7A1	cytochrome c oxidase subunit 7A1	↓-169.24	1.48E-09				
COX7A2	cytochrome c oxidase subunit 7A2			↓ -2.11	1.20E-02	↓ -1.69	2.17E-02
COX7B	cytochrome c oxidase subunit 7B			↓ -3.45	5.00E-04	↓ -2.75	5.00E-04
CYB5A	cytochrome b5 type A			↓ -2.02	6.10E-03	↓ -1.62	3.76E-02
CYC1	cytochrome c1	↑2.62	2.00E-04				
NDUFA1	NADH:ubiquinone oxidoreductase subunit A1			↓ -2.7	3.50E-03	↓ -1.63	1.86E-02
NDUFA2	NADH:ubiquinone oxidoreductase subunit A2			↓ -2.73	2.70E-03	↓ -1.66	1.05E-01
NDUFA3	NADH:ubiquinone oxidoreductase subunit A3			↓ -3.05	2.51E-02	↓ -2.66	1.01E-02
NDUFA7	NADH:ubiquinone oxidoreductase subunit A7	↑2.45	2.69E-02	↓ -2.11	3.52E-02		
NDUFAB1	NADH:ubiquinone oxidoreductase subunit AB1			↓ -2.9	6.80E-03		
NDUFB2	NADH:ubiquinone oxidoreductase subunit B2			↓ -2.83	2.60E-03	↓ -2.69	1.40E-03
NDUFB3	NADH:ubiquinone oxidoreductase subunit B3	↓-2.3	1.55E-02	↓ -2.96	3.30E-03	↓ -2.98	2.60E-03
NDUFB6	NADH:ubiquinone oxidoreductase subunit B6			↓ -2.18	2.29E-02	↓ -1.89	3.92E-02
NDUFB7	NADH:ubiquinone oxidoreductase subunit B7	↑2.88	5.90E-03				
NDUFB8	NADH:ubiquinone oxidoreductase subunit B8			↓ -2.6	2.20E-03	↓ -2.52	1.50E-03
NDUFS4	NADH:ubiquinone oxidoreductase subunit S4			↓ -2.99	2.17E-02	↓ -1.8	7.45E-02
NDUFS6	NADH:ubiquinone oxidoreductase subunit S6			↓ -2.46	1.00E-03	↓ -1.97	6.20E-03
NDUFS7	NADH:ubiquinone oxidoreductase core subunit S7	↑2.3	2.42E-02				
NDUFS8	NADH:ubiquinone oxidoreductase core subunit S8	↑2.23	3.90E-03				
NDUFV2	NADH:ubiquinone oxidoreductase core subunit V2			↓ -3.05	6.80E-03	↓ -1.99	1.28E-02
NDUFV3	NADH:ubiquinone oxidoreductase subunit V3			↓ -2.84	1.33E-02	↓ -2.38	9.81E-02
SDHC	succinate dehydrogenase complex subunit C	↓-2.49	4.26E-02				
UQCR10	ubiquinol-cytochrome c reductase, complex III subunit X			↓ -2	4.09E-02	↓ -2.15	1.05E-02

UQCRB	ubiquinol-cytochrome c reductase binding protein			↓ -2.29	2.47E-02	↓ -1.55	1.35E-01
UQCRC2	ubiquinol-cytochrome c reductase core protein 2			↓ -2.4	2.00E-03	↓ -2.72	8.00E-04
© 2000-2019 QIAGEN. All rights reserved.							

6.2.4.4.2 Upstream Regulator Comparison Analysis

The results of the Upstream Regulator Comparison Analysis are depicted in figure 6.15. The pro-proliferation regulators like CD28, CD3 and MYC were predicted to be inhibited in the FBLN5 overexpressing COS-7 cells but they were either unaffected or slightly activated in the FBLN5 overexpressing CV-1 cells. Surprisingly, other potent proliferative regulators like ERK1/2 and EGF were predicted to be activated in the FBLN5 overexpressing CV-1 cells but not in the FBLN5 overexpressing COS-7 cells. E2F1 was predicted to be inhibited in both FBLN5 overexpressing CV-1 and COS-7 cells but the degree of inhibition was stronger in FBLN5 overexpressing COS-7 cells. Similar was the case with oestrogen receptor 1 (ESR1) and TBX2. Notably, two microRNAs (*let-7* and *mir21*) were predicted to be inhibited upon SV40 transformation, the FBLN5 overexpression not only alleviated the predicted *mir-21* inhibition, but more accurately, it activated this regulator in COS-7 cells. In the interest of time, only MYC will be discussed in more detail. The activation of MYC in wild type COS-7 cells in comparison to wild type CV-1 cells is depicted in figure 6.16 and the schematic diagrams of upstream regulation by MYC in FBLN5 overexpressing COS-7 and CV-1 cells are shown in figures 6.17 and 6.18 respectively. MYC was predicted to be potently activated upon SV40 transformation. Moreover, it appeared that the FBLN5 overexpression resulted in a potent inhibition of MYC in COS-7 cells. Lastly, as the comparison analysis was aimed at pointing the differences among the three conditions, it did not show a complete picture of the predicted activation and inhibition of upstream regulators in wild type COS-7 cells when compared to the wild type CV-1 cells. A subset of results from the Upstream Regulator Analysis of the differentially expressed genes between wild type COS-7 cells *vs* wild type CV-1 cells is presented in supplementary figure S.12. p53 and E2F1 were among the topmost upstream regulators in this analysis, where the former was predicted to be inhibited while the latter was predicted to be activated. Conversely, p53 and the RB are the most well-known targets of SV40 large T-antigen [1, 257]. RB in turn imparts its action largely through inhibition of E2F transcription factors (E2F1, E2F2 and E2F3 in particular) [1]. The results depicting the inhibition of p53 and activation of E2F1 upon transformation by the SV40 DNA were therefore indicative of the authenticity of this analysis in particular and the experimental approach in general. Furthermore, the potent

activation (predicted) of ERBB2 and MYC upon SV40 transformation is worthy of further deliberation but beyond the scope of this thesis.

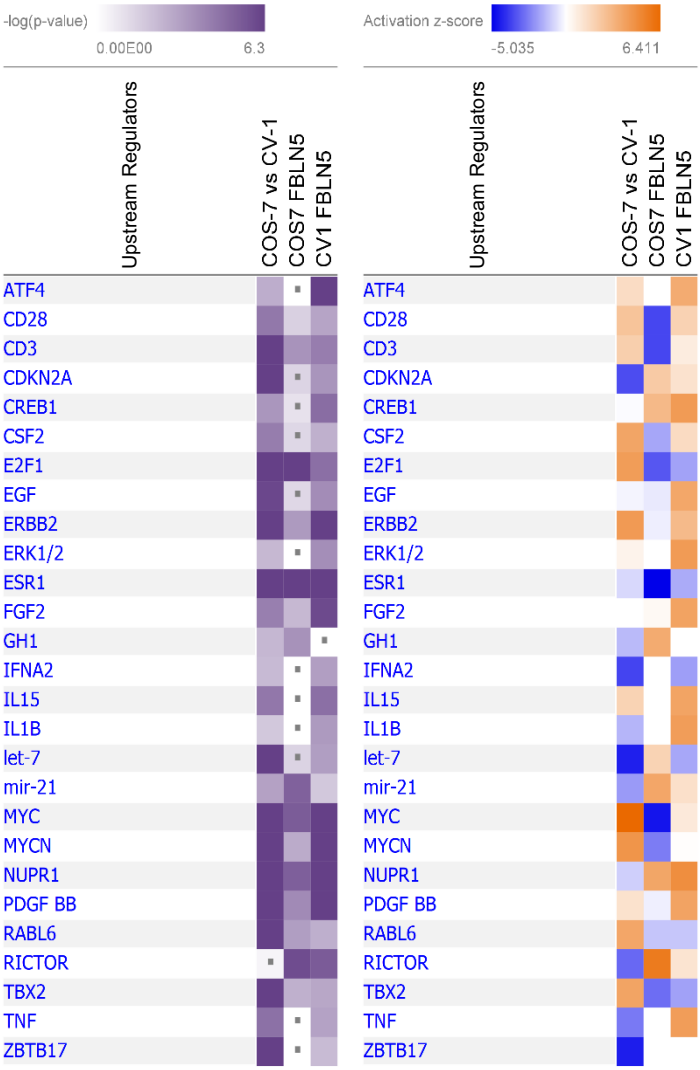


***Figure 6.15*: Upstream Regulator Comparison Analysis.** Comparison analysis results are based on the IPA Upstream Regulator analyses for differentially expressed genes between 'wild type COS-7 cells *vs* wild type CV-1 cells', 'COS-7 cells transiently transfected with FBLN5 *vs* wild type COS-7 cells' and 'CV-1 cells transiently transfected with FBLN5 *vs* wild type CV-1 cells'. These are designated as COS-7 vs CV-1, COS7 FBLN5 and CV1 FBLN5 respectively. The comparison analysis was cut off at $-\log_{10}$(p-value) ≥ 1.3, $-\log_{10}$(Benjimini-Hochberg p-value) ≥ 1.7 and activation z score ≥ 3.5. A threshold of $-\log_{10}$(p-value) ≥ 1.5 was applied to the results thus obtained. The dot represents a $-\log_{10}$(p-value) ≤ 1.5 – the results with the dot did not pass the significance threshold. DGE and pathway Analysis filter: -2 ≤ fold change ≥ 2, p ≤ 0. 05 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. DGE analyses were performed using Transcriptome Analysis Console (TAC) 4.0.1. Pathway analyses and the subsequent comparison analysis were performed using Ingenuity Pathway Analysis - IPA (QIAGEN®).

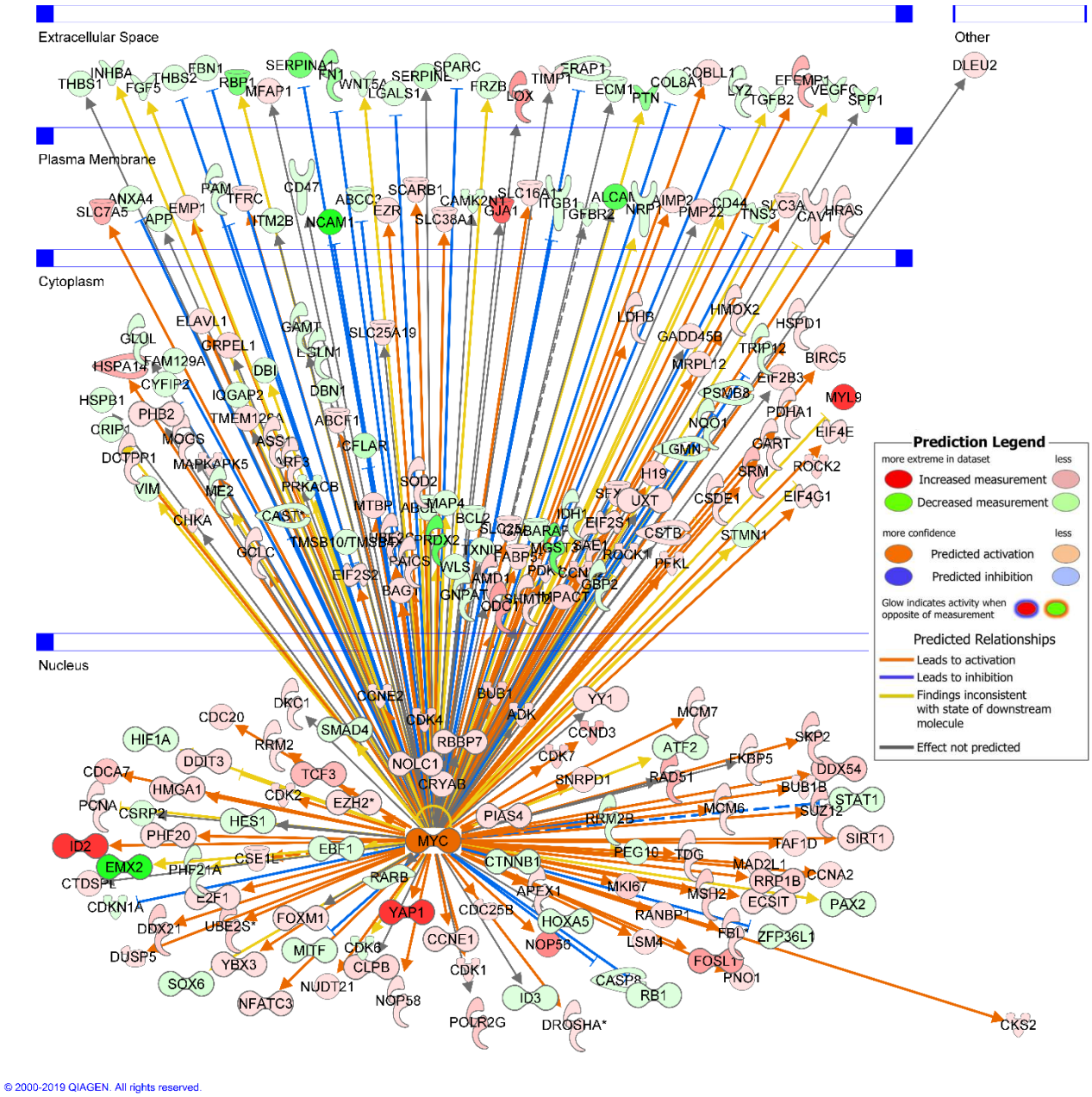
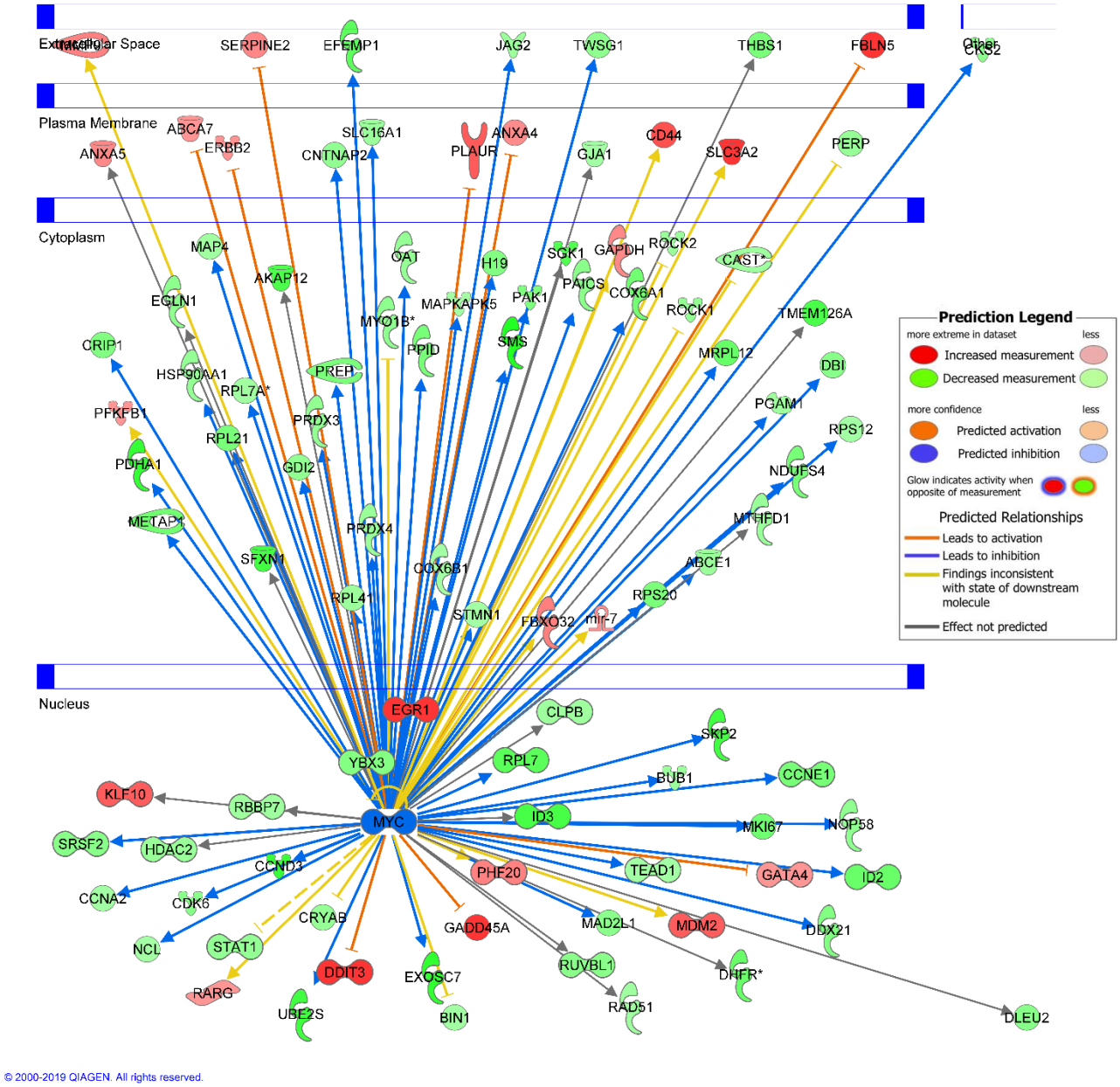
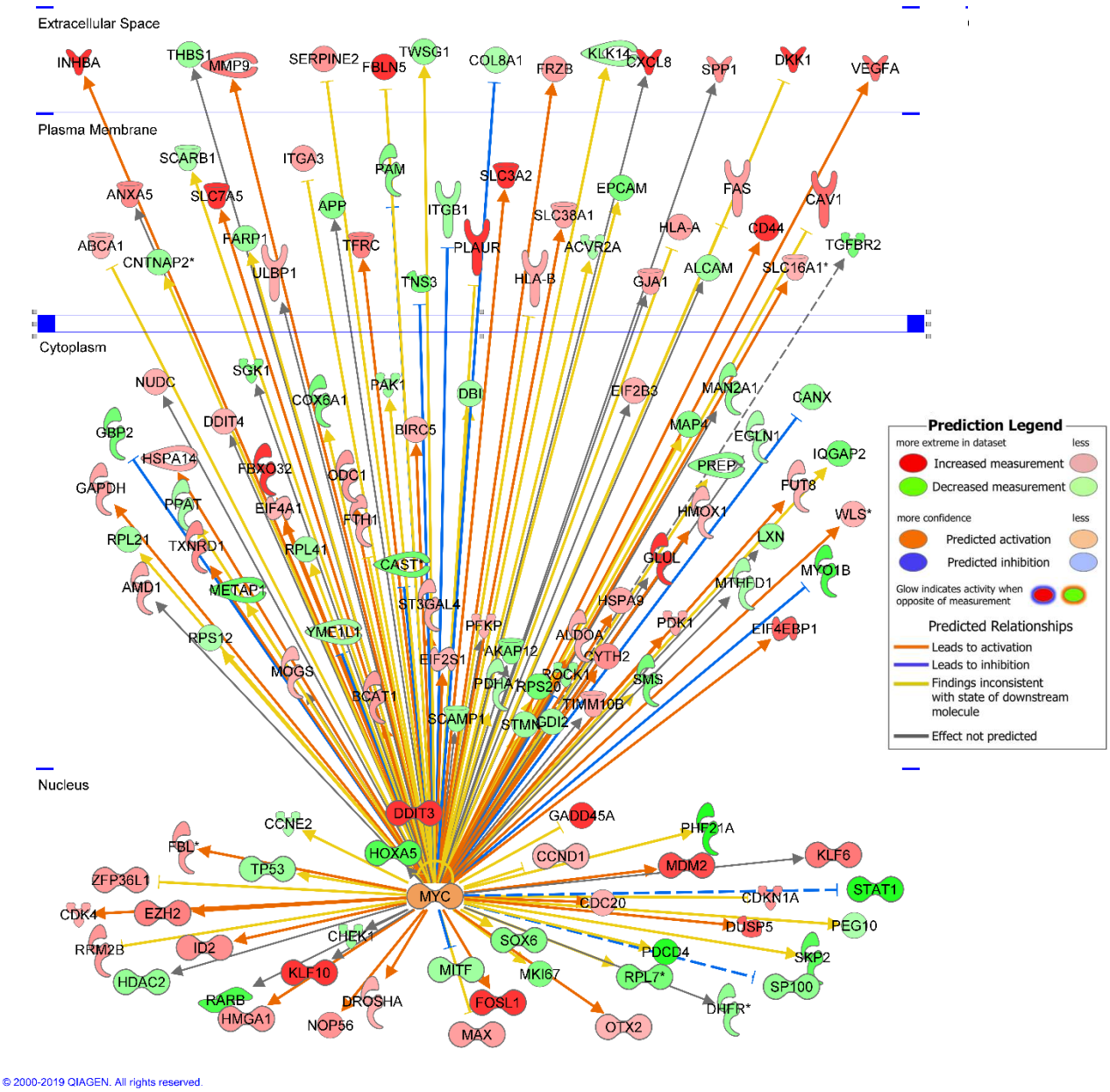


***Figure 6.16*: Network view of MYC as the predicted upstream regulator for the transcriptomic changes upon transformation of CV-1 cells (to COS-7 cells).** IPA Upstream Regulator Analysis was run on the differentially expressed gene set between wild type COS-7 and wild type CV-1 cells. DGE Analysis filter; fold change ≥ 2 to -2 ≤, p ≤ 0. 05 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and source: Ingenuity Pathway Analysis - IPA (QIAGEN ®)



© 2000-2019 QIAGEN. All rights reserved.

***Figure 6.17*: Network view of MYC as the predicted upstream regulator for the transcriptomic changes observed upon FBLN5 overexpression in COS-7 cells**. IPA Upstream Regulator Analysis was run on the differentially expressed gene set between COS-7 cells transfected with FBLN5 vs wild type COS-7 cells. DGE Analysis filter; fold change ≥ 2 to -2 ≤, p ≤ 0. 05 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and source: Ingenuity Pathway Analysis - IPA (QIAGEN ®).



© 2000-2019 QIAGEN. All rights reserved.

***Figure 6.18*: Network view of MYC as the predicted upstream regulator for the transcriptomic changes observed upon FBLN5 overexpression in CV-1 cells**. IPA Upstream Regulator Analysis was run on the differentially expressed gene set between CV-1 cells transfected with FBLN5 vs wild type CV-1 cells. DGE Analysis filter; fold change ≥ 2 to -2 ≤, p ≤ 0. 05 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Pathway analysis and source: Ingenuity Pathway Analysis - IPA (QIAGEN ®).

6.2.4.4.3 Downstream Effects Comparison Analysis

As mentioned before, the Downstream Effects Analysis is primarily useful for investigating the effectors for which the downstream functions are not known. Concerning the current study, the downstream effects of SV40 transformation and FBLN5 overexpression in COS-7 cells were already

known, i.e., FBLN5 caused cell death and reduction in cell viability in COS-7 cells. Instead, I was interested in identifying the upstream causative factors which were leading to this function of FBLN5. Nonetheless, the results of this analysis corroborated the previously identified functions. The results of the Downstream Effects Comparison Analysis as a subset of 'Diseases and Biological Functions' are shown in figure 6.19. The differentially expressed genes between wild type COS-7 cells and wild type CV-1 cells appeared to be involved in diseases/functions like 'liver cancer' and 'replication of virus' which were indicative of the authenticity of data. Similar was the case with functions like 'morbidity or mortality' and 'necrosis' and 'organismal death', which were predicted to be inhibited upon SV40 transformation. Moreover, according to this analysis, the FBLN5 overexpression in COS-7 cells was predicted to activate functions like 'morbidity or mortality' and 'organismal death'. A minute activation of these functions was predicted in FBLN5 overexpressing CV-1 cells as well. Upon a close observation of the sets of molecules involved in 'morbidity or mortality', caspase 7 was upregulated by 2.07-fold upon FBLN5 overexpression in COS-7 cells, whereas, the mRNA levels of caspase 7 remained unaffected either upon SV40 transformation or FBLN5 overexpression in CV-1 cells. Interestingly, 'necrosis' and 'necrosis of tumour' were predicted to be inhibited in FBLN5 overexpressing CV-1 cells and activated in FBLN5 overexpressing COS-7 cells.



***Figure 6.19*: A subset of results from IPA Downstream Effects Comparison Analysis.** Comparison analysis results are based on the IPA pathway analyses for differentially expressed genes between 'wild type COS-7 cells *vs* wild type CV-1 cells', 'COS-7 cells transiently transfected with FBLN5 *vs* wild type COS-7 cells' and 'CV-1 cells transiently transfected with FBLN5 *vs* wild type CV-1 cells'. These are designated as COS-7 vs CV-1, COS7 FBLN5 and CV1 FBLN5 respectively. The comparison analysis was cut off at $-\log_{10}$(p-value) ≥ 1.3, $-\log_{10}$(Benjimini-Hochberg p-value) ≥ 1.3 and activation z score ≥ 2.5. A threshold of $-\log_{10}$(p-value) ≥ 1.5 was applied to the results thus obtained. The dot represents a $-\log_{10}$(p-value) ≤ 1.5 – the results with the dot did not pass the significance threshold. DGE and Pathway Analyses filter: -2 ≤ fold change ≥ 2, p ≤ 0.05 and FDR ≤ 0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human array. DGE analyses were performed using Transcriptome Analysis Console (TAC) 4.0.1. Pathway analyses and the subsequent comparison analysis were performed using Ingenuity Pathway Analysis - IPA (QIAGEN®).

6.3 Discussion

Assessing the differences in function(s) of an effector between two biological systems is a challenging task. The challenges are even more daunting when it comes to comparing the transcriptomic changes between two biological systems upon activation or inhibition of the upstream effector. The biological systems seldom work in isolation; the genetic and biochemical interactions in even the relatively simpler prokaryotic organisms are far more complex than probably the most advanced electronic circuits. In case the downstream transcriptional effects imparted by a stimulant between two eukaryotic organisms or cell lines have to be compared, the number of variables involved would increase by several orders of magnitude. This implies that the differences in transcriptomic changes between two cell lines upon overexpression of a gene would not necessarily reflect the differences in its functioning. In order to undertake such a study where the functioning of an effector between two cell lines needs to be compared through assessing the transcriptomic changes imparted by it, it is important that the native differences between the two biological systems are kept minimal. This in turn is not possible in biological systems as even two clones in a tissue culture of eukaryotic cells will have significant genetic and biological differences. The nearest possible compromise has to be struck. In the case of my study, the availability of COS-7 cells provided such a compromise.

In Stefan Grimm's lab, the genetic screening and investigation of functioning of the anticancer genes had been undertaken in CV-1 cells over several years. The experimental protocols for working with CV-1 cells had been well optimized. Moreover, the genes I was working with were screened for their ability not to impart cell death in CV-1 cells. Although both Ghada AbuAli and Qize Ding mentioned transformation of CV-1 cells in their studies, neither reported anchorage-independent growth of 'transformed' CV-1 cells. The availability of the previously stabled transformed version of CV-1 cells, the COS-7 cells (cell line developed by transforming the CV-1 cells using the origin defective SV40 DNA [128]), provided the biological system with a near enough compromise to assess the functioning of FBLN5 through differences in transcriptomic changes.

The methodology was, however, not devoid of some concerns which need to be addressed. First of all, the cell death assays employed for COS-7 cells were generally unconventional. Assaying cell death in COS-7 cells was very difficult and considerable time, efforts and resources were spent in order to adapt and/or develop an appropriate cell death assay for these cells. Almost all of the attempts were unsuccessful owing to the exceptionally high division rate of the COS-7 cells. Assaying the cell death

in transfected cells is commonly a difficult affair. When cells are treated with lethal small molecular weight compounds, the great majority of cells in the tissue culture get treated equally (compared to each other). However, in the case of transfection, every cell does not take up the equal amount of nucleic acid. In the case of plasmid DNA, even after the uptake, the heterogeneous cell population does not express the transfected genes equally. Concerning the cells with exceptionally high division rate like that of COS-7 cells, it appears that the untransfected population of cells at the end of post-transfection incubation period surpassed the sensitivity threshold of the conventional cell-death assays. This hypothesis is aided by the fact that the majority of previously reported studies where cell death was assayed for the transfected COS-7 cells, manual phenotype inspection was employed [251, 252, 253, 256]. Traditionally, when phenotype inspection is employed in the cell death assays, luciferase is used as the reporter gene (reporter plasmid co-transfection) [42]. I replaced the reporter gene from luciferase to GFP and used co-transfection with luciferase and GFP as the negative control. Moreover, I also employed two additional assays for TMEFF2- and FBLN5-induced cell death in CV-1 and COS-7 cells. First, the reduction in the number of total green fluorescent cells upon co-transfection with GFP reporter plasmid was assayed using automated cell counting and the results corroborated with phenotype inspection. Secondly, I also used the modified version of clonogenicity assay: TMEFF2 and FBLN5 were cloned into pcDNA 3.1 (which harbours a neomycin resistance cassette), transfected into the CV-1 and COS-7 cells and selection pressure was applied by adding G418 to the media. Again, the results corroborated with both the phenotype inspection and the automated green-fluorescent cell-counting assays. Consequently, the inability of either TMEFF2 or FBLN5 in inducing cell death or reduction in viability of CV-1 cells in all three assays is itself suggestive of the authenticity of the techniques.

The second concern is regarding the difference in the degree of overexpression of TMEFF2 and FBLN5 between CV-1 and COS-7 cells as the primary cause behind the selective cell death-inducing capacity of these genes in the COS-7 cells. If this is true, it challenges the whole concept of anticancer genes, especially in the case of TMEFF2 where the degree of overexpression in CV-1 cells was particularly low. There could be three reasons for the failure of TMEFF2 in reaching the appropriate overexpression levels upon transfection in CV-1 cells: first, the inadequate introduction of plasmid into the cells (bad transfection); secondly, the inadequate transcription of the ectopically overexpressed *TMEFF2* in CV-1 cells and thirdly, there could be mechanisms only functioning in CV-1 cells which were actively inhibiting the translation of the gene or destabilizing its mRNA. If it was the

case of inefficient transfection, it is highly unlikely that all the time points transfected would give similar results for two different assays (immunoblot and qRT-PCR). In addition, even if CV-1 cells were unable to take up the plasmids during transfections, the gene expression should have increased over time proportionally, nevertheless. The results indicate otherwise – instead, they are suggestive of some unknown mechanisms in place in the CV-1 cells where, most probably, the TMEFF2 mRNA is destabilized actively. As far as the possibility of inadequate transcription of ectopically overexpressed *TMEFF2* in CV-1 cells is concerned, the overexpression of FBLN5 should be taken into consideration. FBLN5 overexpression in CV-1 cells was substantial and it increased over time proportionally. It is highly unlikely that two genes (FBLN5 and TMEFF2) in the same plasmid, under the same potent SV40 promoter would transcribe differently in the same cell line transfected using the same transfection protocol. Lastly, the possibility of inadequate translation of ectopically overexpressed *TMEFF2* can be overruled by the fact that both mRNA and protein levels of TMEFF2 failed to achieve the adequate levels of overexpression.

As for the difference in the degree of overexpression of FBLN5 between COS-7 and CV-1 cells, CV-1 cells are much more sensitive to cell death, presumably because the cell death machinery in the untransformed cells is largely intact. It is evident from the fact that the transfection with tBID completely eradicated CV-1 cells. COS-7 cells, on the other hand, appeared to be resistant to this very potent apoptosis inducer. Even if we assume that the selective cell death caused by FBLN5 in COS-7 cells was due to the difference in the degree of overexpression, the same should apply to tBID. A vastly higher sensitivity of CV-1 cells to cell death should, in theory, compensate for the marginally lower levels of FBLN5 overexpression. However, the difference in the degree of FBLN5 overexpression between CV-1 and COS-7 cells cannot be considered in isolation. Being a transformed cell line, COS-7 cells proliferate at a much higher rate than CV-1 cells, which would impose certain vulnerabilities, like the rapid metabolic rate, the demand for nutrients and the need for countering the harmful phenomena like ROS production and ER stress (due to unfolded proteins). In this perspective, even a marginal difference in the effects imparted by FBLN5 overexpression could shift the balance for worse. The higher degree of FBLN5 overexpression could contribute to the selective cell death caused by the gene in COS-7 cells but only in combination with several other factors, i.e., it is highly unlikely to be the sole reason behind this phenomenon. This argument is also aided by the fact that FBLN5 overexpression in COS-7 and CV-1 cells, by and large, resulted in the up- and downregulation of different sets of genes (figure 6.12).

Another possible objection on the experimental methodology employed in this chapter could be the use of human DNA microarrays to assess the transcriptomic changes in the cells from the vervet monkey (*C. aethiops*). Firstly, human DNA microarrays (Affymetrix® and Illumina ®) have been employed in the past for transcriptomic studies in vervet monkey tissue samples [258, 259, 260]. Secondly, two very important studies in this regard (Ref [261] and [262]) elaborated the very high degree of similarity between the genomes and transcriptomes of human and vervet monkey. The publication by Warren *et al.* [261] was the result of a vast international collaborative effort to reference the vervet monkey genome. Warren *et al.* [261] state, '*we aligned 97.9% of 46,823 human known RefSeq transcripts to the vervet genome at an average coverage of 98.8%*' (p.1923). In a similar publication, Jasinska *et al.* [262] state, '*given the high degree of conservation between the vervet and human genomes and the ready availability of inexpensive human gene expression microarrays, such arrays have provided a cost effective means for initial genome-wide profiling of vervet gene expression*' (p.131). Notably, these studies were conducted when only the transcriptome level arrays were available. I employed Affymetrix® GeneChip™, Clariom™ D microarrays which are capable of providing exon-level expression profiles of more than 540,000 transcripts. Together with the high degree of similarity between human and vervet monkey genomes, the ability to map multiple sequences at exon level, on an enormous range of transcripts would provide even further sensitivity and depth to analyse the transcriptomic changes in CV-1 and COS-7 cells accurately. In addition, the validity of transcriptomic data and the subsequent pathway analyses is further aided by the fact that when transcriptomic changes between wild type COS-7 cells *vs* wild type CV-1 cells were subjected to the IPA Upstream Regulator analysis, E2F1 was predicted to be activated and p53 was predicted to be inhibited (supplementary figure S.12). Nonetheless, qRT-PCR validation of a subset of differentially expressed transcripts would provide further authentication to the transcriptomic data.

As stated earlier, the biological systems seldom work in isolation. Almost all of the previously reported anticancer genes act at multiple levels to induce the tumour-specific cell death. It would be surprising if FBLN5 were to be any different in this regard. Regarding the results of the IPA Canonical Pathways Comparison Analysis, a strong inhibition of the oxidative phosphorylation (OXPHOS) was observed in the FBLN5 overexpressing COS-7 and CV-1 cells. A closer look at the list of genes involved in the process (table 6.3) revealed that the FBLN5 overexpression resulted in the downregulation of numerous genes involved in this vital process; moreover, the number of downregulated genes pertaining to this pathway was slightly larger in the case of COS-7 cells. It is widely accepted that

cancer cells fulfil their intensified energy demands through aerobic glycolysis rather than oxidative phosphorylation (Warburg effect) [263]. However, to the best of my knowledge, the Warburg effect in COS-7 cells has not been studied. The COS-7 cells exhibit an exceptionally high division rate (personal observation), which in turn would increase their energy demand. Compared to CV-1 cells, the heightened energy demands together with a slightly higher inhibition of oxidative phosphorylation in COS-7 cells could have caused the selective loss of cell viability in the COS-7 cells. The effects of FBLN5 overexpression in the COS-7 cells could thus be comparable to those of PARP1 hyperactivation, i.e., the bioenergetic collapse because of NAD^+ and ATP depletion, ultimately leading to parthanatos (please see section 1.4.6 for details). Moreover, the phenomenon would be similar to the selective cell death induced by MDA-7/IL-24, PAR4, parvovirus NS1 and Orcltl3, where among other factors, it was postulated that these genes induced ER stress and/or ROS production in both normal and cancer cells. The cell death is induced selectively in the cancer cells because these factors are already heightened in cancer cells as compared to normal cells. One might argue that if the inhibition of oxidative phosphorylation is the only cause of the selective cell death induced by FBLN5 in COS-7 cells, how transferable would this phenomenon be to cancer cells where the Warburg effect is largely in place? According to Solaini *et al.* [263], the inefficient aerobic glycolysis only provides a part of the required ATPs in cancer cells and the rest are still generated through oxidative phosphorylation. Besides, when considered in the context of functions of FBLN5 (on its own, not as an anticancer gene), the inhibitory effects of FBLN5 on the oxidative phosphorylation have not been reported (to the best of my knowledge). The potent inhibition of the oxidative phosphorylation may have been the reason behind the reported FBLN5-induced decrease in ROS production in pancreatic ductal adenocarcinoma by Wang *et al.* [248]. Also, the ability of FBLN5 to inhibit the oxidative phosphorylation could possibly be the reason behind the hypoxia induced and HIF-1α-dependant FBLN5 upregulation reported by Gaudall *et al.* [250].

The canonical pathway which showed a contrasting activity between COS-7 and CV-1 cells was ‘cyclins and cell cycle regulation’. Interestingly, cyclins A2, B2, D3, E1, E2 and cyclin H were upregulated in COS-7 cells when compared to CV-1 cells. Upon FBLN5 overexpression, precisely the same list of cyclins was downregulated in COS-7 cells but not in CV-1 cells. In this regard, the results of the IPA Upstream Regulator Comparison Analysis provide further insight. MYC was predicted to be potently activated in wild type COS-7 cells (compared to wild type CV-1 cells); further, upon FBLN5 overexpression, this oncogenic transcription factor was precited to be potently inhibited in COS-7

cells, whereas in CV-1 cells, MYC was predicted to be rather slightly activated (upon FBLN5 overexpression). Cyclins are largely under the transcription control of MYC. It appears that one of the factors that imparts FBLN5-induced selective cell death and loss of viability in COS-7 is the potent inhibition of MYC. The inhibition of MYC could be a contributing factor in the FBLN5-induced reduction of long-term cell viability in COS-7 cells (figure 6.3). In addition, the induction of cell death upon knockdown of MYC has been reported in at least two studies [264] [265].

When I started my PhD, Qize Ding was in the final year (of his PhD) and some of the research aims overlapped between our projects. However, by and large we employed different approaches: I have been interested in investigating the ability of anticancer genes in imparting a multitude of actions in the transformed and cancer cells – I view this quality as the very strength of the anticancer genes and their distinguishing feature among the tumour specific therapies. On the other hand, Qize Ding was largely engrossed in identifying and investigating the specific synthetic lethality partners for the anticancer genes. The putative synthetic lethality partner for the novel anticancer genes that he concentrated on most was cMYC [59].

Qize Ding reported FBLN5 as the synthetic lethality partner of cMYC, that is, FBLN5 overexpression in wild type CV-1 cells or CV-1 cells transfected with control plasmid did not result in cell death, but when transfected in cMYC overexpressing CV-1 cells, FBLN5 caused cell death. Moreover, Ding also demonstrated that the cell death induced by FBLN5 in MCF-7 cells was attenuated upon knockdown of cMYC. Similarly, I observed FBLN5 overexpression to cause cell death in COS-7 cells and not in CV-1 cells; I also observed that in comparison to CV-1 cells, MYC is potently (predicted to be) activated in COS-7 cells. It appears that despite largely different approaches, our results regarding the putative functioning of FBLN5 corroborated in the end; the question, however, remains that if cMYC and FBLN5 are synthetic lethality partners, why does FBLN5 overexpression in COS-7 cells result in a potent inhibition of MYC? As I stated before several times, biological systems seldom work in isolation. Higher order mechanistic network analyses could be employed to hypothesize the possible links between FBLN5, MYC and cell death. However, one could speculate that the cell cycle, G1/S or G2/M checkpoint regulation could be involved in these phenomena. To wrap up the discussion, preliminarily, it appears that potent inhibition of oxidative phosphorylation and inhibition of MYC, could, in part, play a role in the FBLN5 induced selective cell death and/or loss of viability in transformed and cancer cells.

7 FINAL DISCUSSION: summary, prospects and contribution to knowledge

Following the discovery of 22 novel anticancer genes, the subsequent studies presented in this thesis comprise several interesting and novel findings which hold exciting prospects. Four out of the 22 anticancer genes, for example, were known non-coding transcripts; the possible role of non-coding RNAs in tumour-specific cell death could provide a very interesting line of enquiry in the rapidly expanding research field of non-coding RNAs. Ten novel transcripts were also identified – the sequences of 4 of these overlapped with portions of the introns of previously known genes, another 4 overlapped the small portions of the exons of previously annotated genes while 2 appeared to exhibit some novel exon structures. Subjecting these transcripts to the nucleotide-to-protein translation tools could reveal if these are coding or non-coding transcripts. In case novel coding transcripts are identified, the subsequent hydropathy analysis of the predicted amino acid sequences would give indication of the physiological roles of these novel proteins. In addition, 2 transcripts could be novel splice variants of SEPT5 and LRRFIP2; confirmation of this finding will require detailed search among genomic databases. Ensembl genome browser is itself a very comprehensive database. There is a strong chance that if these two novel splice variants have not been annotated by Ensembl, they have not been characterized previously. Additionally, onco-suppressive roles of the 6 protein-coding and 1 non-coding anticancer genes were observed in breast cancer. Two anticancer genes were chosen for further investigation (TMEFF2 and FBLN5). The regulation and functions of these two genes were investigated in MCF-7 and COS-7 cells, to the best of my knowledge, for the very first time. Both TMEFF2 and FBLN5 were found to be positively regulated by FOXO3 in MCF-7 cells. TMEFF2 was studied in detail for its onco-suppressive role in breast cancer and a comprehensive and detailed literature review has also been included in this thesis which can provide substantial information to the prospective researchers. Furthermore, it was established that the TMEFF2 causes p53-independent apoptosis in MCF-7 breast cancer cells which along with the activation of caspases, also exhibited mitochondrial depolarization and morphological features typical of apoptosis. Through in-depth gene expression profiling, a strong interferon induction along with upregulation of TRAIL was observed upon TMEFF2 overexpression in MCF-7 cells. After relating these findings with some closely related (previously published) reports, the putative mode of TMEFF2-induced apoptosis in MCF-7 cells was proposed to be the combined action of TRAIL and interferons. As to how TMEFF2 could be

developed into an effective therapy in breast cancer, there could be several exciting prospects, but each need extensive further experimentation. Firstly, the shedded/secreted ectodomain of TMEFF2 is capable of imparting actions in the cells treated with the TMEFF2 conditioned media. There could also be a possibility of the similar in MCF-7 cells. Treatment of MCF-7 cells with the purified TMEFF2 ectodomain from conditioned media and assaying for apoptosis and gene expression of ISGs could reveal exciting results. (I intended to perform this experiment but the purification of TMEFF2 would need several optimization steps. Merely using conditioned media from TMEFF2 overexpressing cells would not be feasible as there was a strong chance of it containing TRAIL and TNFα, which in turn could induce apoptosis and interferon response on their own). In case the TMEFF2 ectodomain imparts similar effects as the TMEFF2 overexpression in MCF-7 cells, further work could be conducted along the lines of the interatumoral administration of TMEFF2 ectodomain having the therapeutic potential in breast cancer. In case the treatment with ectodomain does not result in the desired effects, the conceivable interatumoral gene therapy with TMEFF2 in breast cancer could be probed further on the lines similar to those of MDA-7/IL-24 incorporated adenoviral vector (Ad.mda-7; INGN 241) [34]. The effects of TMEFF2 overexpression in different breast cancer cell lines need to be investigated. Kumar-Sinha *et al*. [48] demonstrated the augmentation of apoptosis in MCF-7 cells and the triple-negative breast cancer xenograft cell line, SUM149, upon treatment with interferons and TRAIL and Ruiz-Ruiz *et al*. [266] reported that the IFNγ sensitized the breast cancer cells, MDA-MB231 to FAS-induced apoptosis. Together with very low expression levels of TMEFF2 in breast tissue, which appears to get even lower in breast cancer, there is a strong chance that TMEFF2 overexpression in different breast cancer cell lines would result in effects similar to the ones observed in MCF-7 cells, i.e., strong apoptosis induction. Another possible direction of investigation could be the option of interatumoral administration of a cocktail of interferons along with TRAIL in breast cancer, although the safety profile of such treatment may be difficult to optimize. In terms of TMEFF2 as an anticancer gene, using CV-1 cells as the normal cell model, COS-7 cells as the transformed cell model and MCF-7 cells as the model for cancer cells, it was established that TMEFF2 does not cause cell death in CV-1 cells but causes cell death and/or decrease in proliferation of COS-7 cells (and induces potent apoptosis in MCF-7 cells). Moreover, the primary mode behind the activity of TMEFF2 as an anticancer gene seemed to be the difference in its expression levels upon transfection in CV-1 and COS-7 cells, which in turn appeared to be due to the active downregulation of the transfected gene in CV-1 cells through some unknown mechanisms. The effects of TMEFF2 overexpression in various normal cell models need to be studied; in the case that TMEFF2 does not have any toxic effects in a

number of normal cells, the prospects of the intertumoral gene therapy with TMEFF2 discussed above become even more promising.

Similarly, after observing that FBLN5 induces cell death and/or reduction in proliferation of COS-7 cells and not in CV-1 cells, an in-depth gene expression profiling was performed on FBLN5 overexpressing CV-1 and COS-7 cells. It was followed by an innovative comparison-ontology analyses. The results revealed some interesting putative mechanisms behind the COS-7 cells specific cell death upon FBLN5 overexpression. These included inhibition of oxidative phosphorylation and selective downregulation of MYC-dependent genes in COS-7 cells. It was for the first time that the role of FBLN5 in oxidative phosphorylation was reported (to the best of my knowledge). A future line of enquiry into the eventual use of FBLN5 as an anticancer gene would include checking the effects of FBLN5 overexpression in various normal, cancer and transformed cell lines. Moreover, the higher order mechanistic pathways analyses proposed in chapter 6 and the subsequent functional studies could reveal some very promising synthetic lethal targets for developing novel tumour-specific therapies. Another interesting conceivable line of enquiry regarding the FBLN5-induced cell death would be checking the effects of secreted or recombinant FBLN5 treatment on transformed and/or cancer cells. In case the secreted FBLN5 can impart cancer specific cell death, opportunities of developing this secreted glycoprotein into a tumour-specific therapeutic agent would become even more exciting.

Lastly, I have only been able to use a fraction of data generated through the gene expression profiling experiments. The complete set of data will be made available to the public through both Gene Expression Omnibus (GEO) and ArrayExpress. The differentially expressed novel transcripts, non-coding transcripts and novel splice variants could be very useful for further investigation regarding the functions of TMEFF2 and FBLN5. Especially, the vast lists of differentially expressed non-coding and novel transcripts observed upon TMEFF2 overexpression could be of particular interest, not only for their involvement in the functioning of TMEFF2, but also researchers would be able to capitalize on this information for future studies in the exciting and expanding field of interferon signalling. Similarly, the data regarding the differentially expressed splice variant, non-coding and novel transcripts between CV-1 and COS-7 cells could be very valuable in deciphering the neoplastic transformation in general.

8 References

1. Weinberg RA. *The Biology of Cancer*. 2nd ed. New York: Garland Science, Taylor & Francis Group; 2014.

2. Cho H, Mariotto AB, Schwartz LM, Luo J, Woloshin S. When do changes in cancer survival mean progress? The insight from population incidence and mortality. *Journal of the National Cancer Institute. Monographs*. 2014;2014(49):187-197. doi:10.1093/jncimonographs/lgu014.

3. PARASCANDOLA J. The Theoretical Basis of Paul Ehrlich's Chemotherapy. *Journal of the History of Medicine and Allied Sciences*. 1981;XXXVI(1):19-43. doi:10.1093/jhmas/XXXVI.1.19.

4. Coppoc GL. Chemotherapy of Neoplastic Diseases. In: Jim E. Riviere, Mark G. Papich, editors. *Veterinary pharmacology and therapeutics*. 9 ed. Iowa State University Press: John Wiley & Sons; 2013. p.1205-1249.

5. Wallington M, Saxon EB, Bomb M, Smittenaar R, Wickenden M, McPhail S*, et al.* 30-day mortality after systemic anticancer treatment for breast and lung cancer in England: a population-based, observational study. *The Lancet. Oncology*. 2016;17(9):1203-1216. doi:10.1016/s1470-2045(16)30383-7.

6. Grimm S, Noteborn M. Anticancer genes: inducers of tumour-specific cell death signalling. *Trends in Molecular Medicine*. 2010;16(2):88-96. doi:10.1016/j.molmed.2009.12.002.

7. Grimm S. Introduction. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.1-8. doi:10.1007/978-1-4471-6458-6_1.

8. Morgan G, Ward R, Barton M. The contribution of cytotoxic chemotherapy to 5-year survival in adult malignancies. *Clinical Oncology*. 2004;16(8):549-560. doi:10.1016/j.clon.2004.06.007.

9. Chabner BA, Roberts TG. Chemotherapy and the war on cancer. *Nature Reviews Cancer*. 2005;5(1):65-72. doi:10.1038/nrc1529.

10. Druker BJ, Tamura S, Buchdunger E, Ohno S, Segal GM, Fanning S*, et al.* Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nature Medicine*. 1996;2(5):561-566. doi:doi:10.1038/nm0596-561.

11. Huang M, Shen A, Ding J, Geng M. Molecularly targeted cancer therapy: some lessons from the past decade. *Trends in Pharmacological Sciences*. 2014;35(1):41-50. doi:10.1016/j.tips.2013.11.004.

12. Abuali G, Chaisaklert W, Stelloo E, Pazarentzos E, Hwang MS, Qize D*, et al.* The anticancer gene ORCTL3 targets stearoyl-CoA desaturase-1 for tumour-specific apoptosis. *Oncogene*. 2015;34(13):1718–1728. doi:10.1038/onc.2014.93.

13. Bullenkamp J, Tavassoli M. Signalling of Apoptin. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. 1 ed. London: Springer London; 2014. p.11-37. doi:10.1007/978-1-4471-6458-6_2.

14. Roninson IB, Chin JE, Choi KG, Gros P, Housman DE, Fojo A*, et al.* Isolation of human mdr DNA sequences amplified in multidrug-resistant KB carcinoma cells. *Proceedings of the National Academy of Sciences*. 1986;83(12):4538-4542. doi:10.1073/pnas.83.12.4538.

15. Garcia-Cao I, Duran A, Collado M, Carrascosa MJ, Martin-Caballero J, Flores JM*, et al.* Tumour-suppression activity of the proapoptotic regulator Par4. *EMBO Reports*. 2005;6(6):577-583. doi:10.1038/sj.embor.7400421.

16. Stempien-Otero A, Karsan A, Cornejo CJ, Xiang H, Eunson T, Morrison RS*, et al.* Mechanisms of Hypoxia-induced Endothelial Cell Death: ROLE OF p53 IN APOPTOSIS. *Journal of Biological Chemistry*. 1999;274(12):8039-8045. doi:10.1074/jbc.274.12.8039.

17. AbuAli G, Grimm S. Isolation and Characterization of the Anticancer Gene Organic Cation Transporter Like-3 (ORCTL3). In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.213-227. doi:10.1007/978-1-4471-6458-6_11.

18. Geoffrey M. Cooper REH. *The Cell: A Molecular Approach*. 4th ed. Washington, D.C: Sinauer Associates; 2007.

19. Elmore S. Apoptosis: a review of programmed cell death. *Toxicologic Pathology*. 2007;35(4):495-516. doi:10.1080/01926230701320337.

20. Schweichel J-U, Merker H-J. The morphology of various types of cell death in prenatal tissues. *Teratology*. 1973;7(3):253-266. doi:10.1002/tera.1420070306.

21. Kroemer G, Galluzzi L, Vandenabeele P, Abrams J, Alnemri ES, Baehrecke EH*, et al.* Classification of cell death: recommendations of the Nomenclature Committee on Cell Death 2009. *Cell Death & Differentiation*. 2009;16(1):3-11. doi:10.1038/cdd.2008.150.

22. Galluzzi L, Vitale I, Abrams JM, Alnemri ES, Baehrecke EH, Blagosklonny MV*, et al.* Molecular definitions of cell death subroutines: recommendations of the Nomenclature Committee on Cell Death 2012. *Cell Death & Differentiation*. 2012;19(1):107-120. doi:10.1038/cdd.2011.96.

23. Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P*, et al.* Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death & Differentiation*. 2018;25(3):486-541. doi:10.1038/s41418-017-0012-4.

24. Ting AT, Bertrand MJM. More to Life than NF-κB in TNFR1 Signaling. *Trends in Immunology*. 2016;37(8):535-545. doi:10.1016/j.it.2016.06.002.

25. Danen-Van Oorschot AA, Fischer DF, Grimbergen JM, Klein B, Zhuang S, Falkenburg JH*, et al.* Apoptin induces apoptosis in human transformed and malignant cells but not in normal cells. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;94(11):5843-5847. doi:10.1073/pnas.94.11.5843.

26. Lanz HL, Zimmerman RME, Brouwer J, Noteborn MHM, Backendorf C. Mitotic catastrophe triggered in human cancer cells by the viral protein apoptin. *Cell Death & Disease*. 2013;4(2):e487-e487. doi:10.1038/cddis.2013.2.

27. Bai H, Lester GMS, Petishnok LC, Dean DA. Cytoplasmic transport and nuclear import of plasmid DNA. *Bioscience reports*. 2017;37(6):BSR20160616. doi:10.1042/BSR20160616.

28. Guelen L, Paterson H, Gäken J, Meyers M, Farzaneh F, Tavassoli M. TAT-apoptin is efficiently delivered and induces apoptosis in cancer cells. *Oncogene*. 2004;23(5):1153-1165. doi:10.1038/sj.onc.1207224.

29. Kleinberger T. Induction of Cancer-Specific Cell Death by the Adenovirus E4orf4 Protein. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.61-97. doi:10.1007/978-1-4471-6458-6_4.

30. Shtrichman R, Sharf R, Barr H, Dobner T, Kleinberger T. Induction of apoptosis by adenovirus E4orf4 protein is specific to transformed cells and requires an interaction with protein phosphatase 2A. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;96(18):10080-10085. doi:10.1073/pnas.96.18.10080.

31. Rommelaere J, Cornelis JJ. Antineoplastic activity of parvoviruses. *Journal of Virological Methods*. 1991;33(3):233-251. doi:10.1016/0166-0934(91)90024-T.

32. Nüesch JPF, Rommelaere J. Tumor Suppressing Properties of Rodent Parvovirus NS1 Proteins and Their Derivatives. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.99-124. doi:10.1007/978-1-4471-6458-6_5.

33. Geletneky K, Hajda J, Angelova AL, Leuchs B, Capper D, Bartsch AJ*, et al.* Oncolytic H-1 Parvovirus Shows Safety and Signs of Immunogenic Activity in a First Phase I/IIa Glioblastoma Trial. *Molecular Therapy*. 2017;25(12):2620-2634. doi:10.1016/j.ymthe.2017.08.016.

34. Menezes ME, Bhatia S, Bhoopathi P, Das SK, Emdad L, Dasgupta S*, et al.* MDA-7/IL-24: Multifunctional Cancer Killing Cytokine. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.127-153. doi:10.1007/978-1-4471-6458-6_6.

35. Wang CJ, Zhang H, Chen K, Zheng JW, Xiao CW, Ji WW*, et al.* Ad.mda-7 (IL-24) selectively induces apoptosis in hepatocellular carcinoma cell lines, suppresses metastasis, and enhances the effect of doxorubicin on xenograft tumors. *Oncology Research*. 2010;18(11-12):561-574.

36. Huo W, Li ZM, Zhu XM, Bao YM, An LJ. MDA-7/IL-24 suppresses tumor adhesion and invasive potential in hepatocellular carcinoma cell lines. *Oncology Reports*. 2013;30(2):986-992. doi:10.3892/or.2013.2507.

37. Lebedeva IV, Su ZZ, Sarkar D, Kitada S, Dent P, Waxman S*, et al.* Melanoma differentiation associated gene-7, mda-7/interleukin-24, induces apoptosis in prostate cancer cells by promoting mitochondrial dysfunction and inducing reactive oxygen species. *Cancer Research*. 2003;63(23):8138-8144.

38. Hebbar N, Shrestha-Bhattarai T, Rangnekar VM. Cancer-Selective Apoptosis by Tumor Suppressor Par-4. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.155-166. doi:10.1007/978-1-4471-6458-6_7.

39. Chakraborty M, Qiu SG, Vasudevan KM, Rangnekar VM. Par-4 drives trafficking and activation of Fas and Fasl to induce prostate cancer cell apoptosis and tumor regression. *Cancer Research*. 2001;61(19):7255-7263.

40. El-Guendy N, Zhao Y, Gurumurthy S, Burikhanov R, Rangnekar VM. Identification of a unique core domain of par-4 sufficient for selective apoptosis induction in cancer cells. *Molecular and Cellular Biology*. 2003;23(16):5516-5525. doi:10.1128/mcb.23.16.5516-5525.2003.

41. Zhao Y, Burikhanov R, Qiu S, Lele SM, Jennings CD, Bondada S*, et al.* Cancer resistance in transgenic mice expressing the SAC module of Par-4. *Cancer Research*. 2007;67(19):9276-9285. doi:10.1158/0008-5472.Can-07-2124.

42. Irshad S, Mahul-Mellier AL, Kassouf N, Lemarie A, Grimm S. Isolation of ORCTL3 in a novel genetic screen for tumor-specific apoptosis inducers. *Cell Death & Differentiation*. 2009;16(6):890-898. doi:10.1038/cdd.2009.21.

43. Iwasawa R, Mahul-Mellier A-L, Datler C, Pazarentzos E, Grimm S. Fis1 and Bap31 bridge the mitochondria-ER interface to establish a platform for apoptosis induction. *The EMBO Journal*. 2011;30(3):556-568. doi:10.1038/emboj.2010.346.

44. von Karstedt S, Montinaro A, Walczak H. Exploring the TRAILs less travelled: TRAIL in cancer biology and therapy. *Nature Reviews Cancer*. 2017;17:352. doi:10.1038/nrc.2017.28.

45. Mérino D, Lalaoui N, Morizot A, Solary E, Micheau O. TRAIL in cancer therapy: present and future challenges. *Expert Opinion on Therapeutic Targets*. 2007;11(10):1299-1314. doi:10.1517/14728222.11.10.1299.

46. Haselmann V, Kurz A, Bertsch U, Hübner S, Olempska–Müller M, Fritsch J*, et al.* Nuclear Death Receptor TRAIL-R2 Inhibits Maturation of Let-7 and Promotes Proliferation of Pancreatic and Other Tumor Cells. *Gastroenterology*. 2014;146(1):278-290. doi:10.1053/j.gastro.2013.10.009.

47. Bertsch U, Röder C, Kalthoff H, Trauzold A. Compartmentalization of TNF-related apoptosis-inducing ligand (TRAIL) death receptor functions: emerging role of nuclear TRAIL-R2. *Cell Death & Disease*. 2014;5:e1390. doi:10.1038/cddis.2014.351.

48. Kumar-Sinha C, Varambally S, Sreekumar A, Chinnaiyan AM. Molecular cross-talk between the TRAIL and interferon signaling pathways. *Journal of Biological Chemistry*. 2002;277(1):575-585. doi:10.1074/jbc.M107795200.

49. Svensson M, Hakansson A, Mossberg AK, Linse S, Svanborg C. Conversion of alpha-lactalbumin to a protein inducing apoptosis. *Proceedings of the National Academy of Sciences of the United States of America*. 2000;97(8):4221-4226. doi:10.1073/pnas.97.8.4221.

50. Gustafsson L, Aits S, Önnerfjord P, Trulsson M, Storm P, Svanborg C. Changes in Proteasome Structure and Function Caused by HAMLET in Tumor Cells. *PLoS ONE*. 2009;4(4):e5229. doi:10.1371/journal.pone.0005229.

51. Hallgren O, Aits S, Brest P, Gustafsson L, Mossberg AK, Wullt B*, et al.* Apoptosis and tumor cell death in response to HAMLET (human alpha-lactalbumin made lethal to tumor cells). *Advances in Experimental Medicine and Biology*. 2008;606:217-240. doi:10.1007/978-0-387-74087-4_8.

52. Gustafsson L, Leijonhufvud I, Aronsson A, Mossberg A-K, Svanborg C. Treatment of Skin Papillomas with Topical α-Lactalbumin–Oleic Acid. *New England Journal of Medicine*. 2004;350(26):2663-2672. doi:10.1056/NEJMoa032454.

53. Puthia M, Storm P, Nadeem A, Hsiung S, Svanborg C. Prevention and treatment of colon cancer by peroral administration of HAMLET (human α-lactalbumin made lethal to tumour cells). *Gut*. 2014;63(1):131-142. doi:10.1136/gutjnl-2012-303715.

54. Pazarentzos E, Mazarakis ND. Anticancer Gene Transfer for Cancer Gene Therapy. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.255-280. doi:10.1007/978-1-4471-6458-6_13.

55. Rizzuti M, Nizzardo M, Zanetta C, Ramirez A, Corti S. Therapeutic applications of the cell-penetrating HIV-1 Tat peptide. *Drug Discovery Today*. 2015;20(1):76-85. doi:10.1016/j.drudis.2014.09.017.

56. Kalli C, Teoh WC, Leen E. Introduction of Genes via Sonoporation and Electroporation. In: Stefan Grimm, editor. Anticancer Genes. *Advances in Experimental Medicine and Biology*. London: Springer London; 2014. p.231-254. doi:10.1007/978-1-4471-6458-6_12.

57. Grimm S. The art and design of genetic screens: mammalian culture cells. *Nature Reviews Genetics*. 2004;5(3):179-189. doi:10.1038/nrg1291.

58. AbuAli G. *ORCTL3: Insights into the Function and Structure of a Novel Anticancer Gene*. PhD thesis. Imperial College London; 2013.

59. Ding Q. *A Genetic Screening for the Tumour Suppressive Anticancer Genes*. PhD thesis. Imperial College London; 2016.

60. Bevan LKC. *Using the RISCI Genetic Screening Platform for Elucidating Apoptosis Signalling Network*. PhD thesis. Imperial College London; 2012.

61. Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, Irie R*, et al.* Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nature Genetics*. 2004;36(1):40-45. doi:10.1038/ng1285.

62. Brunet A, Bonni A, Zigmond MJ, Lin MZ, Juo P, Hu LS*, et al.* Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell*. 1999;96(6):857-868.

63. Fernandez de Mattos S, Essafi A, Soeiro I, Pietersen AM, Birkenkamp KU, Edwards CS*, et al.* FoxO3a and BCR-ABL regulate cyclin D2 transcription through a STAT5/BCL6-dependent mechanism. *Molecular and Cellular Biology*. 2004;24(22):10058-10071. doi:10.1128/mcb.24.22.10058-10071.2004.

64. Nestal de Moraes G, Delbue D, Silva KL, Robaina MC, Khongkow P, Gomes AR*, et al.* FOXM1 targets XIAP and Survivin to modulate breast cancer survival and chemoresistance. *Cellular Signalling*. 2015;27(12):2496-2505. doi:10.1016/j.cellsig.2015.09.013.

65. Nestal de Moraes G, Khongkow P, Gong C, Yao S, Gomes AR, Ji Z*, et al.* Forkhead box K2 modulates epirubicin and paclitaxel sensitivity through FOXO3a in breast cancer. *Oncogenesis*. 2015;4:e167. doi:10.1038/oncsis.2015.26.

66. Mund T, Gewies A, Schoenfeld N, Bauer MK, Grimm S. Spike, a novel BH3-only protein, regulates apoptosis at the endoplasmic reticulum. *FEBS Journal*. 2003;17(6):696-698. doi:10.1096/fj.02-0657fje.

67. Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*. 2000;7(1-2):203-214. doi:10.1089/10665270050081478.

68. Madden T. The BLAST Sequence Analysis Tool. 2002 In: Jo McEntyre, editor. The NCBI Handbook. Internet: Bethesda (MD): National Center for Biotechnology Information (US); [Chapter 16]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK21097/.

69. Kent WJ. BLAT--the BLAST-like alignment tool. *Genome Research*. 2002;12(4):656-664. doi:10.1101/gr.229202.

70. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S*, et al.* Ensembl 2015. *Nucleic Acids Research*. 2015;43(Database issue):D662-669. doi:10.1093/nar/gku1010.

71. *Kaplan-Meier Plotter*. Available from: http://kmplot.com/analysis/. [Accessed 16 August 2016]

72. Nagy A, Lanczky A, Menyhart O, Gyorffy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific Reports*. 2018;8(1):9227. doi:10.1038/s41598-018-27521-y.

73. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q*, et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Research and Treatment*. 2010;123(3):725-731. doi:10.1007/s10549-009-0674-9.

74. Gyorffy B, Lanczky A, Szallasi Z. Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocrine-Related Cancer*. 2012;19(2):197-208. doi:10.1530/erc-11-0329.

75. Gyorffy B, Surowiak P, Budczies J, Lanczky A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS ONE*. 2013;8(12):e82241. doi:10.1371/journal.pone.0082241.

76. Szasz AM, Lanczky A, Nagy A, Forster S, Hark K, Green JE*, et al.* Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. *Oncotarget*. 2016;7(31):49322-49333. doi:10.18632/oncotarget.10337.

77. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A*, et al.* Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347(6220):1260419. doi:10.1126/science.1260419.

78. Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H*, et al.* A subcellular map of the human proteome. *Science*. 2017;356(6340). doi:10.1126/science.aal3321.

79. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G*, et al.* A pathology atlas of the human cancer transcriptome. *Science*. 2017;357(6352). doi:10.1126/science.aan2507.

80. *Kaplan - Meier Plotter Pan-cancer RNA-seq*. Available from: http://kmplot.com/analysis/index.php?p=service&cancer=pancancer_rnaseq. [Accessed 10 January 2019]

81. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970;227(5259):680-685. doi:10.1038/227680a0.

82. Applied Biosystems. *Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR*. 4371095 Rev B. 2008. Available from: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/cms_042380.pdf.

83. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T*, et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 2012;9(7):676-682. doi:10.1038/nmeth.2019.

84. Applied Biosystems. *Transcriptome Analysis Console (TAC) 4.0.2 USER GUIDE*. 8 ed. ThermoFisher Scientific; 2019. Available from: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/tac_user_manual.pdf.

85. Slenter DN, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N*, et al.* WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Research*. 2018;46(D1):D661-D667. doi:10.1093/nar/gkx1064.

86. Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics*. 2014;30(4):523-530. doi:10.1093/bioinformatics/btt703.

87. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Research*. 2009;37(Database issue):D26-31. doi:10.1093/nar/gkn723.

88. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. 1977;74(12):5463-5467. doi:10.1073/pnas.74.12.5463.

89. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. *International Journal of Ayurveda Research*. 2010;1(4):274-278. doi:10.4103/0974-7788.76794.

90. Kumar V, Abbas AK, Aster JC. *Robbins and Cotran pathologic basis of disease*. 9th ed. Philadelphia, PA: Elsevier/Saunders; 2015.

91. Comsa S, Cimpean AM, Raica M. The Story of MCF-7 Breast Cancer Cell Line: 40 years of Experience in Research. *Anticancer Research*. 2015;35(6):3147-3154.

92. Riccardi C, Nicoletti I. Analysis of apoptosis by propidium iodide staining and flow cytometry. *Nat Protoc*. 2006;1(3):1458-1461. doi:10.1038/nprot.2006.238.

93. Pozarowski P, Grabarek J, Darzynkiewicz Z. Flow cytometry of apoptosis. *Curr Protoc Cell Biol*. 2004;Chapter 18:Unit 18 18. doi:10.1002/0471143030.cb1808s21.

94. Gerlier D, Thomasset N. Use of MTT colorimetric assay to measure cell activation. *Journal of Immunological Methods*. 1986;94(1–2):57-63. doi:10.1016/0022-1759(86)90215-2.

95. Macfarlane L-A, Murphy PR. MicroRNA: Biogenesis, Function and Role in Cancer. *Current genomics*. 2010;11(7):537-561. doi:10.2174/138920210793175895.

96. Morris KV, Mattick JS. The rise of regulatory RNA. *Nature Reviews Genetics*. 2014;15:423-437. doi:10.1038/nrg3722.

97. Midgley AC, Oltean S, Hascall V, Woods EL, Steadman R, Phillips AO*, et al.* Nuclear hyaluronidase 2 drives alternative splicing of CD44 pre-mRNA to determine profibrotic or antifibrotic cell phenotype. *Science Signaling*. 2017;10(506):eaao1822. doi:10.1126/scisignal.aao1822.

98. Muggenthaler MMA, Chowdhury B, Hasan SN, Cross HE, Mark B, Harlalka GV*, et al.* Mutations in HYAL2, Encoding Hyaluronidase 2, Cause a Syndrome of Orofacial Clefting and Cor Triatriatum Sinister in Humans and Mice. *PLoS Genetics*. 2017;13(1):e1006470-e1006470. doi:10.1371/journal.pgen.1006470.

99. Das S, Chattopadhyay R, Bhakat KK, Boldogh I, Kohno K, Prasad R*, et al.* Stimulation of NEIL2-mediated oxidized base excision repair via YB-1 interaction during oxidative stress. *Journal of Biological Chemistry*. 2007;282(39):28474-28484. doi:10.1074/jbc.M704672200.

100. Albig AR, Schiemann WP. Fibulin-5 function during tumorigenesis. *Future Oncology*. 2005;1(1):23-35. doi:10.1517/14796694.1.1.23.

101. Schaefer L, Iozzo RV. Biological functions of the small leucine-rich proteoglycans: from genetics to signal transduction. *The Journal of Biological Chemistry*. 2008;283(31):21305-21309. doi:10.1074/jbc.R800020200.

102. Liang P, Wan Y, Yan Y, Wang Y, Luo N, Deng Y*, et al.* MVP interacts with YPEL4 and inhibits YPEL4-mediated activities of the ERK signal pathway. *Biochemistry and Cell Biology*. 2010;88(3):445-450. doi:10.1139/o09-166.

103. Sun T, Du W, Xiong H, Yu Y, Weng Y, Ren L*, et al.* TMEFF2 deregulation contributes to gastric carcinogenesis and indicates poor survival outcome. *Clinical Cancer Research*. 2014;20(17):4689-4704. doi:10.1158/1078-0432.Ccr-14-0315.

104. Lam EW, Brosens JJ, Gomes AR, Koo CY. Forkhead box proteins: tuning forks for transcriptional harmony. *Nature Reviews Cancer*. 2013;13(7):482-495. doi:10.1038/nrc3539.

105. Weigel D, Jürgens G, Küttner F, Seifert E, Jäckle 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):645-658. doi:10.1016/0092-8674(89)90133-5.

106. Bella L, Zona S, Nestal de Moraes G, Lam EW. FOXM1: A key oncofoetal transcription factor in health and disease. *Seminars in Cancer Biology*. 2014;29:32-39. doi:10.1016/j.semcancer.2014.07.008.

107. Minamino T, Komuro I. Regeneration of the endothelium as a novel therapeutic strategy for acute lung injury. *Journal of Clinical Investigation*. 2006;116(9):2316-2319. doi:10.1172/JCI29637.

108. Koo CY, Muir KW, Lam EW. FOXM1: From cancer initiation to progression and treatment. *Biochimica et Biophysica Acta*. 2012;1819(1):28-37. doi:10.1016/j.bbagrm.2011.09.004.

109. Wang Z, Ahmad A, Li Y, Banerjee S, Kong D, Sarkar FH. Forkhead box M1 transcription factor: a novel target for cancer therapy. *Cancer Treatment Reviews*. 2010;36(2):151-156. doi:10.1016/j.ctrv.2009.11.006.

110. 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):1845-1858. doi:10.1038/onc.2011.368.

111. Zona S, Bella L, Burton MJ, Nestal de Moraes G, Lam EW. FOXM1: an emerging master regulator of DNA damage response and genotoxic agent resistance. *Biochimica et Biophysica Acta*. 2014;1839(11):1316-1322. doi:10.1016/j.bbagrm.2014.09.016.

112. Katayama K, Nakamura A, Sugimoto Y, Tsuruo T, Fujita N. FOXO transcription factor-dependent p15(INK4b) and p19(INK4d) expression. *Oncogene*. 2008;27(12):1677-1686. doi:10.1038/sj.onc.1210813.

113. de Keizer PL, Packer LM, Szypowska AA, Riedl-Polderman PE, van den Broek NJ, de Bruin A*, et al.* Activation of forkhead box O transcription factors by oncogenic BRAF promotes p21cip1-

dependent senescence. *Cancer Research*. 2010;70(21):8526-8536. doi:10.1158/0008-5472.can-10-1563.

114. Dijkers PF, Medema RH, Pals C, Banerji L, Thomas NS, Lam EW*, et al.* Forkhead transcription factor FKHR-L1 modulates cytokine-dependent transcriptional regulation of p27(KIP1). *Molecular and Cellular Biology*. 2000;20(24):9138-9148.

115. Medema RH, Kops GJ, Bos JL, Burgering BM. AFX-like Forkhead transcription factors mediate cell-cycle regulation by Ras and PKB through p27kip1. *Nature*. 2000;404(6779):782-787. doi:10.1038/35008115.

116. Tran H, Brunet A, Grenier JM, Datta SR, Fornace AJ, Jr., DiStefano PS*, et al.* DNA repair pathway stimulated by the forkhead transcription factor FOXO3a through the Gadd45 protein. *Science*. 2002;296(5567):530-534. doi:10.1126/science.1068712.

117. Fu G, Peng C. Nodal enhances the activity of FoxO3a and its synergistic interaction with Smads to regulate cyclin G2 transcription in ovarian cancer cells. *Oncogene*. 2011;30(37):3953-3966. doi:10.1038/onc.2011.127.

118. Essafi A, Fernandez de Mattos S, Hassen YA, Soeiro I, Mufti GJ, Thomas NS*, et al.* Direct transcriptional regulation of Bim by FoxO3a mediates STI571-induced apoptosis in Bcr-Abl-expressing cells. *Oncogene*. 2005;24(14):2317-2329. doi:10.1038/sj.onc.1208421.

119. Calnan DR, Brunet A. The FoxO code. *Oncogene*. 2008;27(16):2276-2288. doi:10.1038/onc.2008.21.

120. Ho KK, McGuire VA, Koo CY, Muir KW, de Olano N, Maifoshie E*, et al.* Phosphorylation of FOXO3a on Ser-7 by p38 promotes its nuclear localization in response to doxorubicin. *Journal of Biological Chemistry*. 2012;287(2):1545-1555. doi:10.1074/jbc.M111.284224.

121. Kutmon M, Riutta A, Nunes N, Hanspers K, Willighagen EL, Bohler A*, et al.* WikiPathways: capturing the full diversity of pathway knowledge. *Nucleic Acids Research*. 2016;44(D1):D488-494. doi:10.1093/nar/gkv1024.

122. Ji Z, Mohammed H, Webber A, Ridsdale J, Han N, Carroll JS*, et al.* The forkhead transcription factor FOXK2 acts as a chromatin targeting factor for the BAP1-containing histone deubiquitinase complex. *Nucleic Acids Research*. 2014;42(10):6232-6242. doi:10.1093/nar/gku274.

123. Bowman CJ, Ayer DE, Dynlacht BD. Foxk proteins repress the initiation of starvation-induced atrophy and autophagy programs. *Nature Cell Biology*. 2014;16(12):1202-1214. doi:10.1038/ncb3062.

124. Krol J, Francis RE, Albergaria A, Sunters A, Polychronis A, Coombes RC*, et al.* The transcription factor FOXO3a is a crucial cellular target of gefitinib (Iressa) in breast cancer cells. *Molecular Cancer Therapeutics*. 2007;6(12):3169-3179. doi:10.1158/1535-7163.mct-07-0507.

125. Marina M, Wang L, Conrad SE. The scaffold protein MEK Partner 1 is required for the survival of estrogen receptor positive breast cancer cells. *Cell Communication and Signaling*. 2012;10(1):18. doi:10.1186/1478-811x-10-18.

126. Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *Journal of General Virology*. 1977;36(1):59-72. doi:10.1099/0022-1317-36-1-59.

127. Pear WS, Nolan GP, Scott ML, Baltimore D. Production of high-titer helper-free retroviruses by transient transfection. *Proceedings of the National Academy of Sciences of the United States of America*. 1993;90(18):8392-8396. doi:10.1073/pnas.90.18.8392.

128. Gluzman Y. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*. 1981;23(1):175-182. doi:10.1016/0092-8674(81)90282-8.

129. Gu M, Meng K, Majerus PW. The effect of overexpression of the protein tyrosine phosphatase PTPMEG on cell growth and on colony formation in soft agar in COS-7 cells. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;93(23):12980-12985. doi:10.1073/pnas.93.23.12980.

130. *HGNC Database,HUGO Gene Nomenclature Committee (HGNC), European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI),Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, United Kingdom*. Available from: https://www.genenames.org/. [Accessed 17 May 2019]

131. Yates B, Braschi B, Gray KA, Seal RL, Tweedie S, Bruford EA. Genenames.org: the HGNC and VGNC resources in 2017. *Nucleic Acids Research*. 2017;45(D1):D619-d625. doi:10.1093/nar/gkw1033.

132. Horie M, Mitsumoto Y, Kyushiki H, Kanemoto N, Watanabe A, Taniguchi Y*, et al.* Identification and Characterization of TMEFF2, a Novel Survival Factor for Hippocampal and Mesencephalic Neurons. *Genomics*. 2000;67(2):146-152. doi:10.1006/geno.2000.6228.

133. Ali N, Knaüper V. Phorbol Ester-induced Shedding of the Prostate Cancer Marker Transmembrane Protein with Epidermal Growth Factor and Two Follistatin Motifs 2 Is Mediated by the Disintegrin and Metalloproteinase-17. *Journal of Biological Chemistry*. 2007;282(52):37378-37388. doi:10.1074/jbc.M702170200.

134. Labeur M, Wolfel B, Stalla J, Stalla GK. TMEFF2 is an endogenous inhibitor of the CRH signal transduction pathway. *Journal of Molecular Endocrinology*. 2015;54(1):51-63. doi:10.1530/jme-14-0225.

135. Hong HS, Maezawa I, Petrlova J, Zhao XY, J CV, Jin LW. Tomoregulin (TMEFF2) Binds Alzheimer's Disease Amyloid-beta (Abeta) Oligomer and AbetaPP and Protects Neurons from Abeta-Induced Toxicity. *Journal of Alzheimer's Disease*. 2015;48(3):731-743. doi:10.3233/jad-150318.

136. Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. DOG 1.0: illustrator of protein domain structures. *Cell Research*. 2009;19(2):271-273. doi:10.1038/cr.2009.6.

137. Liang G, Robertson KD, Talmadge C, Sumegi J, Jones PA. The Gene for a Novel Transmembrane Protein Containing Epidermal Growth Factor and Follistatin Domains Is Frequently Hypermethylated in Human Tumor Cells. *Cancer Research*. 2000;60(17):4907-4912.

138. Young J, Biden KG, Simms LA, Huggard P, Karamatic R, Eyre HJ*, et al.* HPP1: a transmembrane protein-encoding gene commonly methylated in colorectal polyps and cancers. *Proceedings of the National Academy of Sciences of the United States of America*. 2001;98(1):265-270. doi:10.1073/pnas.011415298.

139. Yonezawa M, Uchida T, Wada K, Akamatsu T, Mizoguchi A, Sakamoto C*, et al.* A novel epidermal growth factor-like molecule containing two follistatin modules stimulates tyrosine phosphorylation of ERBB-4 in MKN28 gastric cancer cells. *Gastroenterology*. 2000;118(4, Part 1):A557. doi:10.1016/S0016-5085(00)84361-9.

140. Glynne-Jones E, Harper ME, Seery LT, James R, Anglin I, Morgan HE*, et al.* TENB2, a proteoglycan identified in prostate cancer that is associated with disease progression and androgen independence. *International Journal of Cancer*. 2001;94(2):178-184.

141. Quayle SN, Sadar MD. A truncated isoform of TMEFF2 encodes a secreted protein in prostate cancer cells. *Genomics*. 2006;87(5):633-637. doi:10.1016/j.ygeno.2005.12.004.

142. Uchida T, Wada K, Akamatsu T, Yonezawa M, Noguchi H, Mizoguchi A*, et al.* A Novel Epidermal Growth Factor-like Molecule Containing Two Follistatin Modules Stimulates Tyrosine Phosphorylation of erbB-4 in MKN28 Gastric Cancer Cells. *Biochemical and Biophysical Research Communications*. 1999;266(2):593-602. doi:10.1006/bbrc.1999.1873.

143. Zhao X-Y, Schneider D, Biroc SL, Parry R, Alicke B, Toy P*, et al.* Targeting Tomoregulin for Radioimmunotherapy of Prostate Cancer. *Cancer Research*. 2005;65(7):2846-2853. doi:10.1158/0008-5472.CAN-04-4019.

144. Liang G, Salem CE, Yu MC, Nguyen HD, Gonzales FA, Nguyen TT*, et al.* DNA Methylation Differences Associated with Tumor Tissues Identified by Genome Scanning Analysis. *Genomics*. 1998;53(3):260-268. doi:10.1006/geno.1998.5502.

145. Widschwendter M, Jones PA. DNA methylation and breast carcinogenesis. *Oncogene*. 2002;21(35):5462-5482. doi:10.1038/sj.onc.1205606.

146. Sato F, Shibata D, Harpaz N, Xu Y, Yin J, Mori Y*, et al.* Aberrant methylation of the HPP1 gene in ulcerative colitis-associated colorectal carcinoma. *Cancer Res*. 2002;62(23):6820-6822.

147. Lee SM, Park JY, Kim DS. Methylation of TMEFF2 gene in tissue and serum DNA from patients with non-small cell lung cancer. *Molecules and Cells*. 2012;34(2):171-176. doi:10.1007/s10059-012-0083-5.

148. Young J, Barker M, Fraser L, Walsh MD, Spring K, Biden KG*, et al.* Mutation searching in colorectal cancer studies: experience with a denaturing high-pressure liquid chromatography system for exon-by-exon scanning of tumour suppressor genes. *Pathology*. 2002;34(6):529-533. doi:10.1080/0031302021000035965-1.

149. Fackler MJ, Umbricht CB, Williams D, Argani P, Cruz LA, Merino VF*, et al.* Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of recurrence. *Cancer Research*. 2011;71(19):6195-6207. doi:10.1158/0008-5472.Can-11-1630.

150. Fackler MJ, Lopez Bujanda Z, Umbricht C, Teo WW, Cho S, Zhang Z*, et al.* Novel methylated biomarkers and a robust assay to detect circulating tumor DNA in metastatic breast cancer. *Cancer Research*. 2014;74(8):2160-2170. doi:10.1158/0008-5472.CAN-13-3392.

151. Park SY, Kwon HJ, Lee HE, Ryu HS, Kim S-W, Kim JH*, et al.* Promoter CpG island hypermethylation during breast cancer progression. *Virchows Archiv*. 2011;458(1):73-84. doi:10.1007/s00428-010-1013-6.

152. de Groot JS, Pan X, Meeldijk J, van der Wall E, van Diest PJ, Moelans CBJCO. Validation of DNA promoter hypermethylation biomarkers in breast cancer — a short report. *Cellular Oncology*. 2014;37(4):297-303. doi:10.1007/s13402-014-0189-1.

153. Suzuki M, Shigematsu H, Shames DS, Sunaga N, Takahashi T, Shivapurkar N*, et al.* DNA methylation-associated inactivation of TGFβ-related genes DRM/Gremlin, RUNX3, and HPP1 in human cancers. *British Journal Of Cancer*. 2005;93:1029. doi:10.1038/sj.bjc.6602837.

154. Zhang C, Zhao H, Li J, Liu H, Wang F, Wei Y*, et al.* The Identification of Specific Methylation Patterns across Different Cancers. *PLoS ONE*. 2015;10(3):e0120361. doi:10.1371/journal.pone.0120361.

155. Lin K, Taylor JR, Jr., Wu TD, Gutierrez J, Elliott JM, Vernes J-M*, et al.* TMEFF2 is a PDGF-AA binding protein with methylation-associated gene silencing in multiple cancer types including glioma. *PLoS ONE*. 2011;6(4):e18608-e18608. doi:10.1371/journal.pone.0018608.

156. Belshaw NJ, Elliott GO, Williams EA, Bradburn DM, Mills SJ, Mathers JC*, et al.* Use of DNA from Human Stools to Detect Aberrant CpG Island Methylation of Genes Implicated in Colorectal Cancer. *Cancer Epidemiology, Biomarkers & Prevention*. 2004;13(9):1495-1501.

157. Wynter CV, Walsh MD, Higuchi T, Leggett BA, Young J, Jass JR. Methylation patterns define two types of hyperplastic polyp associated with colorectal cancer. *Gut*. 2004;53(4):573-580. doi:10.1136/gut.2003.030841.

158. Takahashi T, Shivapurkar N, Riquelme E, Shigematsu H, Reddy J, Suzuki M*, et al.* Aberrant promoter hypermethylation of multiple genes in gallbladder carcinoma and chronic cholecystitis. *Clinical Cancer Research*. 2004;10(18 Pt 1):6126-6133. doi:10.1158/1078-0432.Ccr-04-0579.

159. Chen E, Zheng F, Yuan X, Ye Y, Li X, Dai Y*, et al.* The effect of TMEFF2 methylation on the tumor stage and survival outcome of clear cell renal cell carcinoma. *Cancer Biomarkers : Section A of Disease Markers*. 2017;19(2):207-212. doi:10.3233/cbm-161656.

160. Geddert H, Kiel S, Iskender E, Florl AR, Krieg T, Vossen S*, et al.* Correlation of hMLH1 and HPP1 hypermethylation in gastric, but not in esophageal and cardiac adenocarcinoma. *International Journal of Cancer*. 2004;110(2):208-211. doi:10.1002/ijc.20058.

161. Shibata DM, Sato F, Mori Y, Perry K, Yin J, Wang S*, et al.* Hypermethylation of HPP1 is associated with hMLH1 hypermethylation in gastric adenocarcinomas. *Cancer Research*. 2002;62(20):5637-5640.

162. Suzuki M, Toyooka S, Shivapurkar N, Shigematsu H, Miyajima K, Takahashi T*, et al.* Aberrant methylation profile of human malignant mesotheliomas and its relationship to SV40 infection. *Oncogene*. 2005;24(7):1302-1308. doi:10.1038/sj.onc.1208263.

163. Sun T-T, Tang J-Y, Du W, Zhao H-J, Zhao G, Yang S-L*, et al.* Bidirectional regulation between TMEFF2 and STAT3 may contribute to Helicobacter pylori-associated gastric carcinogenesis. *International Journal of Cancer*. 2015;136(5):1053-1064. doi:10.1002/ijc.29061.

164. Sabbioni S, Miotto E, Veronese A, Sattin E, Gramantieri L, Bolondi L*, et al.* Multigene methylation analysis of gastrointestinal tumors: TPEF emerges as a frequent tumor-specific aberrantly methylated marker that can be detected in peripheral blood. *Molecular Diagnosis*. 2003;7(3-4):201-207. doi:10.1007/bf03260039.

165. Wang J, Elahi A, Ajidahun A, Clark W, Hernandez J, Achille A*, et al.* The interplay between histone deacetylases and c-Myc in the transcriptional suppression of HPP1 in colon cancer. *Cancer Biology & Therapy*. 2014;15(9):1198-1207. doi:10.4161/cbt.29500.

166. Gery S, Sawyers CL, Agus DB, Said JW, Koeffler HP. TMEFF2 is an androgen-regulated gene exhibiting antiproliferative effects in prostate cancer cells. *Oncogene*. 2002;21(31):4739-4746. doi:10.1038/sj.onc.1205142.

167. Afar DE, Bhaskar V, Ibsen E, Breinberg D, Henshall SM, Kench JG*, et al.* Preclinical validation of anti-TMEFF2-auristatin E-conjugated antibodies in the treatment of prostate cancer. *Molecular Cancer Therapeutics*. 2004;3(8):921-932.

168. Torrecilla D, Lozano MV, Lallana E, Neissa JI, Novoa-Carballal R, Vidal A*, et al.* Anti-tumor efficacy of chitosan-g-poly(ethylene glycol) nanocapsules containing docetaxel: anti-TMEFF-2 functionalized nanocapsules vs. non-functionalized nanocapsules. *European Journal of Pharmaceutics and Biopharmaceutics*. 2013;83(3):330-337. doi:10.1016/j.ejpb.2012.10.017.

169. Fuentes AM-P, Brazma A, Jarnuczak AF, Füllgrabe A, Barrera E, Vizcaino JA*, et al.* Expression Atlas: gene and protein expression across multiple studies and organisms. *Nucleic Acids Research*. 2017;46(D1):D246-D251. doi:10.1093/nar/gkx1158.

170. Consortium G. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348(6235):648-660. doi:10.1126/science.1262110.

171. Asmann YW, Necela BM, Kalari KR, Hossain A, Baker TR, Carr JM*, et al.* Detection of redundant fusion transcripts as biomarkers or disease-specific therapeutic targets in breast cancer. *Cancer Research*. 2012;72(8):1921-1928. doi:10.1158/0008-5472.can-11-3142.

172. Corbin JM, Overcash RF, Wren JD, Coburn A, Tipton GJ, Ezzell JA*, et al.* Analysis of TMEFF2 allografts and transgenic mouse models reveals roles in prostate regeneration and cancer. *The Prostate*. 2016;76(1):97-113. doi:10.1002/pros.23103.

173. Kanemoto N, Horie M, Omori K, Nishino N, Kondo M, Noguchi K*, et al.* Expression of TMEFF1 mRNA in the mouse central nervous system: precise examination and comparative studies of TMEFF1 and TMEFF2. *Brain research. Molecular Brain Research*. 2001;86(1-2):48-55. doi:10.1016/S0169-328X(00)00257-6.

174. Chen TR, Wang P, Carroll LK, Zhang YJ, Han BX, Wang F. Generation and characterization of Tmeff2 mutant mice. *Biochemical and Biophysical Research Communications*. 2012;425(2):189-194. doi:10.1016/j.bbrc.2012.07.064.

175. Zeng F, Harris RC. Epidermal growth factor, from gene organization to bedside. *Seminars in Cell & Developmental Biology*. 2014;28:2-11. doi:10.1016/j.semcdb.2014.01.011.

176. Normanno N, De Luca A, Bianco C, Strizzi L, Mancino M, Maiello MR*, et al.* Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene*. 2006;366(1):2-16. doi:10.1016/j.gene.2005.10.018.

177. Patel K. Follistatin. *The International Journal of Biochemistry & Cell Biology*. 1998;30(10):1087-1093. doi:10.1016/S1357-2725(98)00064-8.

178. Namwanje M, Brown CW. Activins and Inhibins: Roles in Development, Physiology, and Disease. 2016;8(7). doi:10.1101/cshperspect.a021881.

179. Huang H, Teng P, Mei R, Yang A, Zhang Z, Zhao X*, et al.* Tmeff2 is expressed in differentiating oligodendrocytes but dispensable for their differentiation in vivo. *Scientific Reports*. 2017;7(1):337. doi:10.1038/s41598-017-00407-1.

180. Zhao XY, Liu HL, Liu B, Willuda J, Siemeister G, Mahmoudi M*, et al.* Tomoregulin internalization confers selective cytotoxicity of immunotoxins on prostate cancer cells. *Translational Oncology*. 2008;1(2):102-109. doi:10.1593/tlo.08124.

181. Overcash RF, Chappell VA, Green T, Geyer CB, Asch AS, Ruiz-Echevarria MJ. Androgen signaling promotes translation of TMEFF2 in prostate cancer cells via phosphorylation of the alpha subunit of the translation initiation factor 2. *PLoS ONE*. 2013;8(2):e55257. doi:10.1371/journal.pone.0055257.

182. Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M*, et al.* Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells. *Molecular Cell*. 2000;6(5):1099-1108. doi:10.1016/S1097-2765(00)00108-8.

183. Edwards DR, Handsley MM, Pennington CJ. The ADAM metalloproteinases. *Molecular Aspects of Medicine*. 2008;29(5):258-289. doi:10.1016/j.mam.2008.08.001.

184. Blobel CP. ADAMs: key components in EGFR signalling and development. *Nature Reviews Molecular Cell Biology*. 2005;6(1):32-43. doi:10.1038/nrm1548.

185. Lin H, Wada K, Yonezawa M, Shinoki K, Akamatsu T, Tsukui T*, et al.* Tomoregulin ectodomain shedding by proinflammatory cytokines. *Life Sciences*. 2003;73(13):1617-1627. doi:10.1016/S0024-3205(03)00514-9.

186. Gawel-Beben K, Ali N, Ellis V, Velasco G, Poghosyan Z, Ager A*, et al.* TMEFF2 shedding is regulated by oxidative stress and mediated by ADAMs and transmembrane serine proteases implicated in prostate cancer. *Cell Biology International*. 2018;42(3):273-280. doi:10.1002/cbin.10832.

187. Tanabe LM, List K. The role of type II transmembrane serine protease-mediated signaling in cancer. *The FEBS Journal*. 2017;284(10):1421-1436. doi:10.1111/febs.13971.

188. Gaweł-Bęben K, Ali N, Ellis V, Velasco G, Poghosyan Z, Ager A*, et al.* TMEFF2 shedding is regulated by oxidative stress and mediated by ADAMs and transmembrane serine proteases implicated in prostate cancer. *Cell Biology International*. 2018;42(3):273-280. doi:10.1002/cbin.10832.

189. Chen X, Overcash R, Green T, Hoffman D, Asch AS, Ruiz-Echevarria MJ. The tumor suppressor activity of the transmembrane protein with epidermal growth factor and two follistatin motifs 2 (TMEFF2) correlates with its ability to modulate sarcosine levels. *Journal of Biological Chemistry*. 2011;286(18):16091-16100. doi:10.1074/jbc.M110.193805.

190. Green T, Chen X, Ryan S, Asch AS, Ruiz-Echevarría MJ. TMEFF2 and SARDH cooperate to modulate one-carbon metabolism and invasion of prostate cancer cells. *The Prostate*. 2013;73(14):1561-1575. doi:10.1002/pros.22706.

191. Chen X, Corbin JM, Tipton GJ, Yang LV, Asch AS, Ruiz-Echevarría MJ. The TMEFF2 tumor suppressor modulates integrin expression, RhoA activation and migration of prostate cancer cells. *Biochimica et Biophysica Acta*. 2014;1843(6):1216-1224. doi:10.1016/j.bbamcr.2014.03.005.

192. Elahi A, Zhang L, Yeatman TJ, Gery S, Sebti S, Shibata D. HPP1-mediated tumor suppression requires activation of STAT1 pathways. *International Journal of Cancer*. 2008;122(7):1567-1572. doi:10.1002/ijc.23202.

193. Thyrell L, Erickson S, Zhivotovsky B, Pokrovskaja K, Sangfelt O, Castro J*, et al.* Mechanisms of Interferon-alpha induced apoptosis in malignant cells. *Oncogene*. 2002;21:1251. doi:10.1038/sj.onc.1205179.

194. Hernandez JM, Elahi A, Clark W, Humphries LA, Wang J, Achille A*, et al.* The Tumor Suppressive Effects of HPP1 Are Mediated Through JAK-STAT-Interferon Signaling Pathways. *DNA and Cell Biology*. 2015;34(8):541-549. doi:10.1089/dna.2014.2730.

195. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D*, et al.* ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*. 2004;6(1):1-6. doi:10.1016/s1476-5586(04)80047-2.

196. *Oncomine Research Edition*. Available from: https://www.oncomine.org/resource/login.html. [Accessed 20 May 2019]

197. The Human Protein Atlas. *TMEFF2*. Available from: https://www.proteinatlas.org/ENSG00000144339-TMEFF2/pathology/breast+cancer#imid_4069248. [Accessed 20 May 2019]

198. Janicke RU. MCF-7 breast carcinoma cells do not express caspase-3. *Breast Cancer Research and Treatment*. 2009;117(1):219-221. doi:10.1007/s10549-008-0217-9.

199. Gong Y, Schumacher SE, Wu WH, Tang F, Beroukhim R, Chan TA. Pan-Cancer Analysis Links PARK2 to BCL-XL-Dependent Control of Apoptosis. *Neoplasia*. 2017;19(2):75-83. doi:10.1016/j.neo.2016.12.006.

200. Zhao Y, Zeng X, Tang H, Ye D, Liu J. Combination of metformin and paclitaxel suppresses proliferation and induces apoptosis of human prostate cancer cells via oxidative stress and targeting the mitochondria-dependent pathway. *Oncology Letters*. 2019;17(5):4277-4284. doi:10.3892/ol.2019.10119.

201. Bae J, Leo CP, Hsu SY, Hsueh AJW. MCL-1S, a Splicing Variant of the Antiapoptotic BCL-2 Family Member MCL-1, Encodes a Proapoptotic Protein Possessing Only the BH3 Domain. *Journal of Biological Chemistry*. 2000;275(33):25255-25261. doi:10.1074/jbc.M909826199.

202. Casey G, Lo-Hsueh M, Lopez ME, Vogelstein B, Stanbridge EJ. Growth suppression of human breast cancer cells by the introduction of a wild-type p53 gene. *Oncogene*. 1991;6(10):1791-1797.

203. Komarov PG, Komarova EA, Kondratov RV, Christov-Tselkov K, Coon JS, Chernov MV*, et al.* A chemical inhibitor of p53 that protects mice from the side effects of cancer therapy. *Science*. 1999;285(5434):1733-1737. doi:10.1126/science.285.5434.1733.

204. Ostrakhovitch EA, Cherian MG. Role of p53 and reactive oxygen species in apoptotic response to copper and zinc in epithelial breast cancer cells. *Apoptosis*. 2005;10(1):111-121. doi:10.1007/s10495-005-6066-7.

205. Essmann F, Engels IH, Totzke G, Schulze-Osthoff K, Janicke RU. Apoptosis resistance of MCF-7 breast carcinoma cells to ionizing radiation is independent of p53 and cell cycle control but caused by the lack of caspase-3 and a caffeine-inhibitable event. *Cancer Research*. 2004;64(19):7065-7072. doi:10.1158/0008-5472.Can-04-1082.

206. Wang S, Konorev EA, Kotamraju S, Joseph J, Kalivendi S, Kalyanaraman B. Doxorubicin Induces Apoptosis in Normal and Tumor Cells via Distinctly Different Mechanisms: INTERMEDIACY OF H2O2- AND p53-DEPENDENT PATHWAYS. *Journal of Biological Chemistry*. 2004;279(24):25535-25543. doi:10.1074/jbc.M400944200.

207. Qiagen. *Ingenuity Pathway Analysis (IPA)*. Available from: https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis. [Accessed 20 May 2019]

208. Liu F, Hu X, Zimmerman M, Waller JL, Wu P, Hayes-Jordan A*, et al.* TNFalpha cooperates with IFN-gamma to repress Bcl-xL expression to sensitize metastatic colon carcinoma cells to TRAIL-mediated apoptosis. *PLoS ONE*. 2011;6(1):e16241. doi:10.1371/journal.pone.0016241.

209. Bryan J. Interferon was not the miracle cure for cancer hoped for in its early days. *The Pharmaceutical Journal*. 2018;280:637.

210. Kotredes KP, Gamero AM. Interferons as inducers of apoptosis in malignant cells. *Journal of Interferon & Cytokine Research*. 2013;33(4):162-170. doi:10.1089/jir.2012.0110.

211. Di Franco S, Turdo A, Todaro M, Stassi G. Role of Type I and II Interferons in Colorectal Cancer and Melanoma. *Frontiers in immunology*. 2017;8:878-878. doi:10.3389/fimmu.2017.00878.

212. ClinicalTrials.gov. U.S.National Library of Medicine. *A Phase I-II Study of Interferon-gamma Plus Weekly Paclitaxel, Trastuzumab and Pertuzumab in Patients With HER-2 Positive Breast Cancer*. Available from: https://clinicaltrials.gov/ct2/show/NCT03112590. [Accessed 24 June 2019]

213. Cheon H, Borden EC, Stark GR. Interferons and their stimulated genes in the tumor microenvironment. *Seminars in oncology*. 2014;41(2):156-173. doi:10.1053/j.seminoncol.2014.02.002.

214. Li K, Qu S, Chen X, Wu Q, Shi M. Promising Targets for Cancer Immunotherapy: TLRs, RLRs, and STING-Mediated Innate Immune Pathways. *International Journal of Molecular Sciences*. 2017;18(2):404. doi:10.3390/ijms18020404.

215. Freier CP, Kuhn C, Endres S, Mayr D, Friese K, Jeschke U*, et al.* FOXP3+ Cells Recruited by CCL22 into Breast Cancer Correlates with Less Tumor Nodal Infiltration. *Anticancer Research*. 2016;36(6):3139-3145.

216. Chawla-Sarkar M, Lindner DJ, Liu Y-F, Williams BR, Sen GC, Silverman RH*, et al.* Apoptosis and interferons: Role of interferon-stimulated genes as mediators of apoptosis. *Apoptosis*. 2003;8(3):237-249. doi:10.1023/a:1023668705040.

217. Sahin H, Borkham-Kamphorst E, do ON, Berres ML, Kaldenbach M, Schmitz P*, et al.* Proapoptotic effects of the chemokine, CXCL 10 are mediated by the noncognate receptor TLR4 in hepatocytes. *Hepatology*. 2013;57(2):797-805. doi:10.1002/hep.26069.

218. Paroni F, Domsgen E, Maedler K. CXCL10- a path to beta-cell death. *Islets*. 2009;1(3):256-259. doi:10.4161/isl.1.3.9110.

219. Singh L, Arora SK, Bakshi DK, Majumdar S, Wig JD. Potential role of CXCL10 in the induction of cell injury and mitochondrial dysfunction. *International Journal of Experimental Pathology*. 2010;91(3):210-223. doi:10.1111/j.1365-2613.2009.00697.x.

220. Zhang HM, Yuan J, Cheung P, Chau D, Wong BW, McManus BM*, et al.* Gamma interferon-inducible protein 10 induces HeLa cell apoptosis through a p53-dependent pathway initiated by suppression of human papillomavirus type 18 E6 and E7 expression. *Molecular and Cellular Biology*. 2005;25(14):6247-6258. doi:10.1128/mcb.25.14.6247-6258.2005.

221. Sistigu A, Yamazaki T, Vacchelli E, Chaba K, Enot DP, Adam J*, et al.* Cancer cell–autonomous contribution of type I interferon signaling to the efficacy of chemotherapy. *Nature Medicine*. 2014;20(11):1301-1309. doi:10.1038/nm.3708.

222. O'Shea JJ, Schwartz DM, Villarino AV, Gadina M, McInnes IB, Laurence A. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annual Review of Medicine*. 2015;66:311-328. doi:10.1146/annurev-med-051113-024537.

223. Thyrell L, Hjortsberg L, Arulampalam V, Panaretakis T, Uhles S, Dagnell M*, et al.* Interferon α-induced Apoptosis in Tumor Cells Is Mediated through the Phosphoinositide 3-Kinase/Mammalian Target of Rapamycin Signaling Pathway. *Journal of Biological Chemistry*. 2004;279(23):24152-24162. doi:10.1074/jbc.M312219200.

224. Panaretakis T, Hjortsberg L, Tamm KP, Bjorklund AC, Joseph B, Grander D. Interferon alpha induces nucleus-independent apoptosis by activating extracellular signal-regulated kinase 1/2 and c-Jun NH2-terminal kinase downstream of phosphatidylinositol 3-kinase and mammalian target of rapamycin. *Molecular Biology of the Cell*. 2008;19(1):41-50. doi:10.1091/mbc.e07-04-0358.

225. Thyrell L, Erickson S, Zhivotovsky B, Pokrovskaja K, Sangfelt O, Castro J*, et al.* Mechanisms of Interferon-alpha induced apoptosis in malignant cells. *Oncogene*. 2002;21(8):1251-1262. doi:10.1038/sj.onc.1205179.

226. Stuchbury G, Münch G. Optimizing the generation of stable neuronal cell lines via pre-transfection restriction enzyme digestion of plasmid DNA. *Cytotechnology*. 2010;62(3):189-194. doi:10.1007/s10616-010-9273-1.

227. Singh KP, Treas J, Tyagi T, Gao W. DNA demethylation by 5-aza-2-deoxycytidine treatment abrogates 17 beta-estradiol-induced cell growth and restores expression of DNA repair genes in human breast cancer cells. *Cancer Letters*. 2012;316(1):62-69. doi:10.1016/j.canlet.2011.10.022.

228. Xiong H, Qiu H, Zhuang L, Xiong H, Jiang R, Chen Y. Effects of 5-Aza-CdR on the proliferation of human breast cancer cell line MCF-7 and on the expression of Apaf-1 gene. *Journal of Huazhong*

University of Science and Technology [Medical Sciences]. 2009;29(4):498-502. doi:10.1007/s11596-009-0421-9.

229. Timpl R, Sasaki T, Kostka G, Chu M-L. Fibulins: a versatile family of extracellular matrix proteins. *Nature Reviews Molecular Cell Biology*. 2003;4(6):479-489. doi:10.1038/nrm1130.

230. Nakamura T, Lozano PR, Ikeda Y, Iwanaga Y, Hinek A, Minamisawa S*, et al.* Fibulin-5/DANCE is essential for elastogenesis in vivo. *Nature*. 2002;415(6868):171-175. doi:10.1038/415171a.

231. Markova D, Zou Y, Ringpfeil F, Sasaki T, Kostka G, Timpl R*, et al.* Genetic Heterogeneity of Cutis Laxa: A Heterozygous Tandem Duplication within the Fibulin-5 Gene. *The American Journal of Human Genetics*. 2003;72(4):998-1004. doi:10.1086/373940.

232. Stone EM, Braun TA, Russell SR, Kuehn MH, Lotery AJ, Moore PA*, et al.* Missense variations in the fibulin 5 gene and age-related macular degeneration. *New England Journal of Medicine*. 2004;351(4):346-353. doi:10.1056/NEJMoa040833.

233. Schiemann WP, Blobe GC, Kalume DE, Pandey A, Lodish HF. Context-specific Effects of Fibulin-5 (DANCE/EVEC) on Cell Proliferation, Motility, and Invasion: FIBULIN-5 IS INDUCED BY TRANSFORMING GROWTH FACTOR-β AND AFFECTS PROTEIN KINASE CASCADES. *Journal of Biological Chemistry*. 2002;277(30):27367-27377. doi:10.1074/jbc.M200148200.

234. Albig AR, Schiemann WP. Fibulin-5 Antagonizes Vascular Endothelial Growth Factor (VEGF) Signaling and Angiogenic Sprouting by Endothelial Cells. *DNA and Cell Biology*. 2004;23(6):367-379. doi:10.1089/104454904323145254.

235. Watson JD, Oster SK, Shago M, Khosravi F, Penn LZ. Identifying Genes Regulated in a Myc-dependent Manner. *Journal of Biological Chemistry*. 2002;277(40):36921-36930. doi:10.1074/jbc.M201493200.

236. Kuang P-P, Goldstein RH, Liu Y, Rishikof DC, Jean J-C, Joyce-Brady M. Coordinate expression of fibulin-5/DANCE and elastin during lung injury repair. *American Journal of Physiology-Lung Cellular and Molecular Physiology*. 2003;285(5):L1147-L1152. doi:10.1152/ajplung.00098.2003.

237. Heo JH, Song JY, Jeong JY, Kim G, Kim TH, Kang H*, et al.* Fibulin-5 is a tumour suppressor inhibiting cell migration and invasion in ovarian cancer. *Journal of Clinical Pathology*. 2016;69(2):109-116. doi:10.1136/jclinpath-2015-203129.

238. Hu Z, Ai Q, Xu H, Ma X, Li HZ, Shi TP*, et al.* Fibulin-5 is down-regulated in urothelial carcinoma of bladder and inhibits growth and invasion of human bladder cancer cell line 5637. *Urologic Oncology: Seminars and Original Investigations*. 2011;29(4):430-435. doi:10.1016/j.urolonc.2009.06.004.

239. Tang JC, Liu JH, Liu XL, Liang X, Cai XJ. Effect of fibulin-5 on adhesion, migration and invasion of hepatocellular carcinoma cells via an integrin-dependent mechanism. *World Journal of Gastroenterology*. 2015;21(39):11127-11140. doi:10.3748/wjg.v21.i39.11127.

240. Tu K, Dou C, Zheng X, Li C, Yang W, Yao Y*, et al.* Fibulin-5 inhibits hepatocellular carcinoma cell migration and invasion by down-regulating matrix metalloproteinase-7 expression. *BMC Cancer*. 2014;14(1):938. doi:10.1186/1471-2407-14-938.

241. Chen X, Song X, Yue W, Chen D, Yu J, Yao Z*, et al.* Fibulin-5 inhibits Wnt/β-catenin signaling in lung cancer. *Oncotarget*. 2015;6(17):15022-15034. doi:10.18632/oncotarget.3609.

242. Shi XY, Wang L, Cao CH, Li ZY, Chen J, Li C. Effect of Fibulin-5 on cell proliferation and invasion in human gastric cancer patients. *Asian Pacific Journal of Tropical Medicine*. 2014;7(10):787-791. doi:10.1016/s1995-7645(14)60137-1.

243. Hwang CF, Shiu LY, Su LJ, Yu-Fang Y, Wang WS, Huang SC*, et al.* Oncogenic fibulin-5 promotes nasopharyngeal carcinoma cell metastasis through the FLJ10540/AKT pathway and correlates with poor prognosis. *PLoS ONE*. 2013;8(12):e84218. doi:10.1371/journal.pone.0084218.

244. Lee YH, Schiemann WP. Fibromodulin suppresses nuclear factor-kappaB activity by inducing the delayed degradation of IKBA via a JNK-dependent pathway coupled to fibroblast apoptosis. *Journal of Biological Chemistry*. 2011;286(8):6414-6422. doi:10.1074/jbc.M110.168682.

245. Mohamedi Y, Fontanil T, Solares L, Garcia-Suarez O, Garcia-Piqueras J, Vega JA*, et al.* Fibulin-5 downregulates Ki-67 and inhibits proliferation and invasion of breast cancer cells. *International Journal of Oncology*. 2016;48(4):1447-1456. doi:10.3892/ijo.2016.3394.

246. Devitt Møller H, Ralfkjær U, Cremers N, Frankel M, Troelsgaard Pedersen R, Klingelhöfer J*, et al.* Role of Fibulin-5 in Metastatic Organ Colonization. *Molecular Cancer Research*. 2011;9(5):553-563. doi:10.1158/1541-7786.Mcr-11-0093.

247. Schluterman MK, Chapman SL, Korpanty G, Ozumi K, Fukai T, Yanagisawa H*, et al.* Loss of fibulin-5 binding to beta1 integrins inhibits tumor growth by increasing the level of ROS. *Disease Models & Mechanisms*. 2010;3(5-6):333-342. doi:10.1242/dmm.003707.

248. Wang M, Topalovski M, Toombs JE, Wright CM, Moore ZR, Boothman DA*, et al.* Fibulin-5 Blocks Microenvironmental ROS in Pancreatic Cancer. *Cancer Research*. 2015;75(23):5058-5069. doi:10.1158/0008-5472.Can-15-0744.

249. Topalovski M, Hagopian M, Wang M, Brekken RA. Hypoxia and Transforming Growth Factor beta Cooperate to Induce Fibulin-5 Expression in Pancreatic Cancer. *Journal of Biological Chemistry*. 2016;291(42):22244-22252. doi:10.1074/jbc.M116.730945.

250. Guadall A, Orriols M, Rodríguez-Calvo R, Calvayrac O, Crespo J, Aledo R*, et al.* Fibulin-5 Is Up-regulated by Hypoxia in Endothelial Cells through a Hypoxia-inducible Factor-1 (HIF-1α)-dependent Mechanism. *Journal of Biological Chemistry*. 2011;286(9):7093-7103. doi:10.1074/jbc.M110.162917.

251. Kitanaka C, Namiki T, Noguchi K, Mochizuki T, Kagaya S, Chi S*, et al.* Caspase-dependent apoptosis of COS-7 cells induced by Bax overexpression: differential effects of Bcl-2 and Bcl-xL on Bax-induced caspase activation and apoptosis. *Oncogene*. 1997;15(15):1763-1772. doi:10.1038/sj.onc.1201349.

252. Yoon JY, Baek CW, Kim EJ, Park BS, Yu SB, Yoon JU*, et al.* Propofol protects against oxidative-stress-induced COS-7 cell apoptosis by inducing autophagy. *Journal of Dental Anesthesia and Pain Medicine*. 2017;17(1):37-46. doi:10.17245/jdapm.2017.17.1.37.

253. Tajima H, Tsuchiya K, Yamada M, Kondo K, Katsube N, Ishitani R. Over-expression of GAPDH induces apoptosis in COS-7 cells transfected with cloned GAPDH cDNAs. *NeuroReport*. 1999;10(10):2029-2033.

254. Hiwatashi K, Ueno S, Kubo F, Sakoda M, Tateno T, Hayashi T*, et al.* Relevance of apoptosis and tolerance to hypoxic stress in cells transfected with receptor for advanced glycation end products (RAGE). *Anticancer Research*. 2009;29(4):1287-1294.

255. Guo YX, Wei T, Dallmann K, Kwang J. Induction of caspase-dependent apoptosis by betanodaviruses GGNNV and demonstration of protein α as an apoptosis inducer. *Virology*. 2003;308(1):74-82. doi:10.1016/S0042-6822(02)00098-3.

256. Gu Y, Chen X, Shang C, Singh K, Barzegar M, Mahdavian E*, et al.* Fusarochromanone induces G1 cell cycle arrest and apoptosis in COS7 and HEK293 cells. *PLoS ONE*. 2014;9(11):e112641. doi:10.1371/journal.pone.0112641.

257. Ahuja D, Sáenz-Robles MT, Pipas JM. SV40 large T antigen targets multiple cellular pathways to elicit cellular transformation. *Oncogene*. 2005;24(52):7729-7745. doi:10.1038/sj.onc.1209046.

258. Jasinska AJ, Zelaya I, Service SK, Peterson CB, Cantor RM, Choi O-W*, et al.* Genetic variation and gene expression across multiple tissues and developmental stages in a non-human primate. *Nature Genetics*. 2017;49(12):1714–1721. doi:10.1038/ng.3959.

259. Lachance PED, Chaudhuri A. Microarray analysis of developmental plasticity in monkey primary visual cortex. *Journal of Neurochemistry*. 2004;88(6):1455-1469. doi:10.1046/j.1471-4159.2003.02274.x.

260. Ding Z, Dong J, Liu J, Deng SX. Preferential gene expression in the limbus of the vervet monkey. *Molecular Vision*. 2008;14:2031-2041.

261. Warren WC, Jasinska AJ, Garcia-Perez R, Svardal H, Tomlinson C, Rocchi M*, et al.* The genome of the vervet (Chlorocebus aethiops sabaeus). *Genome Research*. 2015;25(12):1921-1933. doi:10.1101/gr.192922.115.

262. Jasinska AJ, Schmitt CA, Service SK, Cantor RM, Dewar K, Jentsch JD*, et al.* Systems biology of the vervet monkey. *ILAR Journal*. 2013;54(2):122-143. doi:10.1093/ilar/ilt049.

263. Solaini G, Sgarbi G, Baracca A. Oxidative phosphorylation in cancer cells. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 2011;1807(6):534-542. doi:10.1016/j.bbabio.2010.09.003.

264. Zhang J, Zhou L, Nan Z, Yuan Q, Wen J, Xu M*, et al.* Knockdown of cMyc activates Fas-mediated apoptosis and sensitizes A549 cells to radiation. *Oncology Reports*. 2017;38(4):2471-2479. doi:10.3892/or.2017.5897.

265. Luoto KR, Meng AX, Wasylishen AR, Zhao H, Coackley CL, Penn LZ*, et al.* Tumor Cell Kill by c-MYC Depletion: Role of MYC-Regulated Genes that Control DNA Double-Strand Break Repair. *Cancer Research*. 2010;70(21):8748-8759. doi:10.1158/0008-5472.Can-10-0944.

266. Ruiz-Ruiz C, Munoz-Pinedo C, Lopez-Rivas A. Interferon-gamma treatment elevates caspase-8 expression and sensitizes human breast tumor cells to a death receptor-induced mitochondria-operated apoptotic program. *Cancer Research*. 2000;60(20):5673-5680.

9 Appendix

Supplementary Figures.

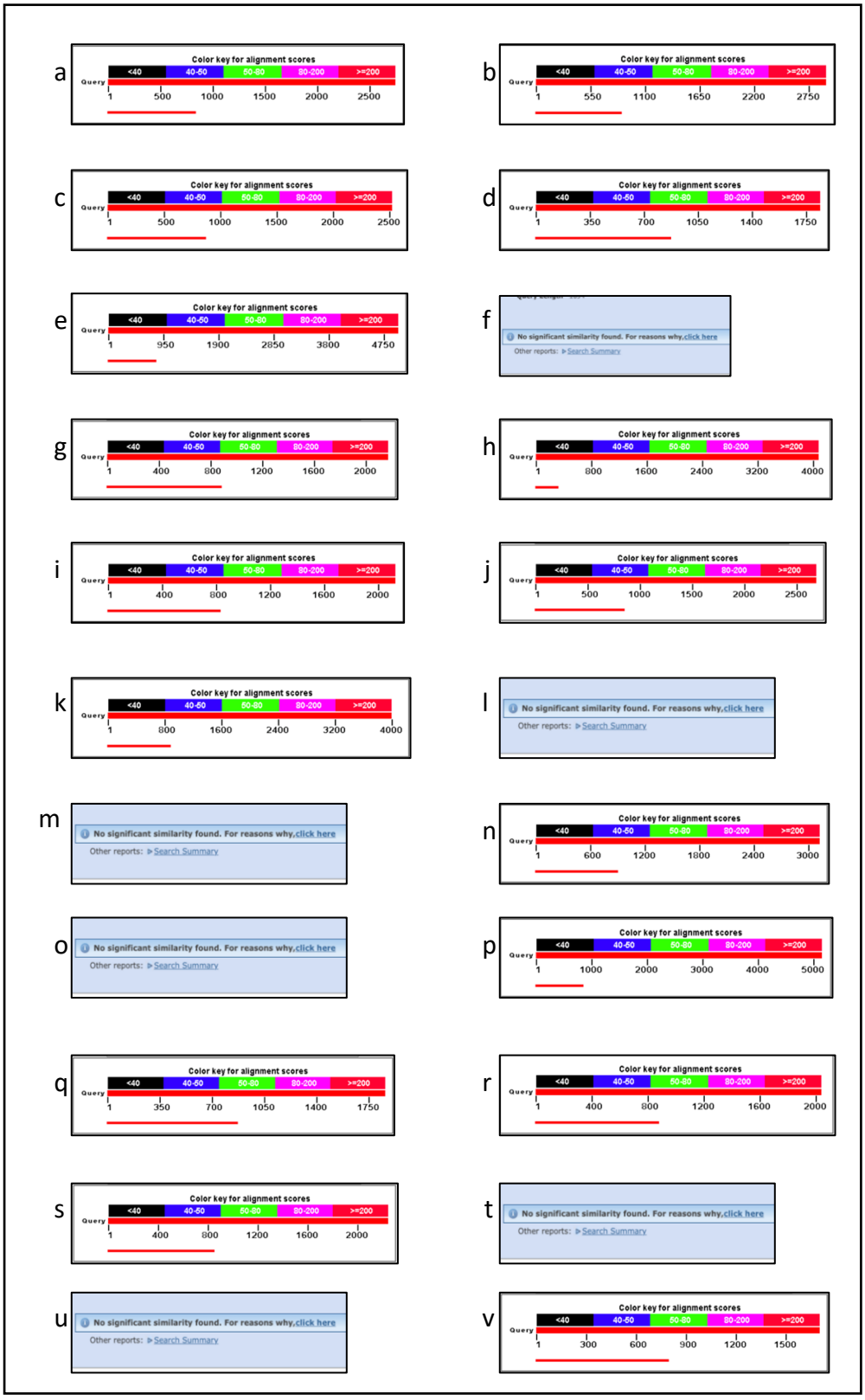


Figure S.1: Sequence alignment between the Sanger sequencing results and the NITE library reference sequences. (a) ACG1, (b) ACG2, (c) ACG3, (d) ACG4, (e) ACG5, (f) ACG6, (g) ACG7, (h) ACG8, (i) ACG9, (j) ACG10, (k) ACG11, (l) ACG12, (m) ACG13, (n) ACG14, (o) ACG15, (p) ACG16, (q) ACG17, (r) ACG18, (s)

ACG19, (t) ACG20, (u) ACG21, (v) ACG22. Screen shots from NCBI nucleotide blast two-sequence alignment tool, BLASTN 2.2.32+.

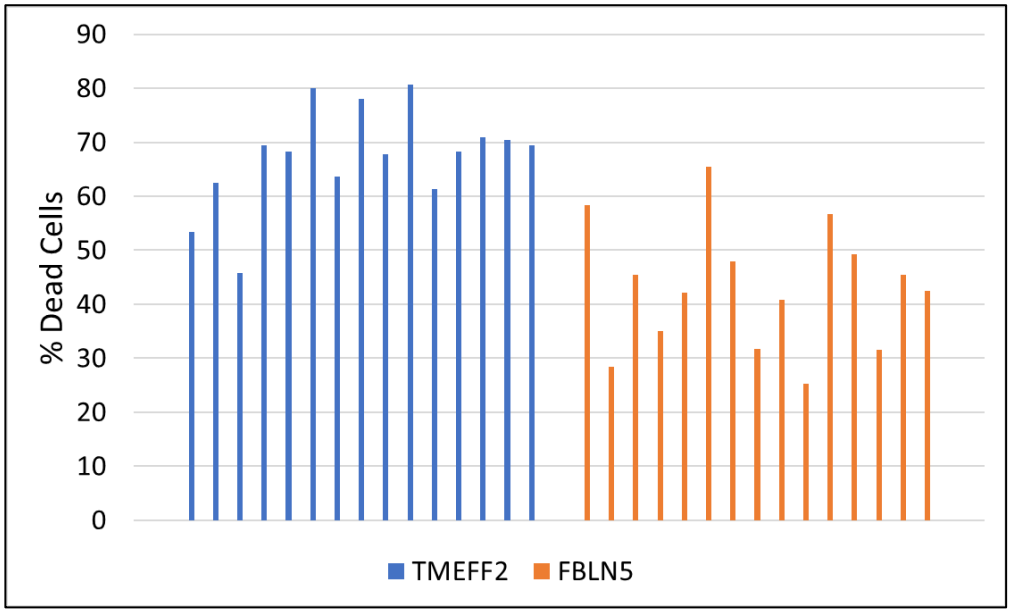


Figure S.2: Cell death in MCF-7 cells upon TMEFF2 and FBLN5 overexpression, detected using TMRE and DRAQ7™ staining through flow cytometry across 15 independently transfected populations of cells. Cells were harvested 48 hours post transfection, stained with TMRE/DRAQ7™ and subjected to flow cytometry using BD LSR II flow cytometer. Excitation lasers: 405nm, 488nm, 561nm and 633nm. TMRE staining detected using 610/20 YG filter and DRAQ7™ using 695/40 R filter. Each bar represents the sum of TMRE -ve and DRAQ7™ +ve events.

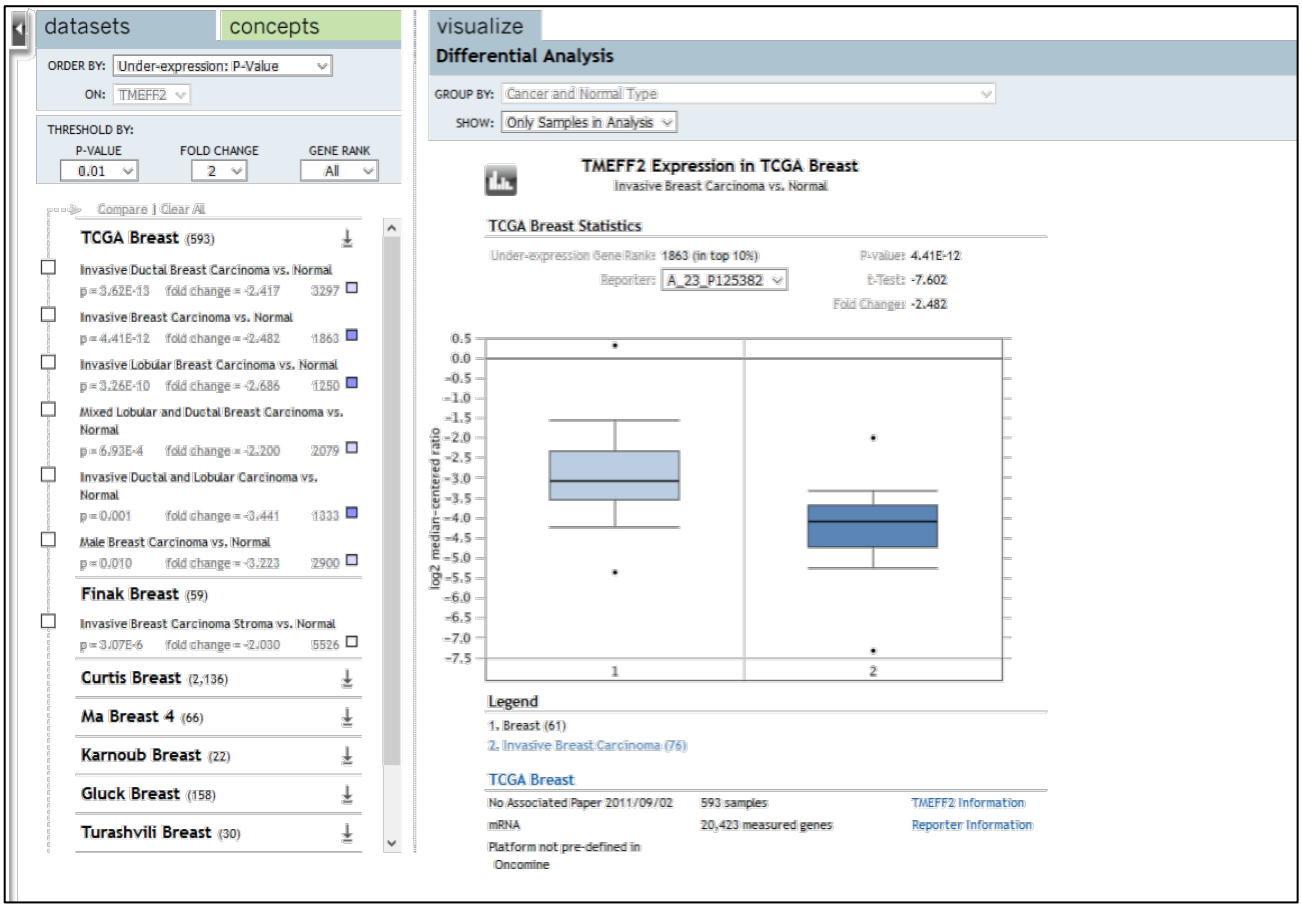


Figure S.3: The screenshot clipped from ONCOMINE™ Research Edition database showing results for the relative mRNA levels of TMEFF2 between normal individuals and Invasive Breast Carcinoma patients. TMEFF2 was ranked among the top 7% of genes downregulated in patients with invasive breast carcinoma according to TGCA data. Source URL: https://www.oncomine.org/resource/login.html.

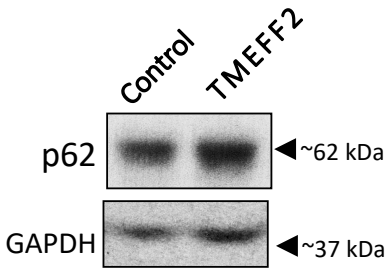


Figure S.4: Immunoblot for p62 upon TMEFF2 overexpression in MCF-7 cells. Cells were transfected with TMEFF2 and GFP (negative 'control') in 10 cm dishes and harvested 30 hours post-transfection. Proteins were separated through SDS-PAGE under reducing conditions and electroblotted onto nitrocellulose membrane. GAPDH was used as a loading control.

File Name count: 6	Labeling Controls Threshold	Hybridization Controls Threshold	Pos vs Neg AUC Threshold	Condition
DC_Motasim_UKB_Sample_13_MCF7_CON_1_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	MCF7 Control
DC_Motasim_UKB_Sample_14_MCF7_CON_2_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	MCF7 Control
DC_Motasim_UKB_Sample_15_MCF7_CON_3_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	MCF7 Control
DC_Motasim_UKB_Sample_16_MCF7_TMF_1_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	MCF7 TMEFF2
DC_Motasim_UKB_Sample_17_MCF7_TMF_2_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	MCF7 TMEFF2
DC_Motasim_UKB_Sample_18_MCF7_TMF_3_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	MCF7 TMEFF2

Figure S.5: Quality control (QC) checks for the microarray analysis of MCF-7 cells transfected with TMEFF2 and MCF-7 cells transfected with luciferase 'control'. Platform: Affymetrix® GeneChip™ Clariom™ D human exome arrays. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1. NB: Labelling controls were not used during library preparation.

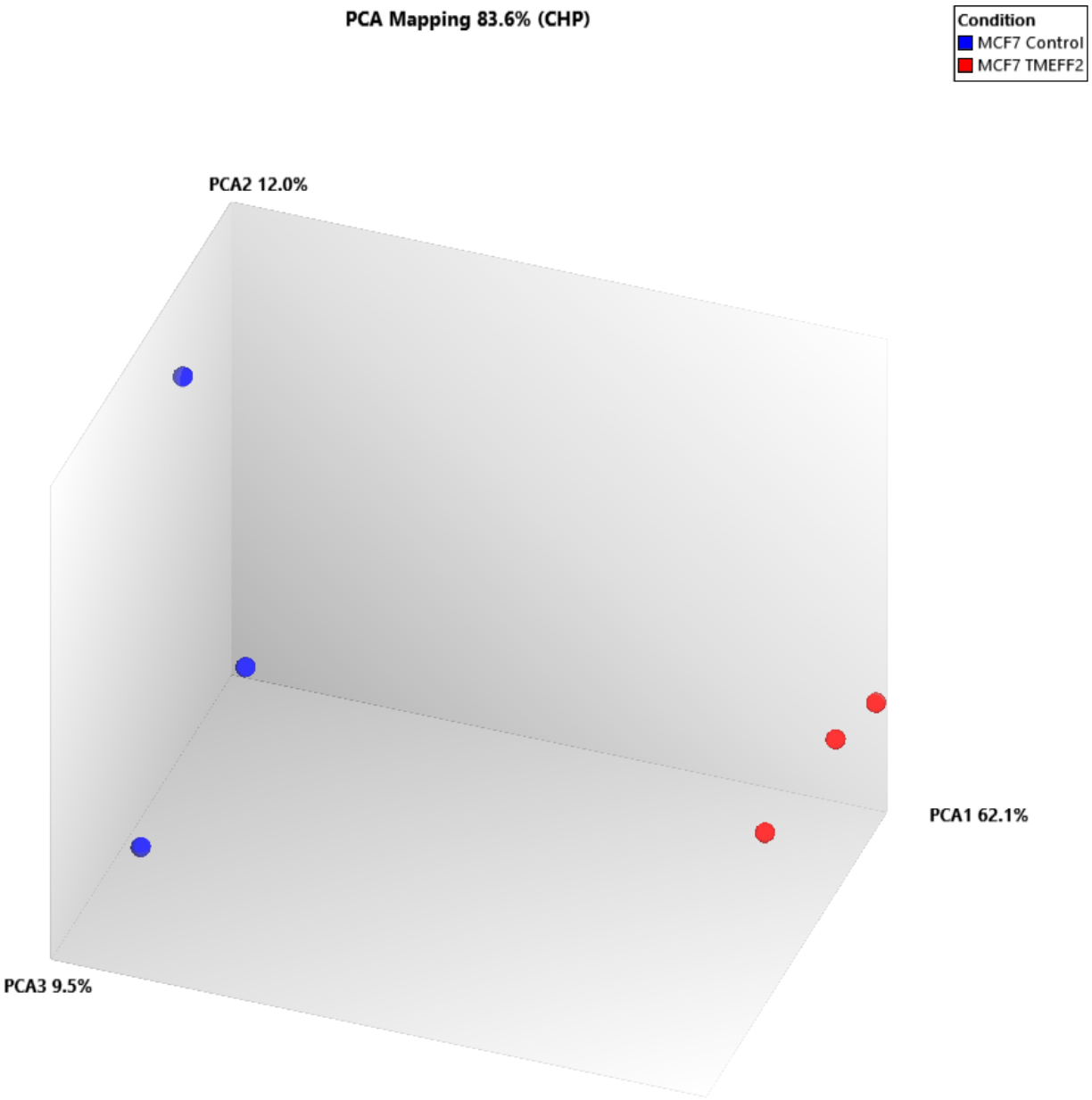


Figure S.6: Principle component analysis (PCA) plot for the microarray analysis of MCF-7 cells transfected with TMEFF2 and MCF-7 cells transfected with luciferase 'control'. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

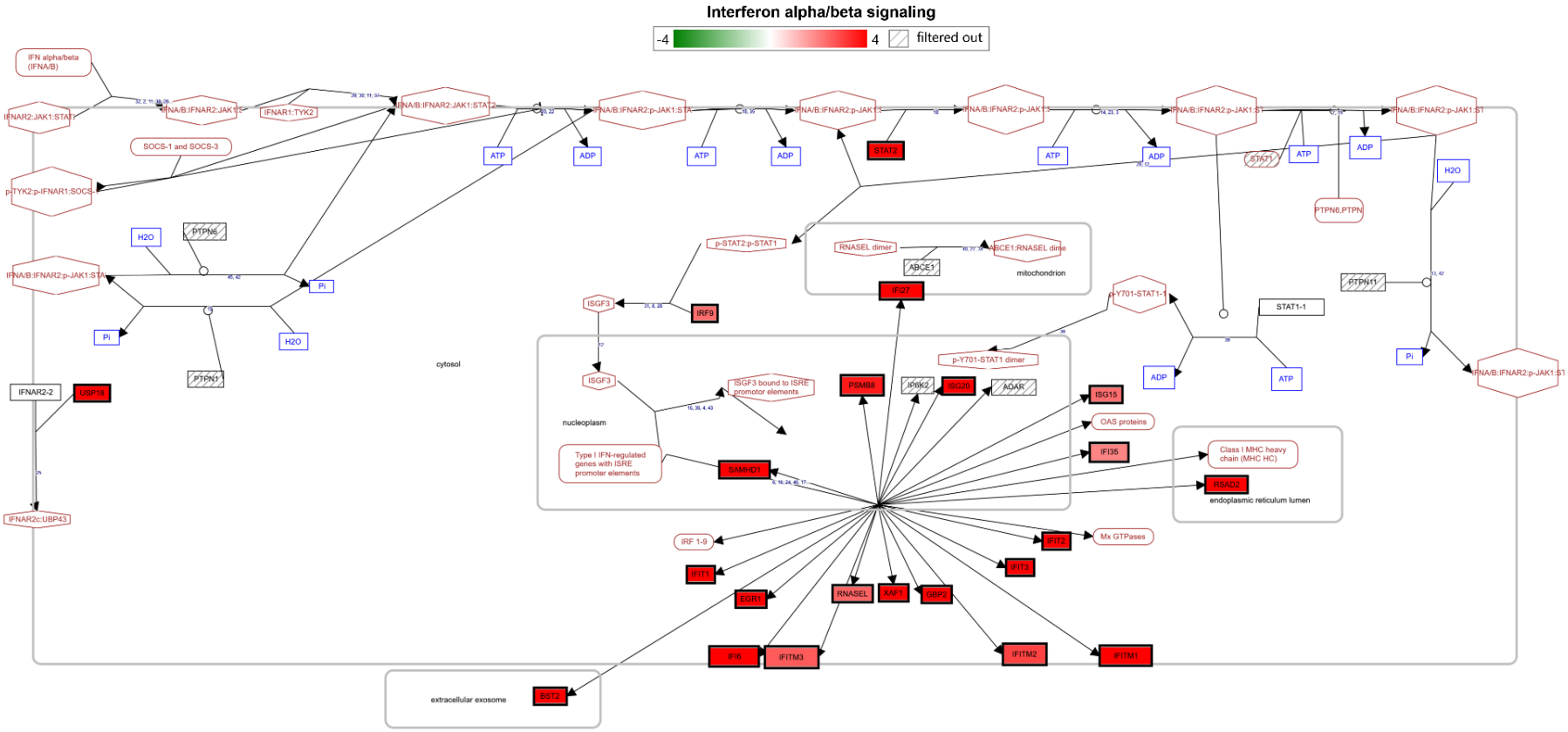


Figure S.7: Interferon alpha/beta pathway (WikiPathway) with overlapping differentially expressed genes between MCF-7 cells transfected with TMEFF2 vs MCF-7 cells transfected with Luciferase 'Control'. Filter applied: Fold change ≥2 to ≤ -2, p ≤0.005 and FDR ≤0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1. with integrated WikiPathways.

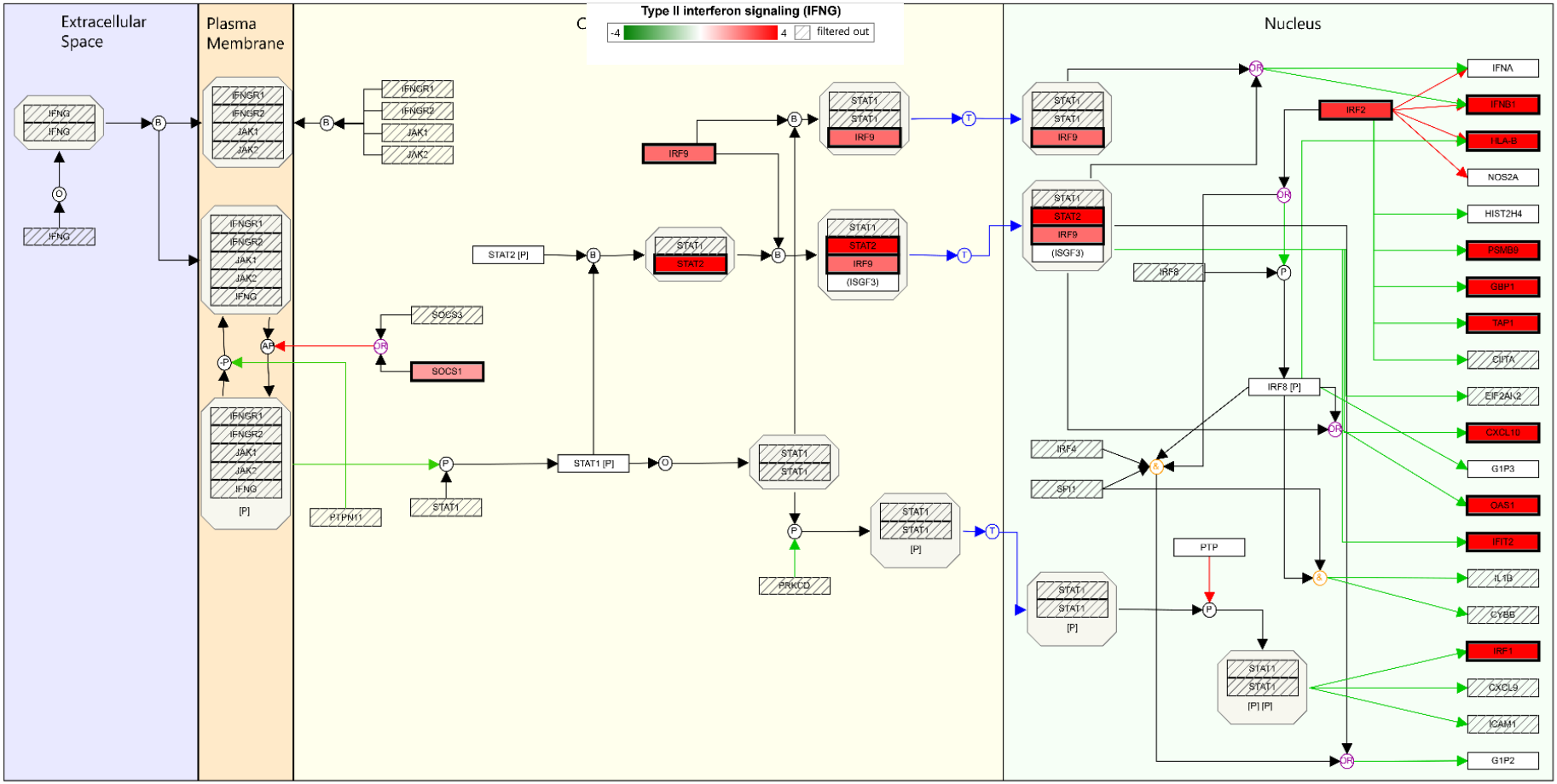


Figure S.8. Type II interferon signalling (IFNG) pathway (WikiPathway) with overlapping differentially expressed genes between MCF-7 cells transfected with TMEFF2 vs MCF-7 cells transfected with Luciferase 'Control'. Filter applied: Fold change ≥2 to ≤ -2, p ≤0.005 and FDR ≤0.05. Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1. with integrated WikiPathways.

File Name count: 12	Labeling Controls Threshold	Hybridization Controls Threshold	Pos vs Neg AUC Threshold	Condition
DC_Motasim_UKB_Sample_1_COS7_FB5_1_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	COS7 FBLN5
DC_Motasim_UKB_Sample_2_COS7_FB5_2_(Clariom_D_Human).sst-rma-combined-dabg.chp	Pass	Pass	Pass	COS7 FBLN5
DC_Motasim_UKB_Sample_3_COS7_FB5_3_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	COS7 FBLN5
DC_Motasim_UKB_Sample_4_COS7_UT_1_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	COS7 Control
DC_Motasim_UKB_Sample_5_COS7_UT_2_(Clariom_D_Human).sst-rma-combined-dabg.chp	Pass	Pass	Pass	COS7 Control
DC_Motasim_UKB_Sample_6_COS7_UT_3_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	COS7 Control
DC_Motasim_UKB_Sample_7_CV1_FB5_1_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	CV1 FBLN5
DC_Motasim_UKB_Sample_8_CV1_FB5_2_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	CV1 FBLN5
DC_Motasim_UKB_Sample_9_CV1_FB5_3_A1_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	CV1 FBLN5
DC_Motasim_UKB_Sample_10_CV1_UT_1_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	CV1 Control
DC_Motasim_UKB_Sample_11_CV1_UT_2_(Clariom_D_Human).sst-rma-combined-dabg.chp	Out	Pass	Pass	CV1 Control
DC_Motasim_UKB_Sample_12_CV1_UT_3_(Clariom_D_Human).sst-rma-combined-dabg.chp	Pass	Pass	Pass	CV1 Control

Figure S.9: Quality control (QC) checks for the microarray analysis of CV-1 cells and COS-7 cells transfected with FBLN5 and wild type CV-1 and COS-7 cells. Platform: Affymetrix® GeneChip™ Clariom™ D human exome arrays. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1. NB: Labelling controls were NOT used during library preparation

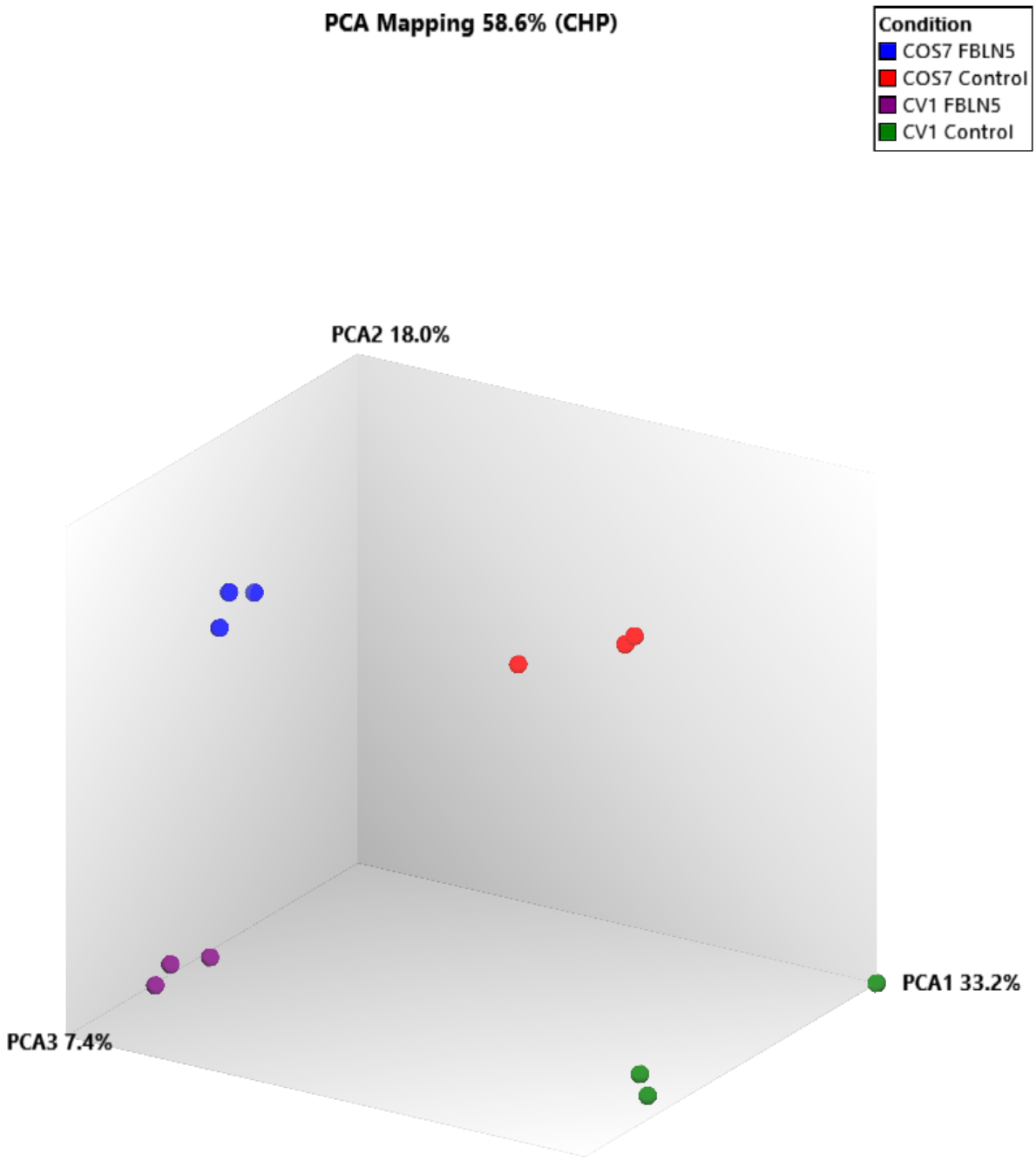


Figure S.10: The principle component analysis (PCA) plot for the microarray analysis of CV-1 cells and COS-7 cells transfected with FBLN5 and wild type CV-1 and COS-7 cells ('control'). Platform: Affymetrix® GeneChip™ Clariom™ D human exome array. Analysis results generated by Transcriptome Analysis Console (TAC) 4.0.1.

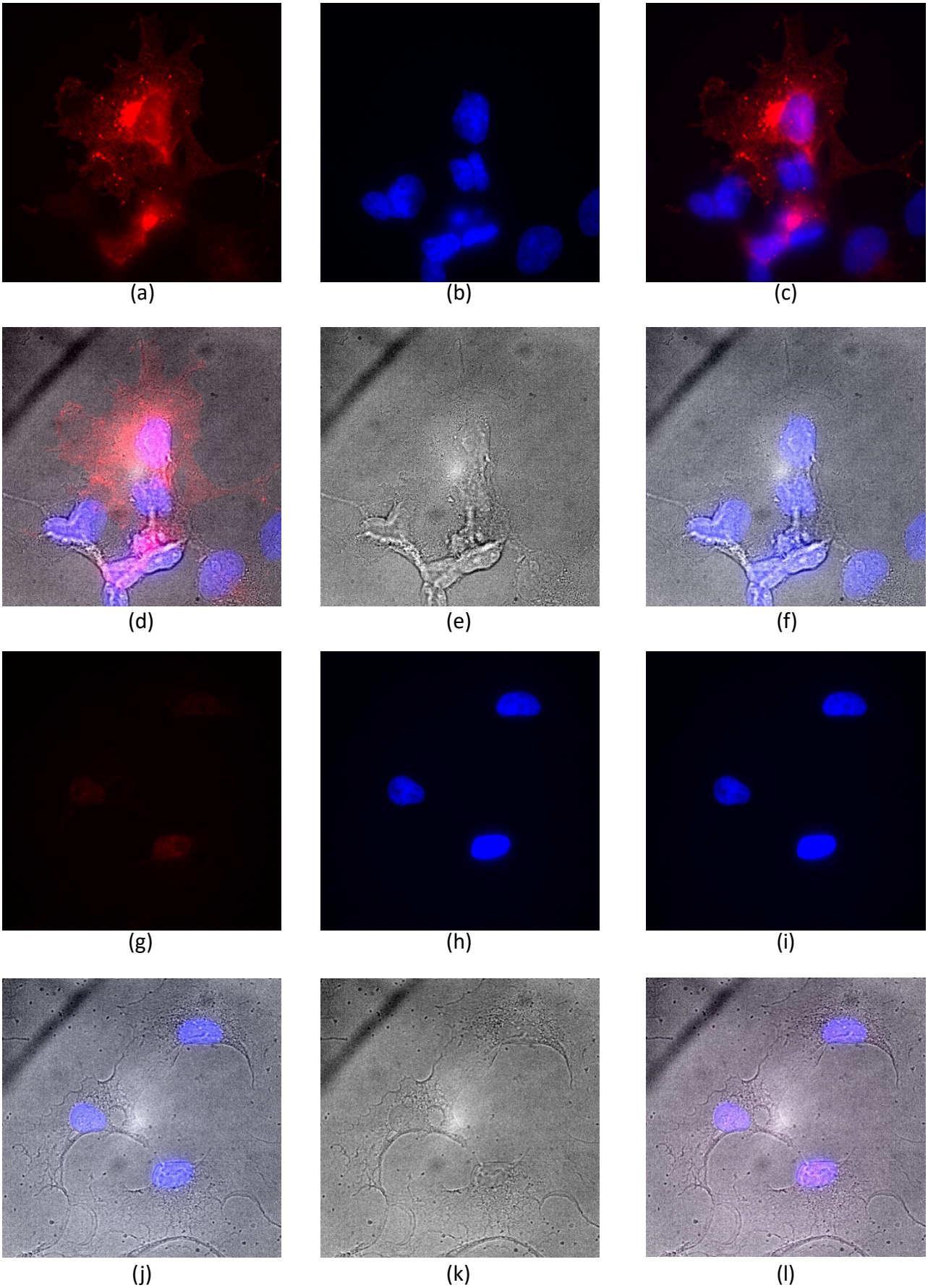
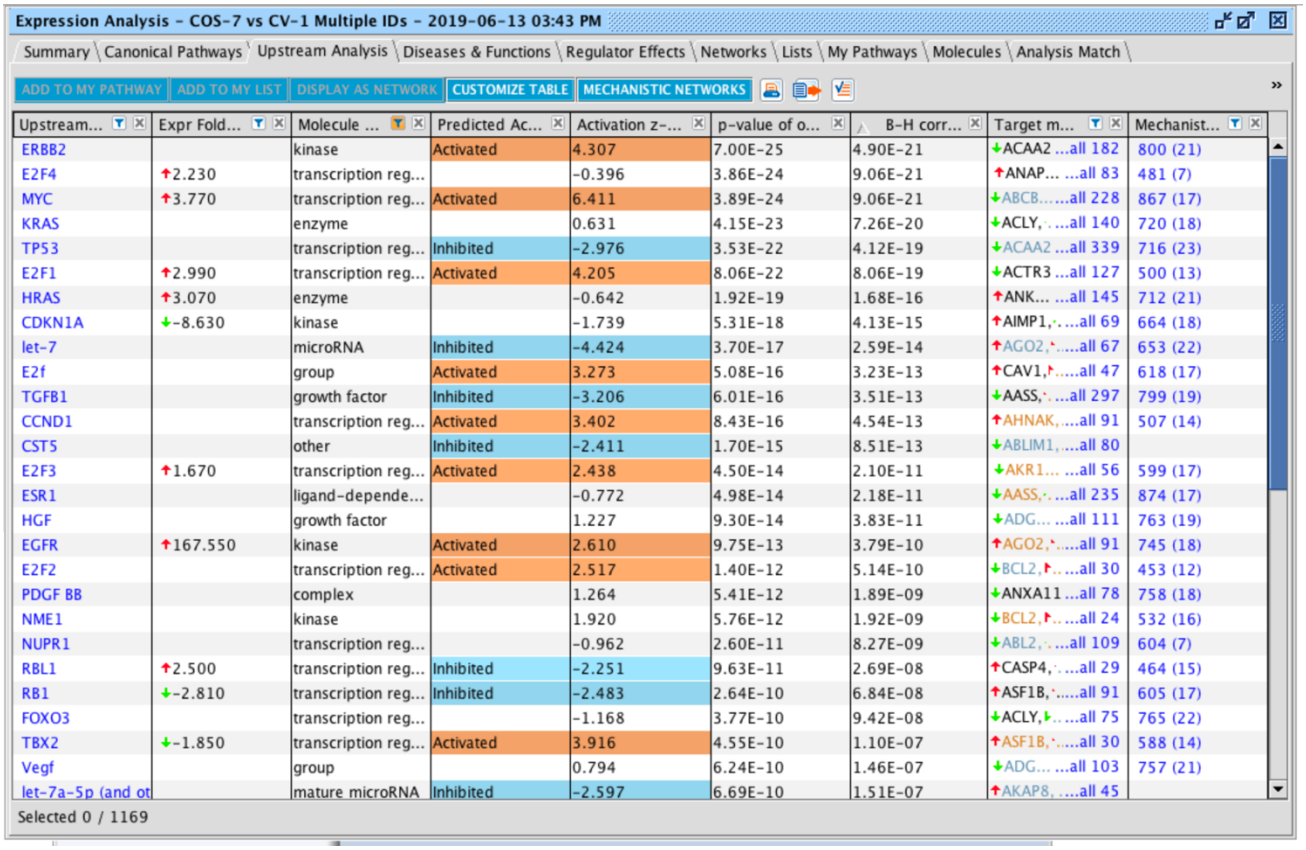


Figure S.11*:* Intracellular localization of the overexpressed TMEFF2 in COS-7 cells (a to f) and CV-1 cells (g to l). Cells were fixed 30 hour post-transfection. Widefield fluorescent micrographs of COS-7 cells: (a) TMEFF2

[Alexa Fluor® 647-Red], (b) DAPI, (c) TMEFF2 and DAPI overlay, (d) TMEFF and DAPI overlay in bright field, (e) bright field, (f) DAPI in bright field. Widefield fluorescent micrographs of CV-1 cells: (g) TMEFF2 [Alexa Fluor® 647-Red] (h) DAPI, (i) TMEFF2 and DAPI overlay, (j) TMEFF and DAPI overlay in bright field, (k) bright field, (l) DAPI in bright field. NB: Imaging was performed by Dr Adam Cawte.



Expression Analysis - COS-7 vs CV-1 Multiple IDs - 2019-06-13 03:43 PM

Summary | Canonical Pathways | Upstream Analysis | Diseases & Functions | Regulator Effects | Networks | Lists | My Pathways | Molecules | Analysis Match

ADD TO MY PATHWAY | ADD TO MY LIST | DISPLAY AS NETWORK | CUSTOMIZE TABLE | MECHANISTIC NETWORKS

Upstream...	Expr Fold...	Molecule ...	Predicted Ac...	Activation z-...	p-value of o...	B-H corr...	Target m...	Mechanist...
ERBB2		kinase	Activated	4.307	7.00E-25	4.90E-21	ACAA2 ...all 182	800 (21)
E2F4	↑2.230	transcription reg...		-0.396	3.86E-24	9.06E-21	ANAP... ...all 83	481 (7)
MYC	↑3.770	transcription reg...	Activated	6.411	3.89E-24	9.06E-21	ABCB... ...all 228	867 (17)
KRAS		enzyme		0.631	4.15E-23	7.26E-20	ACLY, ...all 140	720 (18)
TP53		transcription reg...	Inhibited	-2.976	3.53E-22	4.12E-19	ACAA2 ...all 339	716 (23)
E2F1	↑2.990	transcription reg...	Activated	4.205	8.06E-22	8.06E-19	ACTR3 ...all 127	500 (13)
HRAS	↑3.070	enzyme		-0.642	1.92E-19	1.68E-16	ANK... ...all 145	712 (21)
CDKN1A	↓-8.630	kinase		-1.739	5.31E-18	4.13E-15	AIMP1, ...all 69	664 (18)
let-7		microRNA	Inhibited	-4.424	3.70E-17	2.59E-14	AGO2, ...all 67	653 (22)
E2f		group	Activated	3.273	5.08E-16	3.23E-13	CAV1, ...all 47	618 (17)
TGFB1		growth factor	Inhibited	-3.206	6.01E-16	3.51E-13	AASS, ...all 297	799 (19)
CCND1		transcription reg...	Activated	3.402	8.43E-16	4.54E-13	AHNAK, ...all 91	507 (14)
CST5		other	Inhibited	-2.411	1.70E-15	8.51E-13	ABLIM1, ...all 80	
E2F3	↑1.670	transcription reg...	Activated	2.438	4.50E-14	2.10E-11	AKR1... ...all 56	599 (17)
ESR1		ligand-depende...		-0.772	4.98E-14	2.18E-11	AASS, ...all 235	874 (17)
HGF		growth factor		1.227	9.30E-14	3.83E-11	ADG... ...all 111	763 (19)
EGFR	↑167.550	kinase	Activated	2.610	9.75E-13	3.79E-10	AGO2, ...all 91	745 (18)
E2F2		transcription reg...	Activated	2.517	1.40E-12	5.14E-10	BCL2, ...all 30	453 (12)
PDGF BB		complex		1.264	5.41E-12	1.89E-09	ANXA11 ...all 78	758 (18)
NME1		kinase		1.920	5.76E-12	1.92E-09	BCL2, ...all 24	532 (16)
NUPR1		transcription reg...		-0.962	2.60E-11	8.27E-09	ABL2, ...all 109	604 (7)
RBL1	↑2.500	transcription reg...	Inhibited	-2.251	9.63E-11	2.69E-08	CASP4, ...all 29	464 (15)
RB1	↓-2.810	transcription reg...	Inhibited	-2.483	2.64E-10	6.84E-08	ASF1B, ...all 91	605 (17)
FOXO3		transcription reg...		-1.168	3.77E-10	9.42E-08	ACLY, ...all 75	765 (22)
TBX2	↓-1.850	transcription reg...	Activated	3.916	4.55E-10	1.10E-07	ASF1B, ...all 30	588 (14)
Vegf		group		0.794	6.24E-10	1.46E-07	ADG... ...all 103	757 (21)
let-7a-5p (and ot		mature microRNA	Inhibited	-2.597	6.69E-10	1.51E-07	AKAP8, ...all 45	

Selected 0 / 1169

Figure S.12: Results of the IPA upstream regulator analysis for the transcriptomic differences between wild type COS-7 cells vs wild type CV-1 cells. GeneChip™ Clariom™ D human exome array. Pathway analysis was performed using Ingenuity Pathway Analysis - IPA (QIAGEN®).

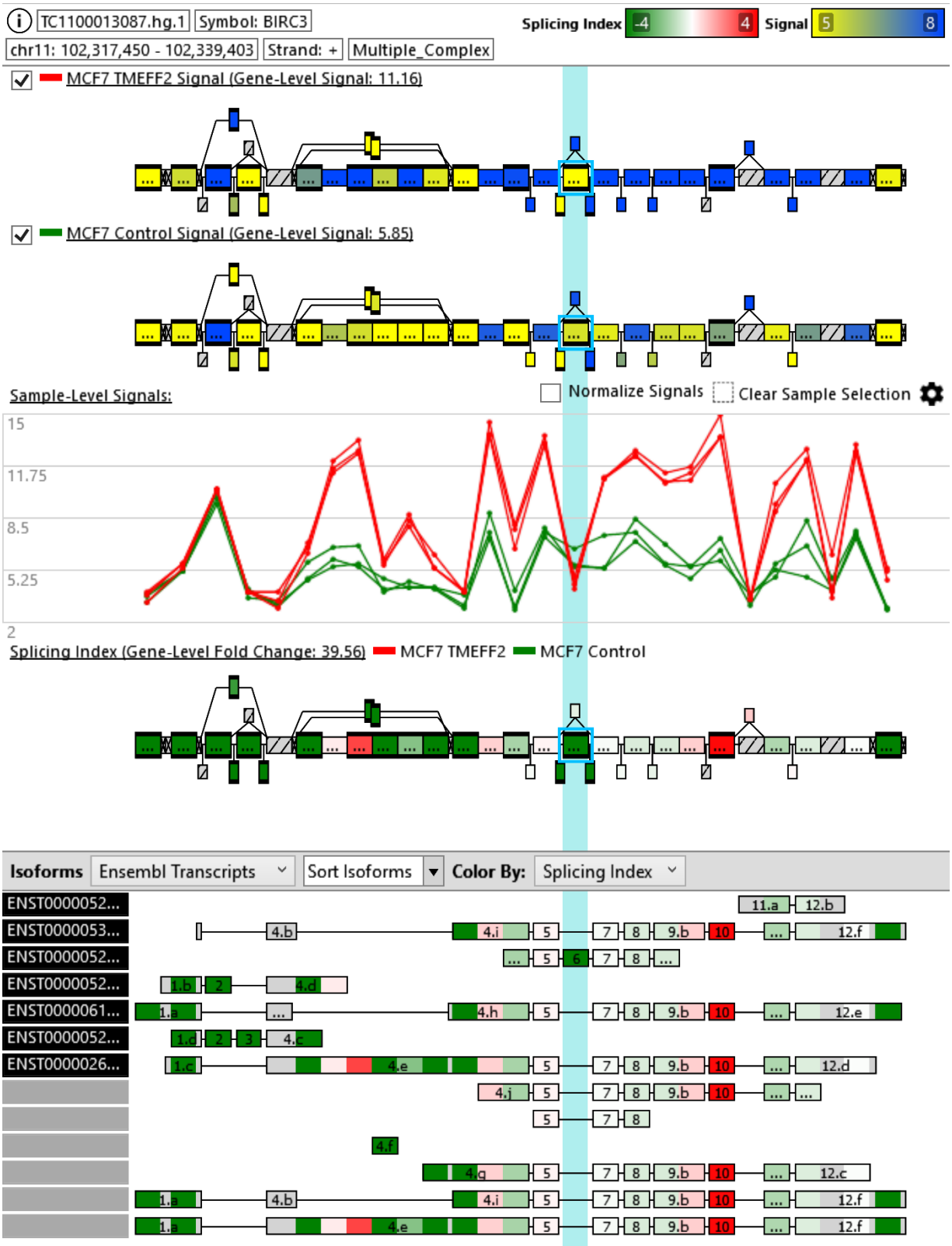


Figure S.13: The alternative-splice variant view for the BIRC3 multiple complex from the transcriptome analysis console (TAC) v.4.0.1.

Permissions

SPRINGER NATURE

Thank you for your order!

Dear Mr. Motasim Masood,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Mr. Motasim Masood
Order Date:	Aug 26, 2019
Order Number:	4656660239195
Publication:	Cell Death and Differentiation
Title:	Stefan Grimm, 1963–2014, a tragic loss for the scientific community
Type of Use:	Thesis/Dissertation
Order Total:	0.00 GBP

View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
https://myaccount.copyright.com

Copyright Clearance Center RightsLink®

SPRINGER NATURE

Thank you for your order!

Dear Mr. Motasim Masood,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Mr. Motasim Masood
Order Date:	Nov 7, 2019
Order Number:	4703810496842
Publication:	Nature Reviews Genetics
Title:	The art and design of genetic screens: mammalian culture cells
Type of Use:	Thesis/Dissertation
Order Ref:	004977862
Order Total:	0.00 GBP

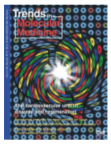
View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
https://myaccount.copyright.com

Copyright Clearance Center RightsLink®



Thank you for your order!

Dear Mr. Motasim Masood,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Mr. Motasim Masood
Order Date:	Nov 4, 2019
Order Number:	4702160462681
Publication:	Trends in Molecular Medicine
Title:	Anticancer genes: inducers of tumour-specific cell death signalling
Type of Use:	reuse in a thesis/dissertation
Order Ref:	00497786
Order Total:	0.00 EUR

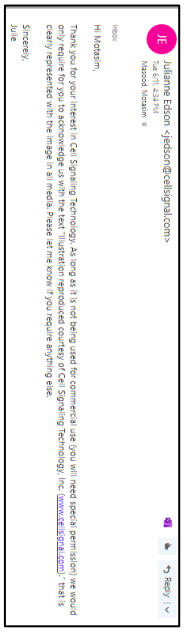
View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
https://myaccount.copyright.com

Copyright Clearance Center RightsLink®


JE
Julianne Edson <jedson@cellsignal.com>
Hi Motasim,
Thank you for your interest in Cell Signaling Technology. As long as it is not being used for commercial use (you will need special permission) we would only require for you to acknowledge us with the text "Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)", that is clearly represented with the image in all media. Please let me know if you require anything else.
Sincerely,
Julie

END