

Investigations into the Functional Significance and Plasticity of O-linked Glycosylation

Joanna Jayne Cull

Degree awarded by Oxford Brookes University

This thesis is submitted in partial fulfilment of the requirements of the award of Doctor of Philosophy

September 2024

Author's Declaration

I, Joanna Jayne Cull, declare that this thesis titled 'Investigations into the Functional Significance and Plasticity of O-linked Glycosylation' and submitted for assessment is an original work of my own. Any works of others used within this thesis, or work presented in it are properly acknowledged. A list of references employed is included.

Acknowledgements

I can't believe this journey that started four years ago in the middle of the COVID-19 pandemic is coming to an end! I am truly grateful for all the opportunities that have been presented to me throughout my PhD, they have shaped me into the scientist and person I am today.

First and foremost, I would like to thank my supervisors, Professor Susan Brooks, Dr Priya Samuel and Dr Ryan Pink. I cannot begin to express my gratitude for your guidance, belief and support over the past four years. Your knowledge, expertise, feedback and experimental understanding have been invaluable, I simply could not have done it without you! A special thank you also goes to all the members of the lab, past and present, for making Oxford Brookes such an enjoyable place to be. In particular, to Ellie, Jamie, Soozana, Natalie, Bethy, Hajrah and Ternisha! In addition, thank you to the Bioimaging team, especially Edward Rea and Dr Alan Marron, for the endless support and guidance. Thank you to the BBSRC and the Oxford Interdisciplinary Bioscience DTP, through which my funding is provided, and without which this work would not have been possible.

Finally, a special thank you to my amazing friends, family and partner. Ant, I could not have done this without you. You have supported me and believed in me every step of the way and I am forever grateful for you. Mum, Dad and Matt thank you for your unconditional support, your unwavering belief in me and for always inspiring me to chase my dreams.

Abstract

Cell surface glycans have important roles in a myriad of biological functions including cell adhesion, signalling and identity. Mucin-type O-linked glycosylation is initiated by the attachment of an N-acetylgalactosamine (GalNAc) to a serine or threonine. This forms an initial structure called Thomsen-nouvelle (Tn) antigen. In healthy tissue, the Tn antigen is cryptic and further elaborated into linear or branched glycans. Failure in O-GalNAc glycan extension results in truncation of O-GalNAc glycans and exposure of O-GalNAc glycans, such as the Tn antigen, which can cause dysregulation of biological functions. For example, O-GalNAc glycans have been associated with poor long-term breast cancer prognosis.

Model breast epithelial cell lines were investigated for their cell surface levels of O-GalNAc glycans. Model epithelial breast cancer cell lines demonstrated significantly higher levels of O-GalNAc glycans compared to the model normal breast epithelial cell line. Levels of O-GalNAc glycans present on the cell surface impacted the observable morphology of a cell. Cell shape changes are key for the epithelial-mesenchymal transition (EMT), which is implicated in wound healing and cancer metastasis. Further investigations into the relationship of EMT with a cell's glycosylation status were conducted.

Cells with high levels of O-GalNAc glycans display a more rounded, epithelial-like morphology, have high levels of E-cadherin (epithelial marker), high adhesive potential to endothelial monolayer in a static adhesion assay, and slower migratory capabilities in wound healing assays. Cells with low levels of O-GalNAc glycans demonstrated an elongated, mesenchymal-like morphology, lower levels of E-cadherin, high levels of vimentin (mesenchymal marker), high migratory capability and less adhesive to endothelial monolayers. Transforming growth factor- β 1 (TGF- β 1) treatment, a known inducer of EMT, of epithelial breast cancer cell lines resulted in the significant reduction of O-GalNAc glycans and E-cadherin. It also resulted in cell morphology changes, to an elongated, mesenchymal-like phenotype.

RNA-sequencing revealed EMT-related genes, relating to mesenchymal markers and EMT signalling pathways, were differentially expressed in O-GalNAc glycan-negative and TGF- β 1 treated cell populations compared to their O-GalNAc glycan-positive counterpart. Differentially expressed gene enrichment analysis also showed enrichment for signalling pathways associated with cancer stem cells and altered proteoglycans and ECM interactions. All these link to the functional migratory differences exhibited in O-GalNAc glycan-negative cells.

The results in this thesis demonstrate the functional significance of a cell's O-GalNAc glycosylation profile in EMT and cell motility, which is of critical importance in cancer metastasis. The exposure or elaboration of O-linked glycan chains also appears to be highly plastic and dynamic allowing cancer cells to access phenotypes optimal for each stage of the metastatic cascade.

Table of Contents

1.	Introduction	2
1.1	Glycosylation of proteins	2
1.1.1	Glycosylation	2
1.1.2	Monosaccharides and nucleotide sugar donors	3
1.1.3	Types of protein glycosylation	4
1.1.4	Mucin-type O-linked protein glycosylation	5
1.1.5	GALNTs.....	14
1.1.6	Control of mucin-type O-linked glycosylation	15
1.2	Lectins	16
1.3	Cancer	17
1.3.1	Breast cancer	17
1.3.2	Glycans as prognostic markers in cancer	19
1.3.3	Hallmarks of cancer.....	19
1.3.4	Metastasis	20
1.3.5	Tumour extracellular matrix.....	22
1.4	Epithelial-mesenchymal transition (EMT)	22
1.4.1	Types of EMT	24
1.4.2	Partial EMT.....	25
1.4.3	Induction of EMT	25
1.4.4	Cell migration	27
1.4.5	Cancer stem cells.....	29
1.5	Project Aims	29
2.	Materials and methods	31
2.1	Cell lines and routine cell culture	31
2.1.1	Epithelial cell lines	31
2.1.2	Primary endothelial cells	31
2.1.3	Re-culturing from frozen cell line stock	33
2.1.4	Maintenance of cell culture	34
2.1.5	Passaging of cell cultures	34
2.1.6	Preparation of frozen cell line stocks	35
2.1.7	Routine mycoplasma testing	35
2.1.8	Cell counting.....	35
2.2	Lectin labelling.....	35
2.2.1	Cell seeding and coverslip preparation	35
2.2.2	Biotinylated lectin cytochemistry with horseradish peroxidase (HRP)	36
2.2.3	Biotinylated lectin cytochemistry with HRP image acquisition and quantification.....	36
2.2.4	Fluorescent lectin labelling	38

2.2.5	Fluorescent lectin labelling image acquisition and quantification	38
2.2.6	Lectin labelling controls	38
2.3	Lectin flow cytometry.....	39
2.3.1	HPA flow cytometry	39
2.3.2	HPA flow cytometry controls	39
2.4	Cell morphology analysis	40
2.5	Cell functional assays	40
2.5.1	Static adhesion assay	41
2.5.2	Static adhesion assay control	41
2.5.3	Static adhesion assay image acquisition and quantification	42
2.5.4	Wound healing assay	43
2.5.5	HPA labelling of wound healing assay	43
2.5.6	E-cadherin, vimentin and HPA labelling of wound healing assay	44
2.5.7	MTT assay.....	45
2.6	EMT marker immunofluorescence	46
2.6.1	E-cadherin, N-cadherin and vimentin immunofluorescence	46
2.6.2	E-cadherin, N-cadherin and vimentin immunofluorescence image acquisition and analysis	47
2.6.3	E-cadherin, N-cadherin and vimentin immunofluorescence controls.....	47
2.7	Western Blotting	48
2.7.1	Protein extraction and quantification from breast epithelial cell lines	48
2.7.2	SDS-PAGE and Western Blotting for EMT markers	49
2.8	TGF- β 1 induced EMT.....	51
2.8.1	TGF- β 1 treatment.....	51
2.8.2	TGF- β 1 treatment and Western Blotting	51
2.8.3	Identifying effects of TGF- β 1 treatment on cell morphology, HPA labelling and EMT markers	51
2.9	Lectin-coated magnetic bead isolation of glycosylation-specific cell populations.....	52
2.9.1	Preparation of cells.....	52
2.9.2	Preparation of HPA-coated magnetic beads.....	53
2.9.3	Separating glycosylation-specific cell populations	53
2.9.4	Dissociation of cells from HPA-coated magnetic beads using competitive inhibition of lectin-binding	53
2.10	Ribonucleic acid (RNA) sequencing	55
2.10.1	Total RNA isolation and DNase treatment	55
2.10.2	Novogene RNA quality control	57
2.10.3	RNA sequencing by Novogene	58
2.10.4	Analysis of sequenced data by Novogene	58
2.11	Statistical analysis	58

3.	Characterisation of model breast epithelial cell lines	61
3.1	Background	61
3.1.1	Altered glycosylation in cancer	61
3.1.2	<i>Helix pomatia</i> agglutinin (HPA) binding as a prognostic marker in cancer	61
3.1.3	<i>Vicia villosa</i> agglutinin (VVA) binding as a prognostic marker in cancer	62
3.2	Aims and Objectives	63
3.3	Methods.....	64
3.3.1	Biotinylated lectin labelling	64
3.3.2	Fluorescent HPA lectin labelling	64
3.3.3	Lectin flow cytometry.....	64
3.3.4	Cell morphology analysis	65
3.4	Results	65
3.4.1	Biotinylated HPA lectin cytochemistry.....	65
3.4.2	Fluorescent HPA lectin labelling	68
3.4.3	VVA lectin cytochemistry.....	71
3.4.4	Cell morphology investigations	74
3.5	Discussion	77
3.5.1	Breast cancer epithelial cell lines synthesize significantly more O-GalNAc glycans than normal breast epithelial cell lines	77
3.5.2	Truncated O-GalNAc glycans are associated with cell morphology	80
3.6	Key findings.....	81
4.	Epithelial-mesenchymal transition and O-GalNAc glycans	83
4.1	Background	83
4.1.1	EMT	83
4.1.2	Functional assays to investigate EMT.....	84
4.1.3	Glycosylation and EMT	86
4.2	Aims and Objectives	87
4.3	Methods.....	87
4.3.1	EMT marker identification	87
4.3.2	Wound healing assay and MTT assay	88
4.3.3	Labelling O-GalNAc glycans in the wound healing assay	88
4.3.4	Labelling E-cadherin and vimentin in the wound healing assay	88
4.3.5	Static adhesion assay	88
4.4	Results	89
4.4.1	EMT marker identification and relationship to O-GalNAc glycans	89
4.4.2	Wound healing assay	93
4.4.3	HPA labelling of the wound healing assay.....	95
4.4.4	E-cadherin labelling of the wound healing assay	99

4.4.5	Static adhesion assay	103
4.5	Discussion	104
4.5.1	O-GalNAc glycans are associated with E-cadherin protein levels	104
4.5.2	O-GalNAc glycans are associated with functional aspects of EMT	105
4.6	Key findings.....	109
5.	Investigating the effect of transforming growth factor- β on O-GalNAc glycans	112
5.1	Background	112
5.1.1	Cytokine-induced EMT <i>in vitro</i>	112
5.1.2	TGF- β pathway	112
5.1.3	Glycosylation changes in cytokine-induced EMT.....	116
5.2	Aims and Objectives	116
5.3	Methods.....	117
5.3.1	TGF- β 1 treatment.....	117
5.3.2	Cell morphology analysis	117
5.3.3	EMT marker analysis	118
5.3.4	Assessment of cellular O-GalNAc glycans following TGF- β 1 treatment.....	118
5.4	Results	119
5.4.1	TGF- β 1 treatment and MCF-7 and T-47D cell morphology	119
5.4.2	TGF- β 1 treatment and EMT marker protein expression	124
5.4.3	TGF- β 1 treatment and O-GalNAc glycans.....	129
5.5	Discussion	132
5.5.1	TGF- β 1 treatment induces an elongated mesenchymal-like cell morphology.....	132
5.5.2	TGF- β 1 treatment affects EMT marker, E-cadherin, protein levels	132
5.5.3	TGF- β 1 treatment reduces cell surface O-GalNAc glycans.....	134
5.6	Key findings.....	135
6.	RNA sequencing.....	138
6.1	Background	138
6.1.1	RNA sequencing	138
6.1.2	Glycosylation and RNA sequencing.....	138
6.2	Aims and Objectives	139
6.3	Methods.....	140
6.3.1	Lectin-coated magnetic bead isolation of glycosylation-specific cell populations.....	140
6.3.2	Total RNA isolation, RNA sequencing and analysis	140
6.4	Results	140
6.4.1	RNA sequencing sample quality overview	140
6.4.2	Gene expression overview.....	141
6.4.3	Principal component analysis (PCA).....	142

6.4.4	Gene co-expression Venn diagrams.....	143
6.4.5	Total number of DEGs across samples	144
6.4.6	MCF-7 HPA-negative and HPA-positive cell differential gene expression analysis ...	145
6.4.7	T-47D HPA-negative and HPA-positive cell differential gene expression analysis	147
6.4.8	MCF-7 TGF- β 1 treated cells and MCF-7 HPA-positive cells differential gene expression analysis	149
6.4.9	KEGG enrichment analysis	150
6.4.10	Mucin-type O-glycan biosynthesis KEGG (hsa00512).....	158
6.4.11	EMT transcription factor DEGs	163
6.5	Discussion	164
6.5.1	MCF-7 HPA-negative DEGs associated with EMT	165
6.5.2	T-47D HPA-negative DEGs associated with EMT	167
6.5.3	TGF- β 1 treated MCF-7 DEGs associated with EMT	168
6.5.4	Mucin-type O-glycan biosynthesis DEGs.....	170
6.6	Key findings.....	174
7.	Final discussion	176
7.1	O-GalNAc glycan profiles of model breast epithelial cell lines.....	178
7.2	Functional implications of mucin-type O-linked glycosylation profiles	180
7.3	Study limitations	182
7.4	Future work	183
7.5	Novel contributions.....	183
8.	References	186
9.	Appendix – Supplementary Materials.....	208

List of Figures

Chapter 1

Figure 1.1 O-linked and N-linked glycosylation.....	4
Figure 1.2 Initiation of mucin-type O-linked glycosylation and synthesis of core structures.	7
Figure 1.3 Mucin-type O-linked glycosylation in the Golgi apparatus.....	8
Figure 1.4 Sialylation of Thomsen nouvelle (Tn) antigen and core 1 glycans.	9
Figure 1.5 Substrate availability can regulate O-glycans created - common competing glycosyltransferases.....	11
Figure 1.6 Elongation pathway of core structures ending in either N-acetylglucosamine (GlcNAc) or Galactose (Gal).	12
Figure 1.7 Type 1 and type 2 chain terminations through the addition of either a fucose (Fuc) or N-acetylneuraminic acid (NeuNAc) monosaccharide.	13
Figure 1.8 Schematic representation of UDP-N- α -D galactosamine: polypeptide N-acetylgalactosaminyltransferases (GALNTs) shared type II membrane structure, located in the Golgi apparatus.	14
Figure 1.9 Multi-level regulation of mucin-type O-linked glycosylation.	16
Figure 1.10 Plant lectin classification based upon their structural diversity.....	17
Figure 1.11 Anatomy of terminal ductal-lobular unit and progression of ductal carcinoma <i>in situ</i> (DCIS) through to invasive ductal carcinoma (IDC).	18
Figure 1.12 Hallmarks of cancer.	20
Figure 1.13 The metastatic cascade.....	21
Figure 1.14 Overview of epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET).....	23
Figure 1.15 Three types of epithelial-mesenchymal transition (EMT).	24
Figure 1.16 Overview of epithelial-mesenchymal transition (EMT) inducing signals, and associated transcriptional regulation.	26
Figure 1.17 Cell migration methods.	28

Chapter 2

Figure 2.1 <i>Helix pomatia</i> agglutinin (HPA) and <i>Vicia villosa</i> agglutinin (VVA) lectin labelling.	37
Figure 2.2 Cell morphology analysis.	40
Figure 2.3 Static adhesion assay.	42
Figure 2.4 Wound healing assay.	43
Figure 2.5 MTT (3-(4,5-dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide) assay.	46
Figure 2.6 Overview of Pierce bicinchoninic acid (BCA) protein assay.	48
Figure 2.7 Western Blotting overview.	50
Figure 2.8 Representative phase contrast images of freehand tool region of interest (ROI) selection overlay of control and transforming growth factor- β 1 (TGF- β 1) treatment conditions in MCF-7 cells..	52

Figure 2.9 Workflow of <i>Helix pomatia</i> agglutinin (HPA)-coated magnetic bead isolation of glycosylation-specific cell populations.	54
Figure 2.10 Total RNA isolation and DNase treatment workflow.	56
Figure 2.11 Workflow of Agilent Technologies 4150 TapeStation measurement of RNA concentration and RNA integrity number equivalent (RINe) value.	57
Figure 2.12 Workflow overview of RNA Sequencing performed by Novogene.	59

Chapter 3

Figure 3.1 <i>Helix pomatia</i> agglutinin (HPA) lectin cytochemistry.	66
Figure 3.2 <i>Helix pomatia</i> agglutinin (HPA) lectin cytochemistry quantification.	67
Figure 3.3 Fluorescent <i>Helix pomatia</i> agglutinin (HPA) lectin labelling.	69
Figure 3.4 Mean corrected total cell fluorescence (CTCF) of <i>Helix pomatia</i> agglutinin (HPA) labelling.	70
Figure 3.5 Optimised <i>Vicia villosa</i> agglutinin (VVA) lectin cytochemistry of 5µg/mL VVA for 5 minutes.	72
Figure 3.6 Optimised <i>Vicia villosa</i> agglutinin (VVA) lectin cytochemistry quantification.	73
Figure 3.7 Morphology examples of MCF-7 and T-47D lectin positive and negative populations.	74
Figure 3.8 <i>Helix pomatia</i> agglutinin (HPA) morphology assessments.	75
Figure 3.9 <i>Vicia villosa</i> agglutinin (VVA) morphology assessments.	76

Chapter 4

Figure 4.1 Levels of E-cadherin in hTERT-HME1, MCF-7 and T-47D cells.	90
Figure 4.2 Levels of vimentin in hTERT-HME1, MCF-7 and T-47D cells.	91
Figure 4.3 Levels of N-cadherin in hTERT-HME1, MCF-7 and T-47D cells.	92
Figure 4.4 Levels of E-cadherin and vimentin protein marker in <i>Helix pomatia</i> agglutinin (HPA) -positive and -negative populations (±SD).	93
Figure 4.5 Wound healing and MTT assays.	94
Figure 4.6 Fluorescent <i>Helix pomatia</i> agglutinin (HPA) labelling of the wound healing assay.	96
Figure 4.7 Regions of interest in fluorescent <i>Helix pomatia</i> agglutinin (HPA) labelling of the wound healing assay.	97
Figure 4.8 Mean fluorescence (AU) of <i>Helix pomatia</i> agglutinin (HPA) labelling at the 'leading edge' of the healing 'wound' and the 'main body' of cells in the wound healing assay.	98
Figure 4.9 E-cadherin immunofluorescence and fluorescent <i>Helix pomatia</i> agglutinin (HPA) labelling of the wound healing assay.	100
Figure 4.10 Regions of interest in E-cadherin immunofluorescence and fluorescent <i>Helix pomatia</i> agglutinin (HPA) labelling of the wound healing assay.	101
Figure 4.11 Mean fluorescence (AU) of E-cadherin labelling at the 'leading edge' of the healing 'wound' and the 'main body' of cells in the wound healing assay.	102

Figure 4.12 Static adhesion assay.....	103
--	-----

Chapter 5

Figure 5.1 Transforming growth factor- β (TGF- β) latent complex formation and activation.....	114
Figure 5.2 Overview of transforming growth factor- β 1 (TGF- β 1) canonical and noncanonical signalling pathways.	115
Figure 5.3 Transforming growth factor- β 1 (TGF- β 1) treatment effect on MCF-7 morphology.	120
Figure 5.4 Transforming growth factor- β 1 (TGF- β 1) treatment morphology assessment in MCF-7 cells.	121
Figure 5.5 Transforming growth factor- β 1 (TGF- β 1) treatment effect on T-47D morphology.	122
Figure 5.6 Transforming growth factor- β 1 (TGF- β 1) treatment morphology assessment in T-47D cells.	123
Figure 5.7 Transforming growth factor- β 1 (TGF- β 1) treatment and E-cadherin immunofluorescence.	125
Figure 5.8 E-cadherin immunofluorescence after transforming growth factor- β 1 (TGF- β 1) treatment.	126
Figure 5.9 Transforming growth factor- β 1 (TGF- β 1) treatment and N-cadherin immunofluorescence.	127
Figure 5.10 Transforming growth factor- β 1 (TGF- β 1) treatment and vimentin immunofluorescence.	128
Figure 5.11 Transforming growth factor- β 1 (TGF- β 1) treatment and Western Blotting for epithelial-mesenchymal transition (EMT) markers.	129
Figure 5.12 Transforming growth factor- β 1 (TGF- β 1) treatment and fluorescent <i>Helix pomatia</i> agglutinin (HPA) lectin labelling.	130
Figure 5.13 Mean corrected total cell fluorescence (CTCF) of <i>Helix pomatia</i> agglutinin (HPA) labelling after transforming growth factor- β 1 (TGF- β 1) treatment.....	131

Chapter 6

Figure 6.1 Sample gene expression distribution box plot of MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive, MCF-7 HPA-negative, T-47D HPA-positive, T-47D HPA-negative and MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells.....	141
Figure 6.2 Principal component analysis (PCA) of MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive, MCF-7 HPA-negative, T-47D HPA-positive, T-47D HPA-negative and MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells.....	142
Figure 6.3 Gene co-expression Venn diagrams, an average of four biological replicates.	143
Figure 6.4 Total differentially expressed genes (DEGs), down-regulated DEG count and up-regulated DEG counts, an average of four biological replicates.	144

Figure 6.5 Volcano plot of significant differentially expressed genes (DEGs) in MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells, an average of four biological replicates.	145
Figure 6.6 Volcano plot of significant differentially expressed genes (DEGs) in T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to T-47D HPA-positive cells, an average of four biological replicates.	147
Figure 6.7 Volcano plot of significant differentially expressed genes (DEGs) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive cells, an average of four biological replicates.	149
Figure 6.8 KEGG enrichment scatter plots of the 20 most significant KEGG pathways for upregulated differentially expressed genes (DEGs), each sample is an average of four biological replicates.	152
Figure 6.9 KEGG enrichment scatter plots of the 20 most significant KEGG pathways for downregulated differentially expressed genes (DEGs), each sample is an average of four biological replicates.	155
Figure 6.10 Mucin-type O-glycan biosynthesis KEGG (hsa00512) identified 6/2,043 differentially expressed genes (DEGs) ($p=0.81$) in MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells.	159
Figure 6.11 Mucin-type O-glycan biosynthesis KEGG (hsa00512) identified 6/577 differentially expressed genes (DEGs) ($p<0.05$) in T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to T-47D HPA-positive cells.	161
Figure 6.12 Mucin-type O-glycan biosynthesis KEGG (hsa00512) identified 16/3,589 differentially expressed genes (DEGs) ($p=0.29$) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive cells.	162
Figure 6.13 Significant differentially expressed genes (DEGs) associated with proteoglycans in cancer KEGG (hsa05205) highlighted from all comparison groups.	166
Figure 6.14 Significant differentially expressed genes (DEGs) associated with signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) highlighted from all comparison groups.	167
Figure 6.15 Differentially expressed genes (DEGs) in mucin-type O-glycan biosynthesis in MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells and T-47D HPA-negative cells compared to their HPA-positive counterpart.	173
Figure 6.16 Differentially expressed genes (DEGs) in mucin-type O-glycan biosynthesis in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive cells.	173

Chapter 7

Figure 7.1 Metastatic cascade: proposed cell and glycan dynamics during the epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET).	177
---	-----

Appendix

Supplementary Figure 1 FIJI (NIH, Bethesda, MD) ImageJ macro script used to aid the analysis of fluorescent <i>Helix pomatia</i> agglutinin (HPA) labelling.	210
Supplementary Figure 2 Representative images of MinError(I) thresholding of the model breast epithelial cell lines used to identify cells from the background.	210
Supplementary Figure 3 Fluorescent <i>Helix pomatia</i> agglutinin (HPA) labelling controls.	211
Supplementary Figure 4 <i>Vicia villosa</i> agglutinin (VVA) lectin cytochemistry concentration optimisation.	212
Supplementary Figure 5 <i>Helix pomatia</i> agglutinin (HPA) lectin cytochemistry of non-permeabilized cells.	213
Supplementary Figure 6 Quantification of <i>Helix pomatia</i> agglutinin (HPA) lectin cytochemistry in non-permeabilized cells.	214
Supplementary Figure 7 <i>Vicia villosa</i> agglutinin (VVA) lectin cytochemistry of non-permeabilized cells.	215
Supplementary Figure 8 Quantification of <i>Vicia villosa</i> agglutinin (VVA) lectin cytochemistry of non-permeabilized cells.	216
Supplementary Figure 9 Comparison of biotinylated lectin cytochemistry of permeabilized and non-permeabilized cells.	217
Supplementary Figure 10 <i>Helix pomatia</i> agglutinin (HPA) flow cytometry gating strategy.	218
Supplementary Figure 11 Average Allophycocyanin-Area (APC-A) <i>Helix pomatia</i> agglutinin (HPA) values ($\pm$ SD) assessed by flow cytometry.	218
Supplementary Figure 12 E-cadherin primary antibody optimisation for epithelial-mesenchymal transition (EMT) marker immunofluorescence.	219
Supplementary Figure 13 Vimentin primary antibody optimisation for epithelial-mesenchymal transition (EMT) marker immunofluorescence.	220
Supplementary Figure 14 N-cadherin primary antibody optimisation for epithelial-mesenchymal transition (EMT) marker immunofluorescence.	221
Supplementary Figure 15 E-cadherin and vimentin immunofluorescence controls.	222
Supplementary Figure 16 N-cadherin immunofluorescence controls.	223
Supplementary Figure 17 Vimentin immunofluorescence and fluorescent <i>Helix pomatia</i> agglutinin (HPA) labelling of the wound healing assay.	224
Supplementary Figure 18 T-47D cell line transforming growth factor- β 1 (TGF- β 1) treatment optimisation – morphology assessment.	225
Supplementary Figure 19 T-47D cell line transforming growth factor- β 1 (TGF- β 1) treatment optimisation – E-cadherin, vimentin and <i>Helix pomatia</i> agglutinin (HPA) assessment.	226
Supplementary Figure 20 Novogene Agilent 5400 RNA sample quality and integrity control – RNA integrity number (RIN) results.	227

Supplementary Figure 21 Novogene RNA sequencing error rate distribution for MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor- β 1 (TGF- β 1) treated, T-47D HPA-positive and T-47D HPA-negative samples.	230
Supplementary Figure 22 Novogene RNA sequencing GC content distribution for MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor- β 1 (TGF- β 1) treated, T-47D HPA-positive and T-47D HPA-negative samples.	231
Supplementary Figure 23 Novogene RNA sequencing sample data filtering for MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor- β 1 (TGF- β 1) treated, T-47D HPA-positive and T-47D HPA-negative samples.	232
Supplementary Figure 24 Novogene RNA sequencing in the genomic region for MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor- β 1 (TGF- β 1) treated, T-47D HPA-positive and T-47D HPA-negative samples.	233
Supplementary Figure 25 Proteoglycans in cancer KEGG (hsa05205) identified 64/2,043 differentially expressed genes (DEGs) ($p < 0.025$) in MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells.	234
Supplementary Figure 26 Proteoglycans in cancer KEGG (hsa05205) identified 119/3,589 differentially expressed genes (DEGs) ($p < 0.025$) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive cells.	237
Supplementary Figure 27 Proteoglycans in cancer KEGG (hsa05205) identified 19/577 differentially expressed genes (DEGs) ($p = 0.07$) in T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to T-47D HPA-positive cells.	241
Supplementary Figure 28 Signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) identified 39/2,043 differentially expressed genes (DEGs) ($p < 0.05$) in MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells.	243
Supplementary Figure 29 Signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) identified 62/3,589 differentially expressed genes (DEGs) ($p < 0.05$) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive cells.	245
Supplementary Figure 30 Signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) identified 17/577 differentially expressed genes (DEGs) ($p < 0.025$) in T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative cells compared to T-47D HPA-positive cells.	248

List of Tables

Chapter 1

Table 1.1 The seven monosaccharides used to create human glycans and associated symbol system.	3
Table 1.2 Nucleotide sugar donor substrates utilised in mucin-type O-linked glycosylation.	4

Chapter 2

Table 2.1 Characteristics of epithelial cell lines.	32
Table 2.2 Details of cell line complete growth medium and supplements used in this work.	33

Chapter 3

Table 3.1 Lectin binding specificities.	63
--	----

Chapter 4

Table 4.1 Common epithelial-mesenchymal transition (EMT) markers.	83
Table 4.2 Evaluation of some functional assays commonly used to investigate epithelial-mesenchymal transition (EMT).....	85

Chapter 6

Table 6.1 Top 10 significant differentially expressed genes (DEGs) and fold-changes between MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative compared to MCF-7 HPA-positive cells.	146
Table 6.2 Top 10 significant differentially expressed genes (DEGs) and fold-changes between T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative compared to T-47D HPA-positive cells.	148
Table 6.3 Top 10 significant differentially expressed genes (DEGs) and fold-changes between MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells and MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-positive cells.	150
Table 6.4 MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells mucin-type O-glycan biosynthesis KEGG (hsa00512) differentially expressed genes (DEGs).	160
Table 6.5 T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative cells mucin-type O-glycan biosynthesis KEGG (hsa00512) differentially expressed genes (DEGs).	160
Table 6.6 MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cell mucin-type O-glycan biosynthesis KEGG (hsa00512) differentially expressed genes (DEGs).	163
Table 6.7 Epithelial-mesenchymal transition (EMT)-related transcription factors differentially expressed genes (DEGs) identified in MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative, T-47D HPA-negative and MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells.	164

Appendix

Supplementary Table 1 List of items and reagents with accompanying catalogue number and company supplier.	208
Supplementary Table 2 Novogene Agilent 5400 RNA sample quality control result summary information.....	228
Supplementary Table 3 Overview of sample sampling data quality.	229
Supplementary Table 4 MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells proteoglycans in cancer KEGG (hsa05205) differentially expressed genes (DEGs).....	236
Supplementary Table 5 MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells proteoglycans in cancer KEGG (hsa05205) differentially expressed genes (DEGs).....	239
Supplementary Table 6 T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative cells proteoglycans in cancer KEGG (hsa05205) differentially expressed genes (DEGs).....	240
Supplementary Table 7 MCF-7 <i>Helix pomatia</i> agglutinin (HPA)-negative cells signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) differentially expressed genes (DEGs).	244
Supplementary Table 8 MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) differentially expressed genes (DEGs).	246
Supplementary Table 9 T-47D <i>Helix pomatia</i> agglutinin (HPA)-negative cells signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) differentially expressed genes (DEGs).	247

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
α -GalNAc	α -N-acetyl-D-galactosamine
AMT	Amoeboid-mesenchymal transition
APC-A	Allophycocyanin-Area
Asn	Asparagine
BCA	Bicinchoninic acid
BMP	Bone morphogenetic proteins
bp	Base pair
BSA	Bovine serum albumin
C1GALT1	Core 1 beta 1,3-galactosyltransferase
CD44	Cluster of differentiation 44
CMP-NeuNAc	Cytidine 5'-monophospho- β -neuraminic acid
CO ₂	Carbon dioxide
COP	Coat protein complex
Co-SMAD	Common SMAD
ConA	Concanavalin A
COSMC	Core 1 beta 1,3-galactosyltransferase specific molecular chaperone
CTCF	Corrected total cell fluorescence
DAB	3,3'-diaminobenzidine tetrahydrochloride hydrate
DAVID	The Database for Annotation, Visualisation and Integrated Discovery
DCIS	Ductal carcinoma <i>in situ</i>
DEGs	Differentially expressed genes
dH ₂ O	Distilled water
DMEM	Dulbecco's Modified Eagle Medium

DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
ECACC	European Collection of Authenticated Cell Cultures
ECL	Enhanced Chemiluminescence
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
eGFP	Enhanced green fluorescent protein
EMT	Epithelial-mesenchymal transition
ERK8	Extracellular signal regulated kinase 8
ER	Endoplasmic reticulum
F12	Nutrient Mixture F12
FACS	Fluorescence-activated cell sorting
FCS	Foetal calf serum
FGF	Fibroblast growth factor
FPKM	Fragments per kilobase of transcript per million mapped reads
Fuc	Fucose
Gal	Galactose
GalNAc	N-acetylgalactosamine
GALNTs	UDP-N- α -D galactosamine: polypeptide N-acetylgalactosaminyltransferases
GDF	Growth differentiation factors
GDP-Fuc	Guanosine 5'-diphospho- β -L-fucose
Glc	Glucose
GlcNAc	N-acetylglucosamine
GPI	Glycosylphosphatidylinositol
H ₂ O ₂	Hydrogen peroxide
HCl	Hydrochloric acid

HGF	Hepatocyte growth factor
HPA	<i>Helix pomatia</i> agglutinin
HRP	Horseradish peroxidase
hTERT	Human telomerase reverse transcriptase
HPTS	8-Hydroxypyrene-1,3,6-trisulfonic acid
HUVEC	Human umbilical vein endothelial cells
IDC	Invasive ductal carcinoma
IGF	Insulin-like growth factor
IMS	Industrial methylated spirits
IL-6	Interleukin-6
IL-8	Interleukin-8
ILC	Invasive lobular carcinoma
I-SMAD	Inhibitory SMAD
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LAP	Latency associated peptide
LCIS	Lobular carcinoma <i>in situ</i>
LLC	Large latent complex
LSM	Laser scanning microscope
MAA	<i>Maackia amurensis</i> agglutinin
Man	Mannose
ManNAc	N-acetylmannosamine
MAT	Mesenchymal-amoeboid transition
MET	Mesenchymal-epithelial transition
MicroRNAs	miRNAs
MMPs	Matrix metalloproteinases
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide
NGS	Next-generation sequencing

NeuNAc	Sialic acid
OST	Oligosaccharyltransferase
p	Probability
PBS	Phosphate-buffered saline
PC1	First principal component
PC2	Second principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
pEMT	Partial EMT
PFA	Paraformaldehyde
PNA	Peanut agglutinin
Pro	Proline
PTS	Proline, Threonine and Serine sequence repeats
PVDF	Polyvinylidene difluoride
Q-Q	Quantile-Quantile
RIN	RNA integrity number
RINe	RNA integrity number equivalent
RIPA	Radioimmunoprecipitation assay
RPKM	Reads per kilobase per million
RNA	Ribonucleic acid
RPMI-1640	Roswell Park Memorial Institute Medium-1640
ROIs	Regions of interest
R-SMADs	Receptor-regulated SMADs
RTK	Receptor tyrosine kinase
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Ser	Serine
shRNA	Short hairpin RNA

SLC	Small latent complexes
SNA	<i>Sambucus nigra</i> agglutinin
STAT3	Signal transducer and activator of transcription 3
sTn	Sialyl-Thomsen nouvelle
TBS	Tris-buffered saline
TCGA	The Cancer Genome Atlas
TGF- β	Transforming growth factor-beta
TGS	Tris-Glycine-SDS buffer
Thr	Threonine
Tn	Thomsen nouvelle antigen
TNF- α	Tumour necrosis factor-alpha
UDP	Uridine diphosphate
UDP-Gal	Uridine 5'-diphospho- α -galactose
UDP-GalNAc	Uridine 5'-diphospho- α -D-N-acetylgalactosamine
UDP-GlcNAc	Uridine 5'-diphospho- α -D-N-acetylglucosamine
VVA	<i>Vicia villosa</i> agglutinin
WGA	Wheat germ agglutinin

Units of measurement

A	Ampere
°C	Degrees Celsius
hr	Hour
μ g	Microgram
μ L	Microlitre
μ m	Micrometre
mg	Milligram
mL	Millilitre

mm	Millimetre
mM	Millimolar
ng	Nanogram
nm	Nanometre
g	G-force
RPM	Revolutions per minute
V	Volt
v/v	Volume per volume
w/v	Weight per volume

Chapter 1

Introduction

1. Introduction

1.1 Glycosylation of proteins

1.1.1 Glycosylation

Glycosylation is a dynamic, heterogeneous and complex form of post-translational modification that is common to all eukaryotic cells. It is estimated that more than half of all proteins found in humans will have been glycosylated [1]. Glycosylation results in the covalent addition of glycans, produced by the coordinated, sequential enzymatic addition and removal of monosaccharides, to both proteins (to form glycoproteins, proteoglycans and glycosylphosphatidylinositol (GPI)-anchored proteins) and lipids (to form glycolipids). This thesis will focus on protein glycosylation. In eukaryotic cells, protein glycosylation primarily occurs in the secretory pathway (endoplasmic reticulum (ER) and Golgi apparatus) and the cytoplasm. Monosaccharides are linked together through glycosidic bonds between the anomeric carbon of one monosaccharide and the hydroxyl group of the other, in either an α or β orientation (reviewed by [2]). Glycans can range greatly in size and shape, from a single monosaccharide to something far more complex, comprising multiple monosaccharides in a linear or branching structure (reviewed by [3]).

Glycans produced as a result of glycosylation therefore are abundant, diverse and are fundamental to a range of biological functions. These can be divided into three broad categories: i) structural functions, ii) functional regulation, and iii) information carriers as ligands. Glycans at the cell surface form the glycocalyx, a dense layer of membrane-bound glycoconjugates, often extending to more than 30nm into the extracellular matrix (ECM). It is in prime position to provide a protective barrier, mediate cell adhesion and facilitate the identification of 'self' vs 'non-self' involved in innate immune system responses (reviewed by [2]).

The complete repertoire of glycoconjugates in a single cell is called the glycome and is estimated to be up to 10^4 times larger than the proteome (reviewed by [4]). Glycans, unlike proteins, are not primary template-driven products of gene expression; instead they are formed on protein or lipid scaffolds by enzymes, substrates and chaperones that are themselves coded for by the mammalian genome, accounting for approximately 3-4% of all human genes [5]. Glycosylation results in a flexible and dynamic means of mediating protein function that is not under direct genetic control. However, understanding glycosylation is technically complex as we cannot directly predict glycan sequences from a DNA template or genetically delete glycans to understand function. In addition, they demonstrate staggering structural complexity and heterogeneity, owing to the anomericity of their constituent monosaccharides, the variety of monosaccharides from which they are composed, and

the diverse, and sometimes large and branched structures that result. Understanding of both their function and control has lagged behind that of other molecules (reviewed by [6]).

1.1.2 Monosaccharides and nucleotide sugar donors

Glycosyltransferases catalyse the transfer of monosaccharides from an activated donor nucleotide sugar to the hydroxyl group of an acceptor molecule. The monosaccharides are typically salvaged from degraded glycans or created from other sugars present in the cell by enzymes (reviewed by [7]). Seven monosaccharides: glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), fucose (Fuc), sialic acid (NeuNAc) and mannose (Man) are the main constituents of glycans in humans, see Table 1.1. Of these, Gal, GalNAc, GlcNAc, Fuc and NeuNAc are represented in mucin-type O-linked glycosylation (reviewed by [8]). Monosaccharides are activated to a high-energy donor, a process that occurs mostly in the cytoplasm, and then transported into the Golgi apparatus and ER by nucleotide sugar donors to be utilised in glycosylation. Nucleotide sugar donors are actively transported via an antiporter that reciprocally transports free sugar-nucleotides, into the ER and Golgi apparatus after synthesis in the cytosolic or nuclear compartments for usage in glycosylation pathways [7]. Nucleotide sugar donor substrates utilised in mucin-type O-linked glycosylation are listed in Table 1.2.

Table 1.1 The seven monosaccharides used to create human glycans and associated symbol system. Adapted from [9].

Name of Monosaccharide	Abbreviation	Symbol
β -D-Glucose	Glc	
β -D-Mannose	Man	
β -D-Galactose	Gal	
β -D-N-Acetylglucosamine	GlcNAc	
β -D-N-Acetylgalactosamine	GalNAc	
N-Acetylneuraminic acid (sialic acid)	NeuNAc	
β -L-Fucose	Fuc	

Table 1.2 Nucleotide sugar donor substrates utilised in mucin-type O-linked glycosylation. Adapted from [10].

Nucleotide Sugar Donor Substrates Utilised in Mucin-Type O-Linked Glycosylation	
Uridine 5'-diphospho- α -galactose	UDP-Gal
Uridine 5'-diphospho- α -D-N-acetylglucosamine	UDP-GlcNAc
Uridine 5'-diphospho- α -D-N-acetylgalactosamine	UDP-GalNAc
Cytidine 5'-monophospho- β -neuraminic acid	CMP-NeuNAc
Guanosine 5'-diphospho- β -L-fucose	GDP-Fuc

1.1.3 Types of protein glycosylation

The two most common forms of protein glycosylation, forming glycoproteins, in eukaryotes are N-linked and O-linked glycosylation. These are distinguished via the sugar-protein linkage, the initial monosaccharide linked to the proteins, and the unique initiating enzyme, see Figure 1.1. There are several types of O-linked protein glycosylation, which include mucin-type (reviewed by [11]), O-GlcNAc (reviewed by [12]), O-Man (reviewed by [13]), O-Fuc (reviewed by [14]) and O-Glc (reviewed by [15]) linkages to either a Serine (Ser) or Threonine (Thr) in a polypeptide chain. This thesis will focus on mucin-type O-linked glycosylation, which is discussed below in '1.1.4 Mucin-type O-linked protein glycosylation'.

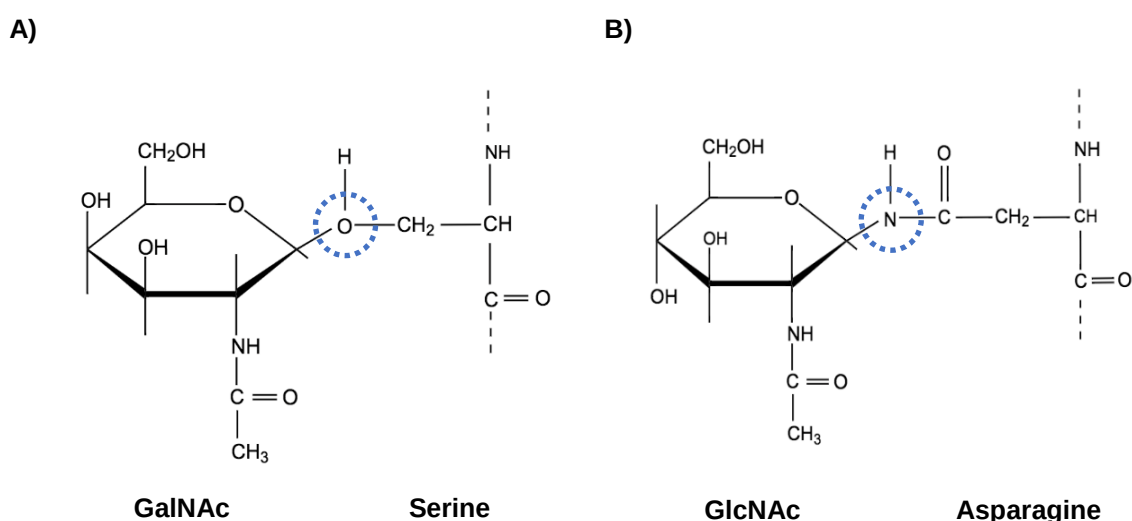


Figure 1.1 O-linked and N-linked glycosylation. A) O-linkage – an N-acetylgalactosamine (GalNAc) residue is added to the hydroxyl group of a serine (Ser) or threonine (Thr) amino acid in a polypeptide chain. B) N-linkage – an N-acetylglucosamine (GlcNAc) residue is added to an asparagine (Asn) amino acid in a polypeptide chain.

In mucin-type O-linked glycosylation, an N-acetylgalactosamine (GalNAc) residue is added to the hydroxyl group of Ser/Thr residue on the polypeptide [16]. It is initiated in, and glycans elongated, in the Golgi apparatus. During recycling of cell membrane glycoproteins, further glycosylation, notably sialylation, can also occur [17, 18]. Whereas, in N-linked glycosylation, an N-acetylglucosamine (GlcNAc) residue is added to an asparagine (Asn) residue within the consensus sequence Asn-X-Ser/Thr where X can be any amino acid except proline (Pro). It is a co-translational event initiated in the cytoplasmic side of the ER onto a lipid-linked dolichol carrier where there is a stepwise transfer of two GlcNAc and five Man residues (reviewed by [19] and [20]). This is then 'flipped' to the luminal side of the ER by the enzyme flippase where it is further extended through the addition of four Man and three Glc residues to form the intermediate $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$. Aided by the action of the enzyme oligosaccharyltransferase (OST), this 14-sugar intermediate is transferred to the Asn on a polypeptide [21]. Once transferred, the oligosaccharide intermediate $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ is trimmed by three sequential steps. Firstly, α -glucosidase I cleaves a terminal Glc residue, this is followed by the action of α -glucosidase II which cleaves a further two Glc residues leaving the intermediate $\text{Man}_9\text{GlcNAc}_2$. Finally, terminal Man is removed via α -mannosidase to yield $\text{Man}_8\text{GlcNAc}_2$ which leaves the ER and enters the cis-Golgi apparatus. Within the cis, medial and trans cisternae of the Golgi apparatus are glycosidases and glycosyltransferases which can extend or cleave the N-linked oligosaccharide. However, all N-linked oligosaccharides contain a common trimannosyl core, $\text{Man}_3\text{GlcNAc}_2\text{-Asn-}$. Depending on the order, type and linkages of the monosaccharide residues attached to this trimannosyl core, the final heterogeneous structures can be grouped into three main groups: high mannose type, complex type and hybrid type [22].

1.1.4 Mucin-type O-linked protein glycosylation

Mucin-type O-linked glycosylation is a wholly post-translational modification of proteins in their tertiary structure that takes place in the Golgi apparatus. Unlike N-linked glycosylation, there is no requisite for a consensus sequence or a lipid-linked dolichol precursor for the initiation of glycosylation. However, the Ser/Thr residues that are glycosylated tend to lie in Pro-rich sequences, in areas of low hydrophobicity and in areas where there is low steric hindrance due to large neighbouring amino acids (reviewed by [23]). Mucin-type O-glycans demonstrate greater structural diversity than N-glycans due to a lack of trimming steps and glycans can range from one monosaccharide long to 20 monosaccharides long [24]. Mucins are large transmembrane and secreted proteins that are the principal component of mucus and are heavily glycosylated in this way due to their high density of Pro, Thr and Ser sequence (PTS) repeats, resulting in O-glycans accounting for 50-80% of their molecular weight and this is therefore where this form of O-glycosylation derives its name [11]. However, these glycans are not only found on mucins, it is estimated that over 85% of all proteins in the secretory pathway are thought to be glycosylated in this way [25].

Mucin-type O-linked glycosylation is initiated in the cis-Golgi apparatus through the addition of a single GalNAc monosaccharide from a nucleotide sugar donor, UDP-GalNAc, via an α -glycosidic bond to the hydroxyl group of either a Ser or Thr in a polypeptide [26-28]. This forms the Thomsen nouvelle (Tn) antigen (GalNAc- α -Ser/Thr), see Figure 1.2. This process is catalysed by a family of 20 isoenzymes called UDP-N- α -D galactosamine: polypeptide N-acetylgalactosaminyltransferases (GALNTs), which are discussed in section '1.1.5 GALNTs' [29]. Mucin-type O-linked glycosylation occurs in a step-wise manner, so the formation of Tn antigen always represents the first step [30]. Stepwise elongation and modification of C3 or C6 of GalNAc in the Tn antigen is performed by specific glycosyltransferases to form a further eight core structures, and this occurs in the medial and trans-Golgi apparatus, see Figure 1.2 and 1.3.

If elongation does not occur, then truncated O-glycans, like Tn antigen, can be present (reviewed by [20]). The Tn antigen itself can also be modified by a family of α 2-6 sialyltransferases (ST6GALNAC1-ST6GALNAC4) to form another truncated O-GalNAc glycan, by α 2-6 sialylation, sialyl-Tn (sTn) (Sia α 2-6GalNAc- α -Ser/Thr), which is a chain terminating event, see Figure 1.4A. The Tn antigen only remains exposed in its unelaborated form at high levels during embryogenesis and in pathologies, such as cancer. It is always further elaborated in healthy, adult tissues, with cores 1 and 2 being the most common type of core structure, found almost ubiquitously on glycoproteins [11]. Cores 3 and 4 are typically only found on glycoproteins in specific tissues, such as the gastrointestinal tract [31, 32] and bronchial tissues [33] but, are still considered to be prevalent enough to be included in the four major subtypes alongside core 1 and 2. Cores 5-8 are deemed to be rare forms of core structures only described in a limited number of tissues. For example, core 5 has been described in the human intestine [34] and colon [35], core 6 in meconium mucins [36], core 7 in bovine submaxillary mucins [37], and core 8 in human bronchial tissues [38].

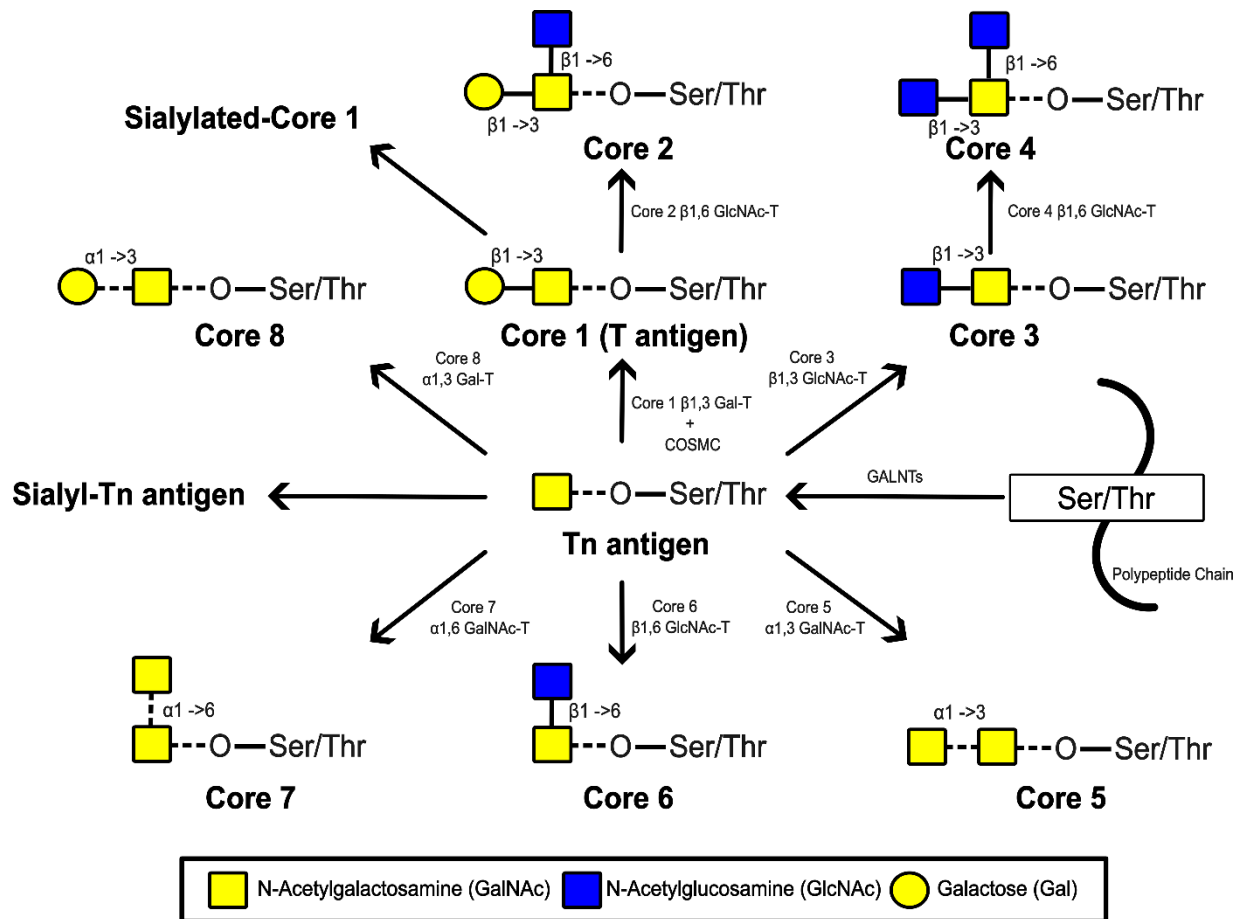


Figure 1.2 Initiation of mucin-type O-linked glycosylation and synthesis of core structures. Mucin-type O-linked glycosylation is initiated through the addition of an N-acetylgalactosamine (GalNAc) monosaccharide onto a serine (Ser) or threonine (Thr) in a polypeptide chain, catalysed by the UDP-N- α -D galactosamine: polypeptide N-acetylgalactosaminyltransferases (GALNTs) to form the Thomsen nouvelle (Tn) antigen (GalNAc- α -Ser/Thr). In healthy adult cells, the Tn antigen is further elaborated by the action of specific glycosyltransferases to form one of eight core structures, which themselves can be extended to form more complex glycans. Adapted from [39].

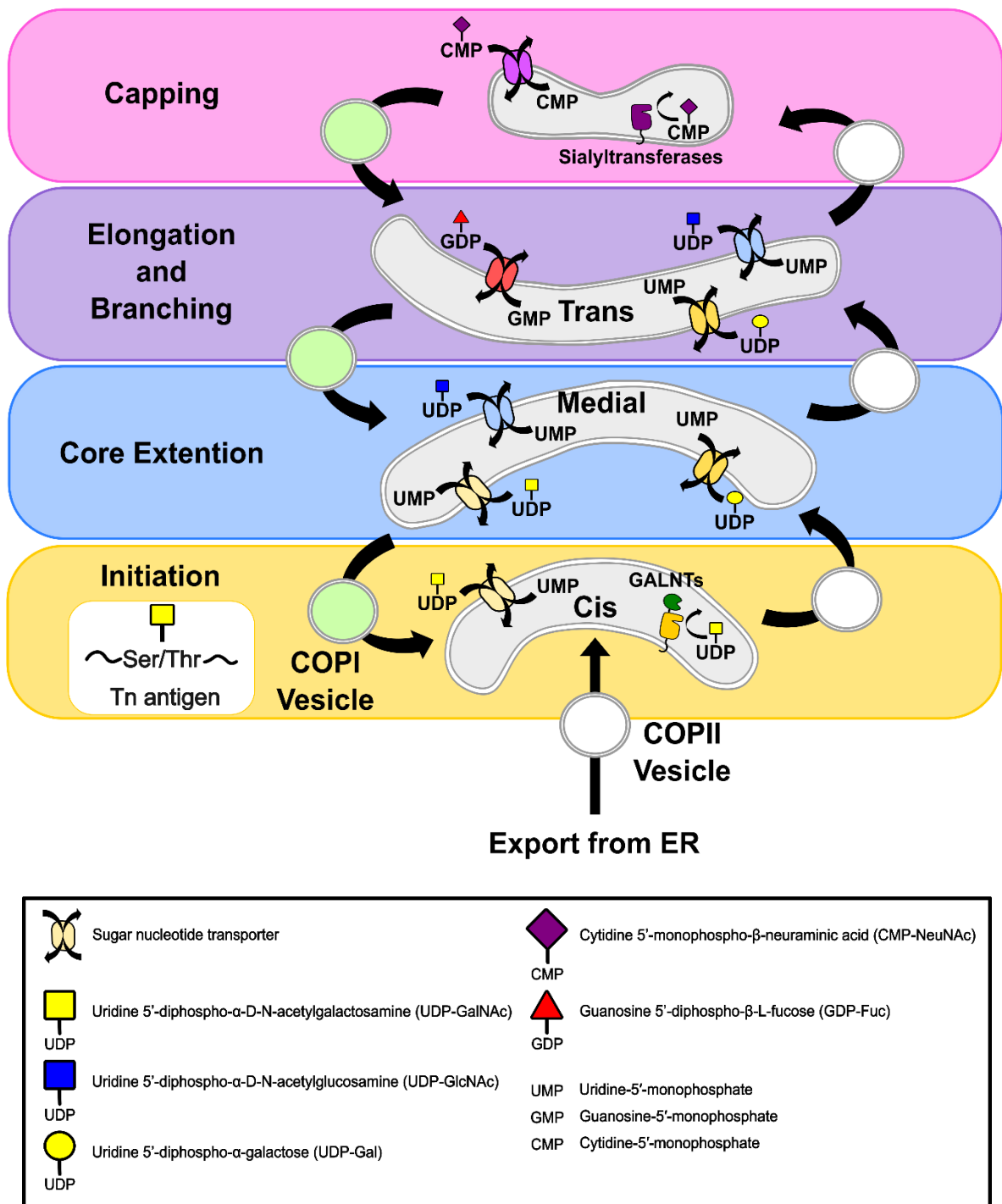


Figure 1.3 Mucin-type O-linked glycosylation in the Golgi apparatus. Initiation of mucin-type O-linked glycosylation occurs in the cis-Golgi apparatus. Core glycan extension, elongation and capping occur throughout the medial and trans-Golgi apparatus. Coat protein complex (COP)-I and -II vesicular transport is responsible for the retrograde of Golgi-resident glycosyltransferases, and transport of ER to Golgi apparatus, respectively. Adapted from [40].

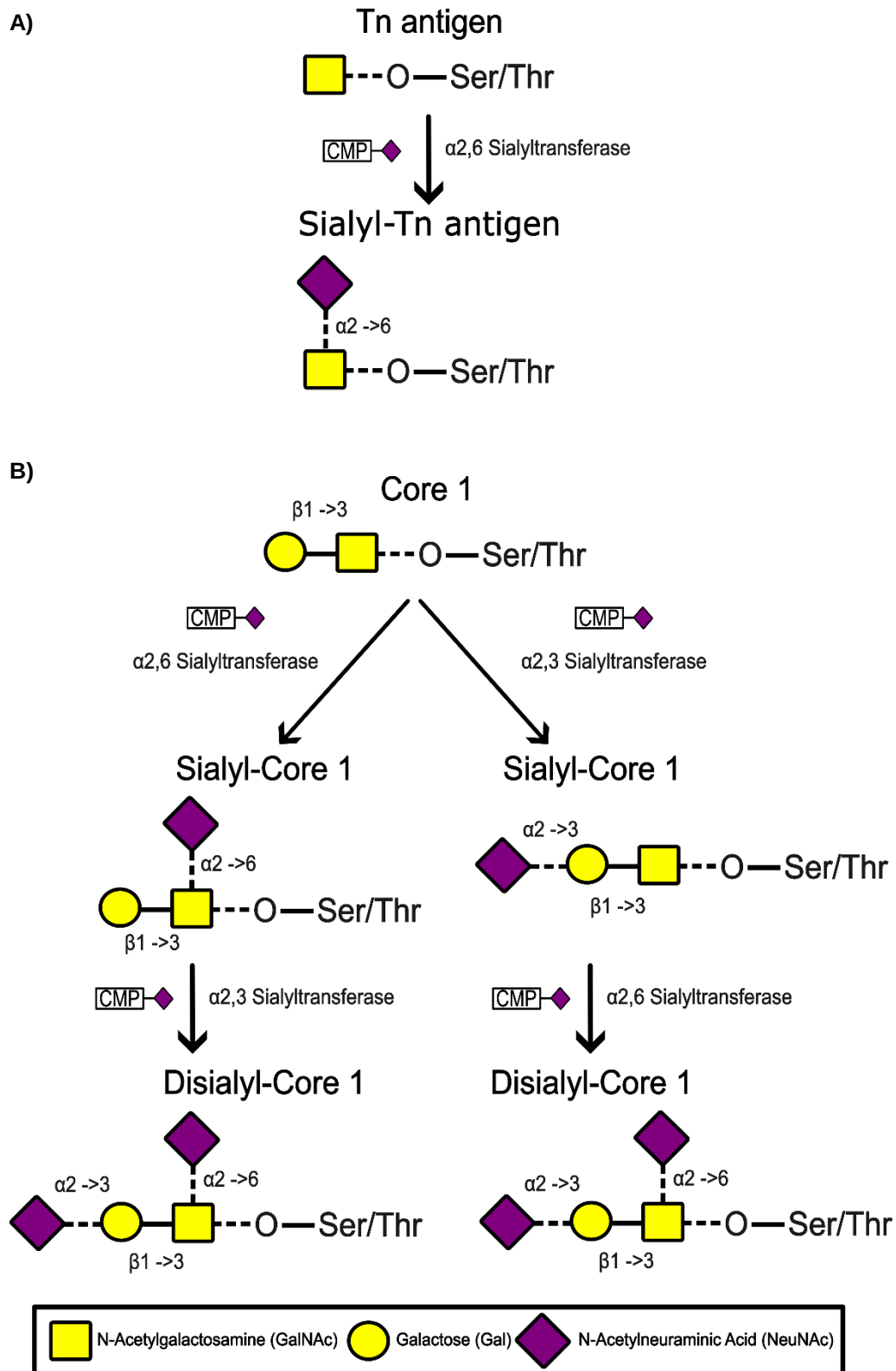


Figure 1.4 Sialylation of Thomsen nouvelle (Tn) antigen and core 1 glycans. A) Tn antigen can be elaborated by $\alpha 2$ -6 sialyltransferases, adding an N-acetylneuraminic acid (NeuNAc) to yield sialyl-Tn antigen. B) Core 1 glycans can be elaborated by either $\alpha 2$ -6 sialyltransferases or $\alpha 2$ -3 sialyltransferases, adding a NeuNAc to yield sialyl-core 1 or further to yield disialyl-core 1. Adapted from [10].

A Gal monosaccharide can be added by core 1 beta 1,3-galactosyltransferase (T synthase or C1GALT1) to Tn antigen to form core 1 (Gal β 1-3GalNAc) also known as the T antigen or Thomson–Friedenrich antigen, see Figure 1.2 (reviewed by [23]). C1GALT1 requires an ER molecular chaperone for correct folding, core 1 β 3-galactosyltransferase specific molecular chaperone (COSMC) encoded by the X-linked (Xq23) C1GALT1C1 gene [41]. COSMC itself does not possess any enzymatic activity [42]. Mutation and loss of COSMC results in the transportation of C1GALT1 into the cytoplasm where it is ubiquitinated and degraded [43, 44]. Core 1 glycans can be sialylated to form sialyl-core 1 or disialyl-core 1 by either family of α 2-3 sialyltransferases, mainly ST3GAL1, or α 2-6 sialyltransferases, see Figure 1.4B.

A GlcNAc monosaccharide can be added by core 2 β 1,6 GlcNAc-T to core 1 to form core 2 (Gal β 1-3(GlcNAc β 1-6)GalNAc). The Tn antigen can also be elaborated through the addition of a GlcNAc monosaccharide by core 3 β 1,3 GlcNAc-T to form core 3 (GlcNAc β 1-3GalNAc). A GlcNAc monosaccharide can then be added to core 3 by core 4 β 1,6 GlcNAc-T to form core 4 (GlcNAc β 1-3(GlcNAc β 1-6)GalNAc). Other ways the Tn antigen can be elaborated is through the addition of a GalNAc monosaccharide by core 5 α 1,3 GalNAc-T to form core 5 (GalNAc α 1-3GalNAc) or a GlcNAc monosaccharide can be added by core 6 β 1,6 GlcNAc-T to form core 6 (GlcNAc β 1-6GalNAc). A GalNAc monosaccharide can be added to the Tn antigen by core 7 α 1,6 GalNAc-T to form core 7 (GalNAc α 1-6GalNAc). Finally, a Gal monosaccharide can be added to the Tn antigen by an α 1,3 Gal-T to form core 8 (Gal α 1-3GalNAc) (reviewed by [45]). Many glycosyltransferases compete for glycans, such as Tn antigen or core 1, which act as substrates for further elongation, see Figure 1.5 for an overview of the most commonly competing enzymes.

These core structures can also be further modified by sialylation, acetylation, and fucosylation. Some examples of extended complex O-glycans include those featuring the ABO and Lewis blood group determinants (reviewed by [40]). Terminal GlcNAc monosaccharides in core structures can be extended by either type 1 chain (Gal with a β 1-3 linkage) or type 2 chain (Gal with β 1-4 linkage), see Figure 1.6A. Whereas, terminal Gal monosaccharides can be elongated through the addition of GlcNAc, see Figure 1.6B. Both type 1 and type 2 chains can then be terminated by the addition of either a Fuc or NeuNAc monosaccharide, see Figure 1.7.

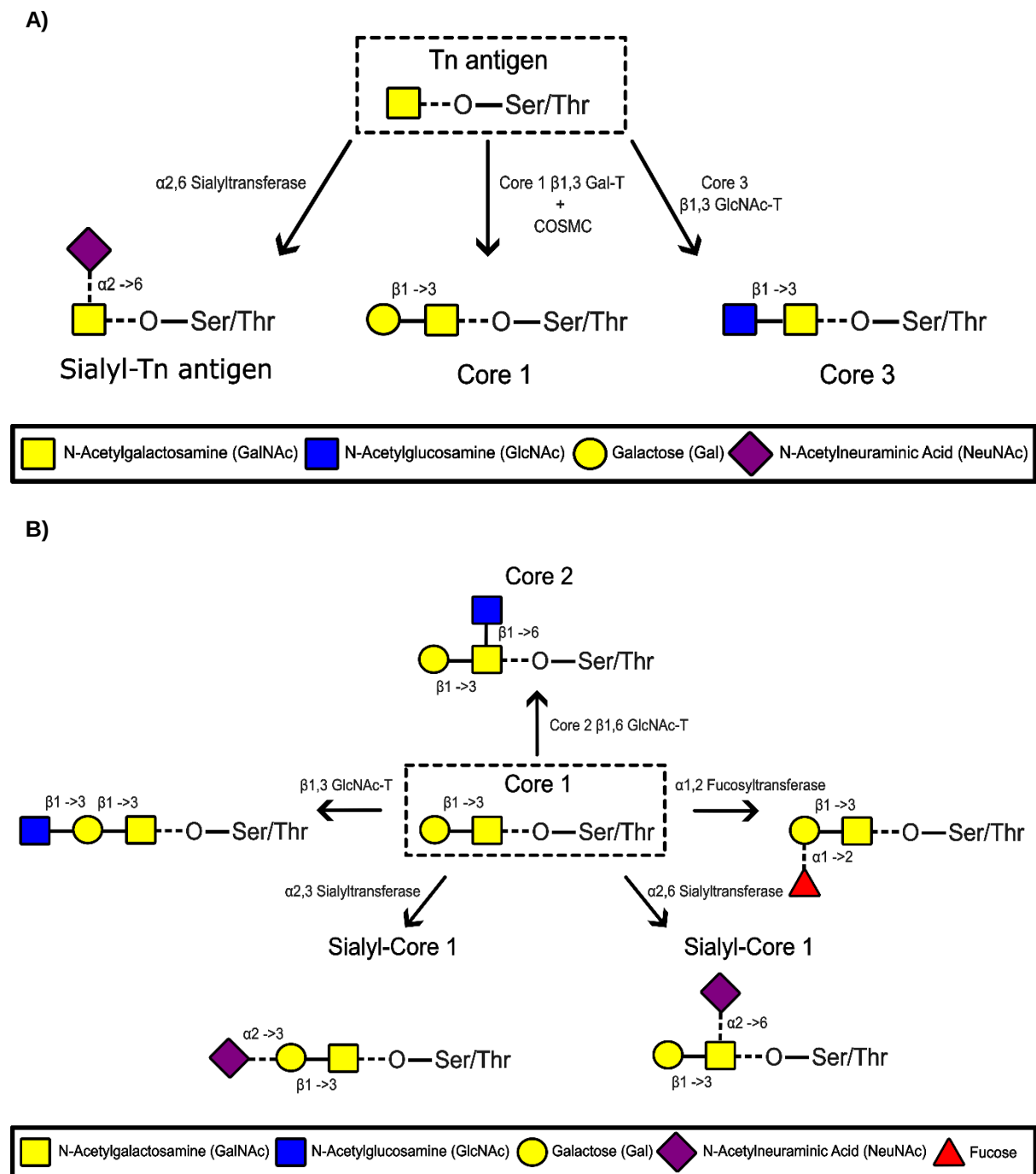
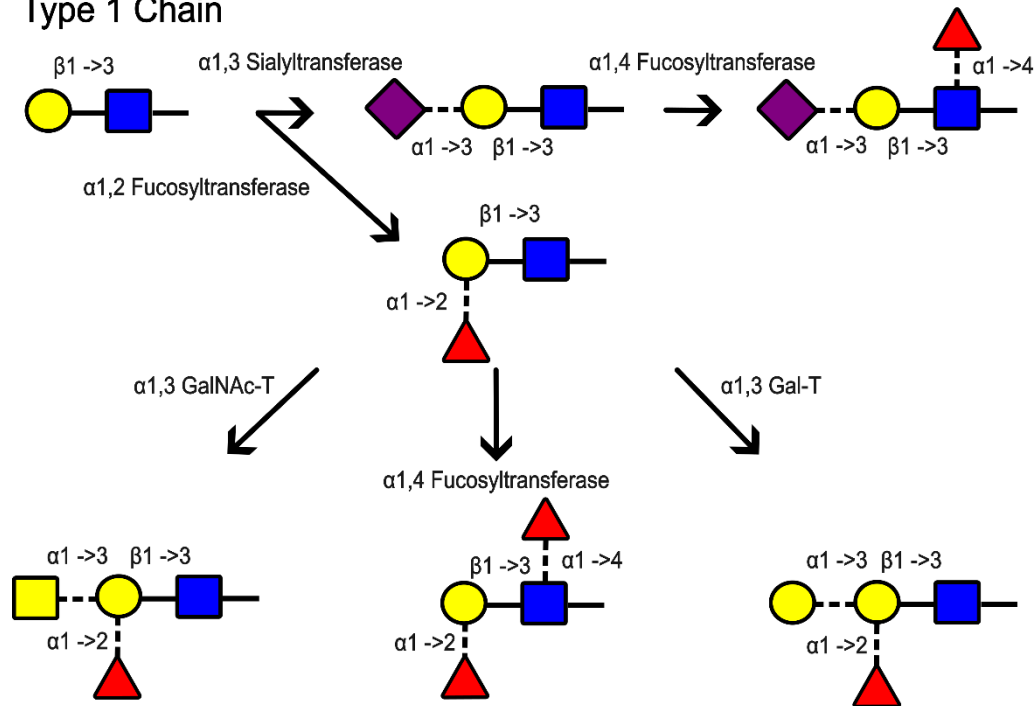


Figure 1.5 Substrate availability can regulate O-glycans created - common competing glycosyltransferases. A) Commonly competing glycosyltransferases for Thomsen nouvelle (Tn) antigen. B) Commonly competing glycosyltransferases for core 1 glycans. Adapted from [10].

Type 1 Chain



Type 2 Chain

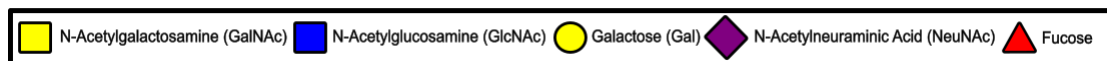
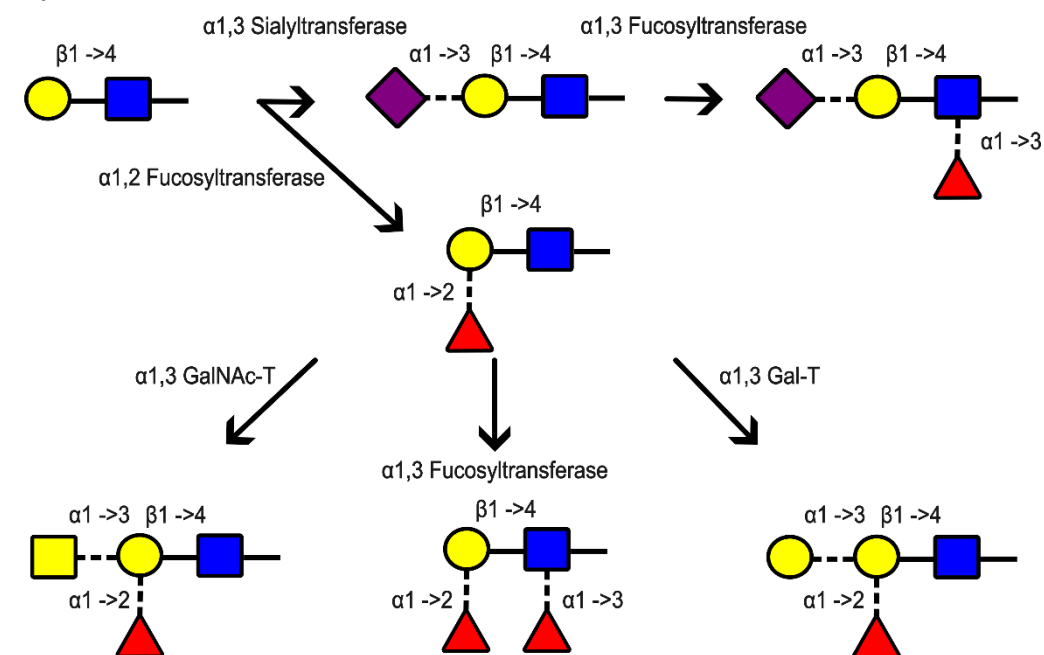


Figure 1.7 Type 1 and type 2 chain terminations through the addition of either a fucose (Fuc) or N-acetylneuraminic acid (NeuNAc) monosaccharide. Adapted from [10].

1.1.5 GALNTs

The GALNTs, responsible for the addition of a GalNAc monosaccharide to a Ser or Thr in a polypeptide to initiate mucin-type O-linked glycosylation, are the largest known family of glycosyltransferases (reviewed by [46] and [47]). To date, 20 GALNTs have been identified in humans, of which 17 have been characterised functionally. Uniquely to eukaryotes these enzymes, except GALNT20, contain both a catalytic and a lectin (ricin-like) domain [48]. They are type II Golgi apparatus localised proteins [49, 50] that have a short N-terminal cytoplasmic tail (4-24 amino acids long), single hydrophobic transmembrane domain, variable length stem region (from ~90 amino acids long in GALNT1, -T13, -T16 to ~170 amino acids long in GALNT3, -T8 and -T15), with exception of an extended 470 amino acid stem region in GALNT5, and both a catalytic (~230 amino acids long) and C-terminal ricin-like lectin domain which demonstrates binding specificity for α -N-acetyl-D-galactosamine (α -GalNAc) (~120 amino acids long) [51, 52], see Figure 1.8. GALNT10 has been shown to have an additional catalytic domain to this common structure [53].

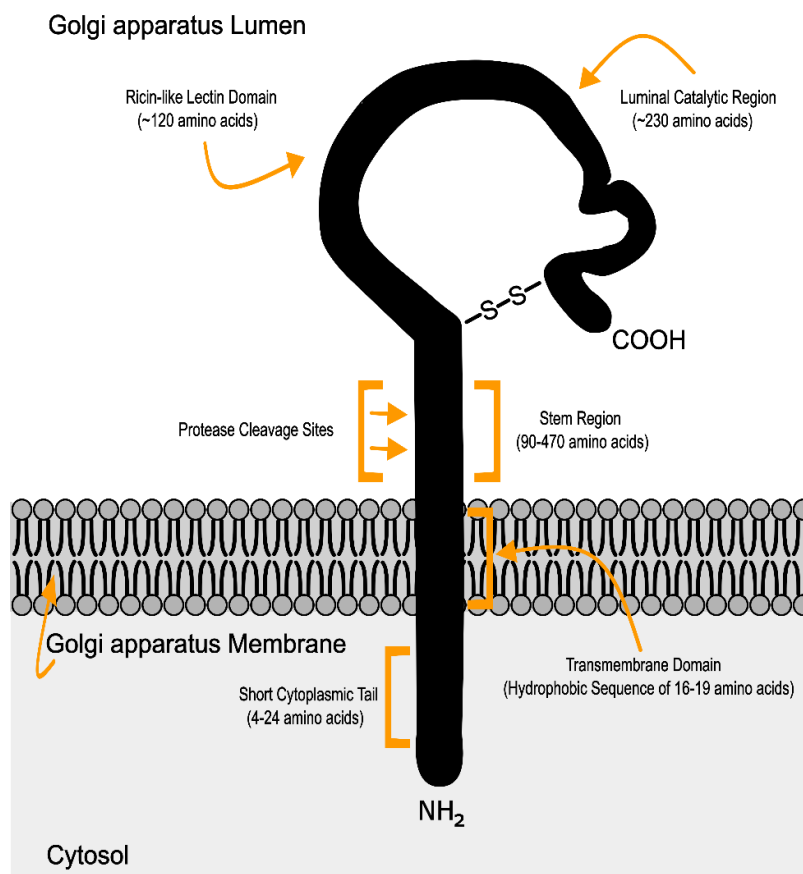


Figure 1.8 Schematic representation of UDP-N- α -D galactosamine: polypeptide N-acetylgalactosaminyltransferases (GALNTs) shared type II membrane structure, located in the Golgi apparatus. A hydrophobic transmembrane domain followed by a short N-terminal tail holds the luminal portion in place. The luminal portion comprises a catalytic domain, a ricin-like domain, and a variable-length stem region. Adapted from [54].

GALNT nomenclature was assigned according to the order of description within the literature, and the various GALNT gene paralogs are also thought to all, except for GALNT4, have arisen from one ancestral gene [55]. GALNTs are functionally very conserved with GALNTs identified in *Caenorhabditis elegans* [56] and *Drosophila melanogaster* [57]. Why so many GALNTs exist is not fully understood however; they seem to demonstrate distinct acceptor peptide substrates, which can be either unmodified or partially glycosylated O-GalNAc-peptide substrates. Some GALNTs, like GALNT1 and -T2 are classed as “early acting”, as they act on naked peptides. Others, such as GALNT4, are classified as “mid-acting” as they demonstrate a preference for peptides with neighbouring O-GalNAc-glycosylation sites. Finally, “late acting” GALNTs such as GALNT7 and -T10 act only on areas of high O-GalNAc glycosylation [29]. GALNTs also exhibit specific differential temporal, spatial and tissue expression. Some GALNTs such as GALNT1 and -T2 are ubiquitously distributed in human tissues [58, 59], other GALNTs such as GALNT3, -T4, -T6, -T7, -T10, -T11, -T12, -T14, -T15 and -T16 are widely expressed by a range of tissues, whilst the remaining GALNTs demonstrate highly regulated expression in limited tissues (reviewed by [46] and [54]).

1.1.6 Control of mucin-type O-linked glycosylation

Regulation of mucin-type O-linked glycans is complex and not fully understood. Increasing evidence suggests that it could occur over a myriad of steps, see Figure 1.9. For example, mutational events in glycosylation-related genes, such as C1GALT1 (reviewed by [44]), COSMC [41, 43, 60, 61] and core 3 β 1,3 GlcNAc-T [62, 63] can affect the glycosylation profile of a cell. Transcription factors, chromatin structure and its regulators can also affect the regulation of the glycome (reviewed by [64]). MicroRNAs (miRNAs) have been documented to regulate over 80 glycozymes across all forms of glycosylation (reviewed comprehensively by [65] and [66]). Other studies have demonstrated that epigenetic modifications, such as methylation, can impact glycozyme expression for example, hypermethylation of GALNTs [67, 68] and COSMC [69, 70] result in transcriptional silencing and subsequently impact glycosylation.

Relocation of GALNTs from the Golgi apparatus to the ER, known as GALNT activation or GALA pathway, results in large increases of unelaborated Tn antigen [71-79]. This type of relocation has not been reported for other glycosyltransferases [75, 78]. There is also evidence that the availability of sugar nucleotide donors, such as UDP-Gal, in the Golgi apparatus lumen and monosaccharide availability may also, unsurprisingly, influence the glycosylation profiles of cells [80-83]. Finally, there is also evidence that the glycosylation profile of a cell can be altered due to changes in Golgi apparatus structure (reviewed by [84] and [85]) and environment (pH) [86-89].

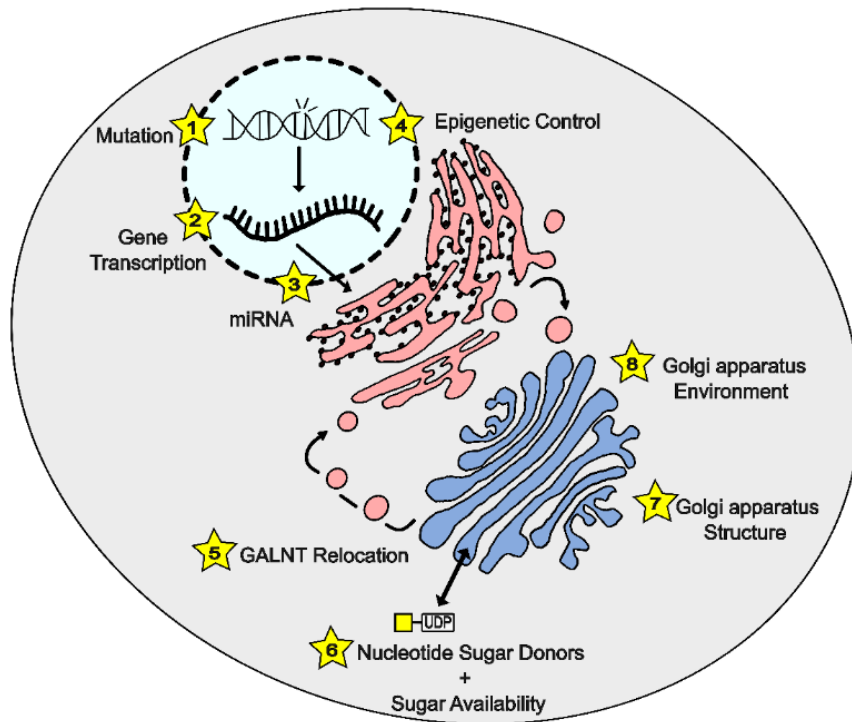


Figure 1.9 Multi-level regulation of mucin-type O-linked glycosylation. Alterations in glycosylation can occur via (1) mutation events of glycozymes, (2) changes in gene transcription (transcription factors), (3) post-transcriptional control through miRNAs, (4) epigenetic alterations, (5) UDP-N- α -D galactosamine: polypeptide N-acetylgalactosaminyltransferase (GALNT) enzyme relocation (GALA pathway), (6) availability of monosaccharides and nucleotide sugar donors, (7) alterations in Golgi apparatus structure, and (8) alterations in Golgi apparatus environment.

1.2 Lectins

Lectins are a family of naturally occurring carbohydrate-binding proteins, or glycoproteins, of non-immune origin that specifically recognize glycans and, because they are multivalent, agglutinate cells or precipitate glycoconjugates [90]. Lectins bind to glycans in a non-covalent, reversible, highly specific manner (reviewed by [91]). Their carbohydrate-binding specificity is typically characterised by the monosaccharide, or occasionally di- or tri-saccharides, they show the greatest affinity towards, and most effectively inhibit agglutination. Lectin specificity also depends upon other factors such as proximity and spatial distribution of neighbouring glycans, and electrostatic and hydrophobic interactions [92]. Lectins can also be grouped according to their molecular structure, see Figure 1.10. Stillmark [93] first isolated the highly toxic protein ricin, derived from the castor bean (*Ricinus communis* L.) seed, in 1888 and noted its ability to cause hemagglutination. Since then, a wide variety of lectins have been isolated from diverse sources - mostly from plants, but also from microbes and animals. Lectins are ubiquitous in living organisms [94] and are a useful tool for glycobiologists in a range of applications, including lectin cytochemistry and flow cytometry, helping to identify the

glycosylation profiles of cells and tissues and to uncover the variety of glycosylation functions in biological systems (reviewed by [95]).

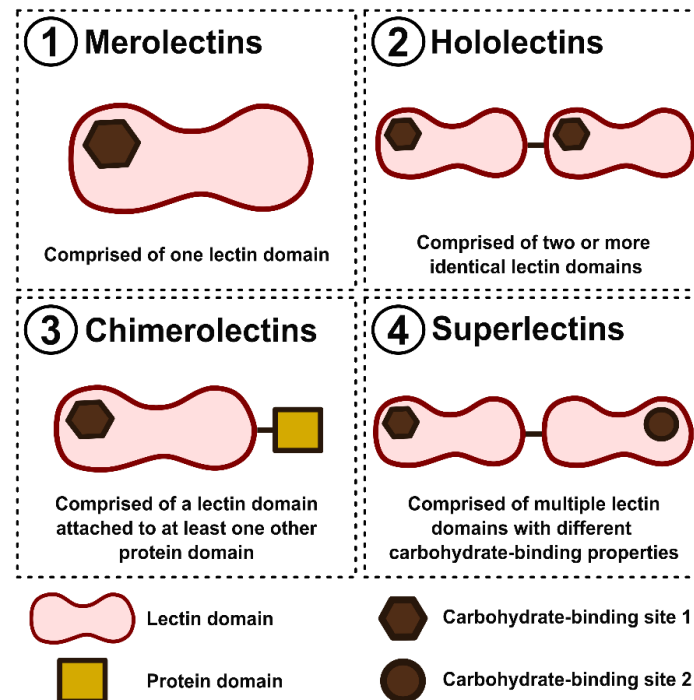


Figure 1.10 Plant lectin classification based upon their structural diversity. (1) Merolectins are lectins that consist of one lectin domain. (2) Hololectins are lectins that are comprised of two or more identical lectin domains. (3) Chimerolectins are lectins that have a lectin domain attached to at least one other protein domain. (4) Superlectins are lectins that are comprised of multiple lectin domains with different carbohydrate-binding properties. Adapted from [96].

1.3 Cancer

1.3.1 Breast cancer

The breast comprises functional units called terminal ductal-lobular units that are made of ducts that branch out into several lobules, the small rounded glandular structures responsible for secreting milk, see Figure 1.11A. Two types of epithelium cells are present in the breast, luminal epithelial cells that form the inner layer of the duct, and myoepithelial cells which form the outer layer of the mammary ducts, see Figure 1.11B. These are surrounded by the basement membrane which separates the epithelium from the surrounding stroma and ECM (reviewed by [97]).

Breast cancer is the most common form of cancer in the United Kingdom, with approximately 55,900 females and 390 males diagnosed every year [98]. Most breast cancers arise from epithelial cells that line the ducts (ductal carcinoma *in situ* (DCIS) and invasive ductal carcinoma (IDC)) or lobules (lobular carcinoma *in situ* (LCIS) and invasive lobular carcinoma (ILC)). The two main forms of breast

cancer are, carcinoma *in situ* (non-invasive) whereby cancer cells remain within the basement membrane of the lobules, see Figure 1.11C or invasive carcinoma in which cancers disseminate outside of the basement membrane, see Figure 1.11D. IDC represents the most common form of breast cancer, followed by ILC [98].

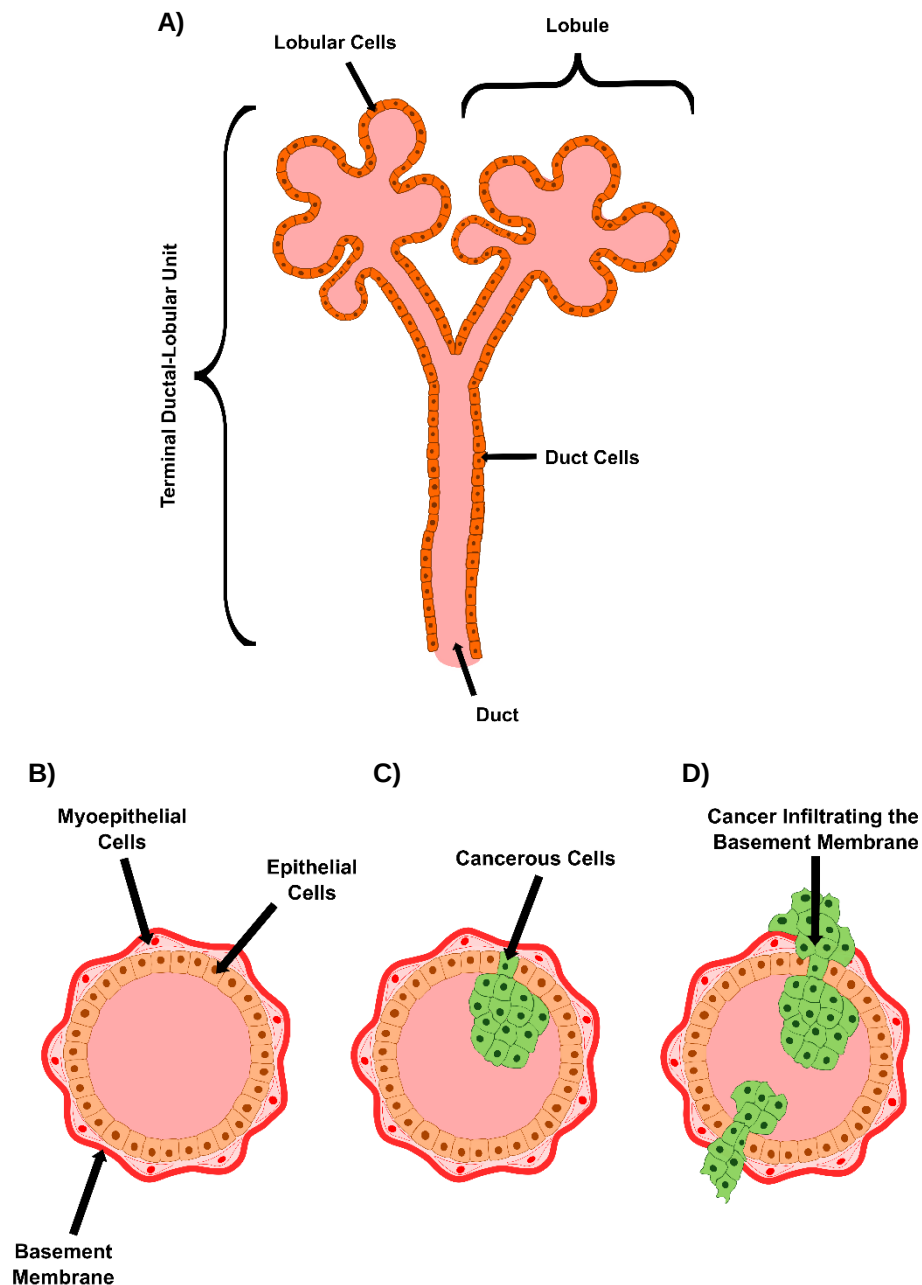


Figure 1.11 Anatomy of terminal ductal-lobular unit and progression of ductal carcinoma *in situ* (DCIS) through to invasive ductal carcinoma (IDC). A) Terminal ductal-lobular unit of normal breast. B) Both milk ducts and lobules are comprised of myoepithelial cells lined by epithelial cells both surrounded by a basement membrane. Most breast cancers arise from the epithelial cells. C) Transformation of the epithelial cell to DCIS. D) Disease progression to IDC.

1.3.2 Glycans as prognostic markers in cancer

The Tn antigen is usually present in significant levels during embryogenesis and is cryptic in healthy adult tissue due to chain extension, as described previously. Intriguingly, it is present in a variety of human pathologies including Tn syndrome, IgA nephropathy and, notably, in many cancers (reviewed by [99] and [11]). The Tn antigen was first described in human cancers by Prokop and Uhlenbruck [100], and literature suggests that as many as 90% of cancers of epithelial origin synthesise unelaborated Tn antigen, including carcinomas of the breast, lung, bladder, cervix, ovary, stomach and prostate [101, 102]. Recently, Rømer et al., [103] reported high levels of Tn antigen in a survey of over 700 tissue specimens of cancers of the pancreas, colorectum and breast, and negligible levels in normal tissue.

The presence of Tn antigen in cancers can be revealed by binding of the lectins from *Helix pomatia* agglutinin (HPA) (79kDa, hexameric glycoprotein, with 6 carbohydrate binding sites, derived from the albumen gland of the Roman snail) or *Vicia villosa* agglutinin (VVA) (114kDa, tetrameric glycoprotein, with four carbohydrate binding sites, derived from hairy vetch seeds). The presence of Tn antigen, detected through binding of these lectins has frequently been shown to be associated with more aggressive behaviour, metastasis and poor patient prognosis (see [104], [16] and [95] for review), discussed in detail in Chapter 3. Similarly, the presence of high levels of T antigen (core 1), often detected by binding of the lectin peanut agglutinin (PNA) (110kDa, tetrameric glycoprotein, with four carbohydrate binding sites, derived from *Arachis hypogaea* fruit, peanuts), or sialyl-core 1 have also been commonly described in a range of cancers and also associated with aggressive biology and poor prognosis (reviewed by [105]).

1.3.3 Hallmarks of cancer

Cancer is a disease that is characterised by the accumulation of genetic alterations which cause uncontrolled cell growth, leading to the formation of tumours. This includes activation of proto-oncogenes, which drive cell cycle forwards, and/or inactivation of tumour-suppressor genes, which normally operate to restrict cellular growth (reviewed by [106]). Hanahan and Weinberg [107] first proposed the ‘hallmarks of cancer’, six different functional capabilities shared by most cancers that are essential in both the development and progression of cancer, intending to unify the complexities of cancer. These ‘hallmarks’ have since been developed through an additional two iterations, the first, Hanahan and Weinberg [108], and the currently proposed hallmarks, Hanahan [109], see Figure 1.12.

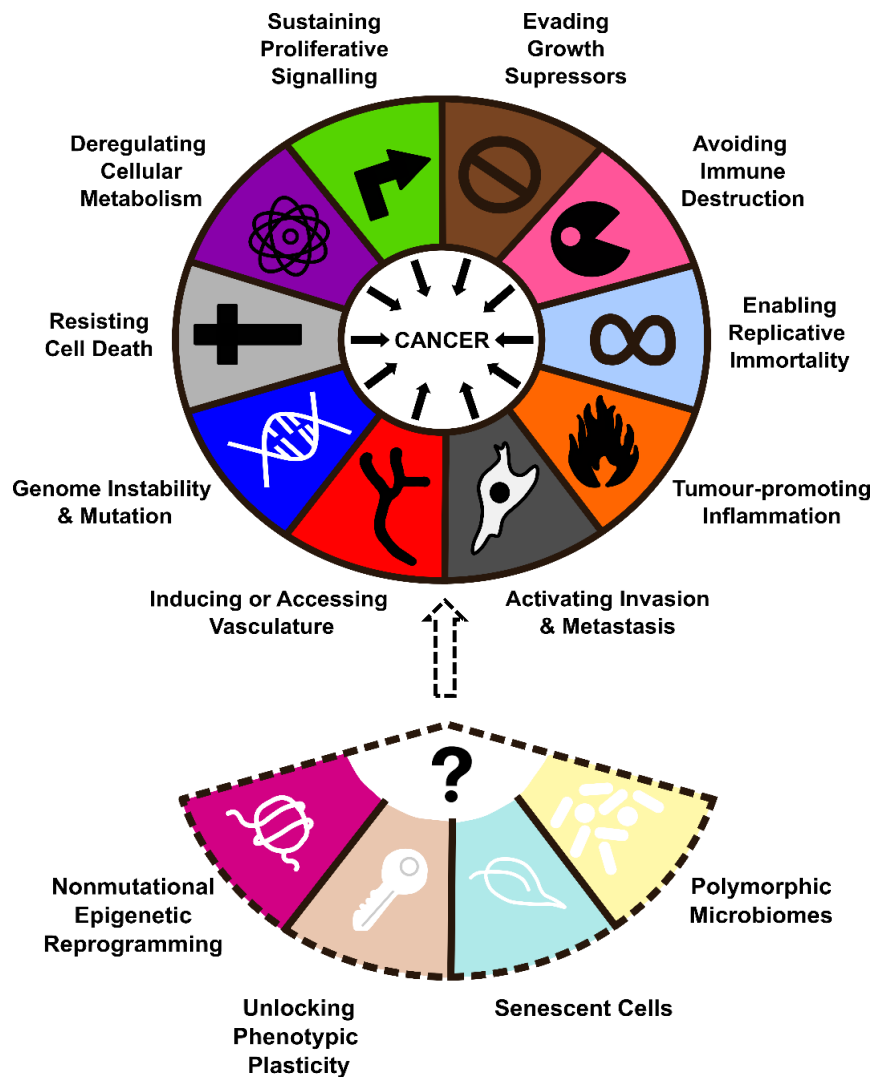


Figure 1.12 Hallmarks of cancer. The ten hallmarks of cancer encompass eight acquired capabilities, evading growth suppressors, avoiding immune destruction, enabling replicative immortality, activating invasion and metastasis, inducing or accessing vasculature, resisting cell death, deregulating cellular metabolism and sustaining proliferative signalling. It also encompasses two enabling characteristics, tumour-promoting inflammation and genome instability and mutation. Four additional emerging hallmarks and enabling characteristics have been proposed which include; unlocking phenotypic plasticity, non-mutational epigenetic reprogramming, polymorphic microbiomes and senescent cells. Adapted from [109].

1.3.4 Metastasis

Cancer metastasis occurs as a result of a cascade of sequential events that culminates in the dissemination of cancer cells from the primary tumour site to another, usually distant, secondary site. Metastasis is the primary cause of cancer-related death, often cited as accounting for up to 90% of cancer deaths [107] however, the process is still not fully understood (reviewed by [110] and [111]). The metastatic cascade is characterised by a series of events including; angiogenesis, disaggregation

of cancer cells from the primary tumour, invasion and migration through the basement membrane and ECM, intravasation into a blood vessel, adhesion to vascular endothelium and extravasation from the blood vessel to form a secondary tumour at a new site, see Figure 1.13.

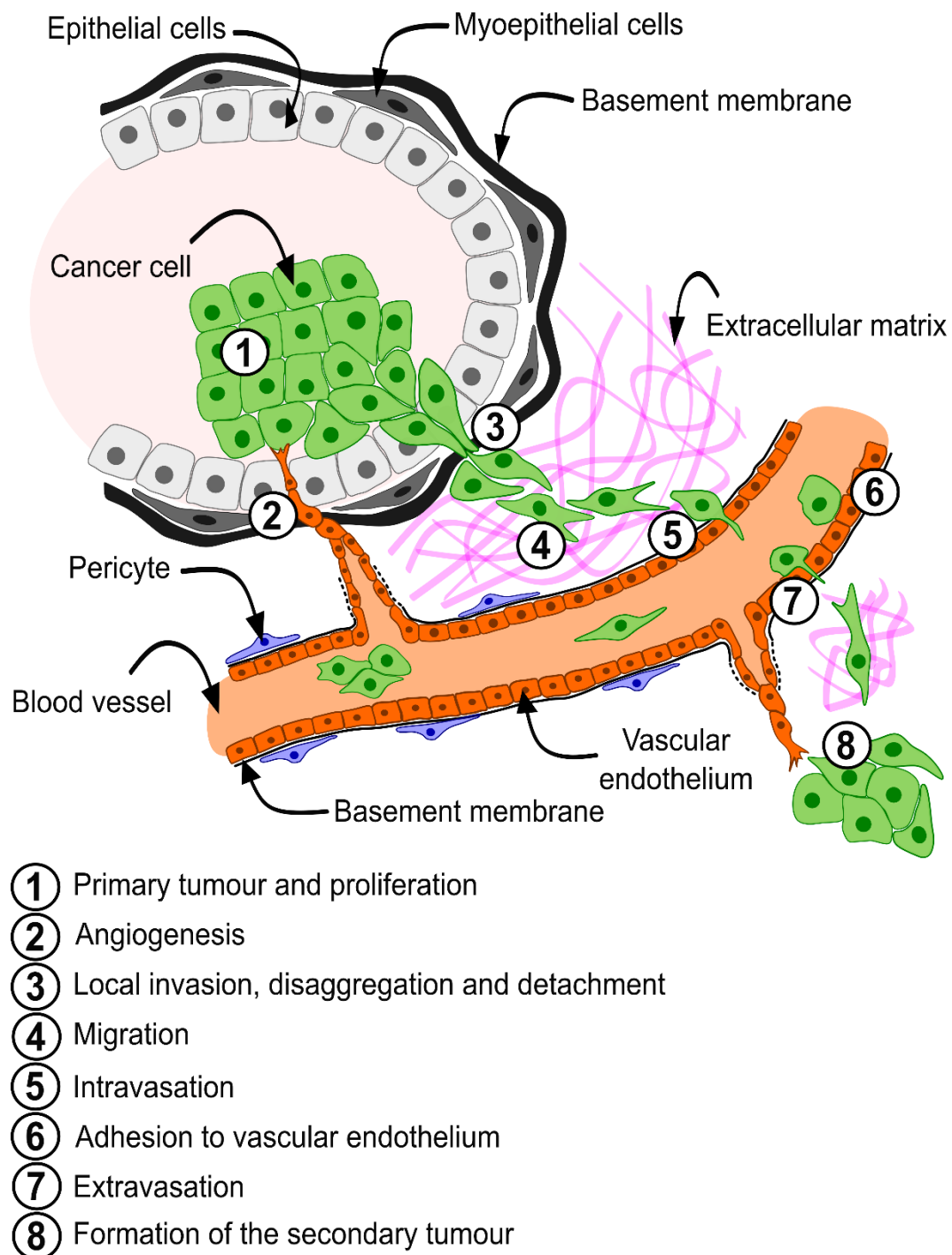


Figure 1.13 The metastatic cascade. Distinct steps in the metastatic cascade that must be completed successfully for cancer cells to spread include; angiogenesis, disaggregation of cancer cells from the primary tumour, invasion and migration through the basement membrane and extracellular matrix, intravasation into a blood vessel, adhesion to vascular endothelium, extravasation from the blood vessel to form a secondary tumour at a new site.

1.3.5 Tumour extracellular matrix

The ECM is a dynamic, three-dimensional (3D), complex network of approximately 300 cross-linked proteins. It provides both biophysical and biochemical cues to modulate cell signalling, morphology, adhesion and migration in a bi-directional manner with cells, which is important in the metastatic cascade [112, 113]. Alongside core ECM fibrous proteins, such as fibronectin, collagens and laminins, and glycoproteins, such as hyaluronic acid, the ECM also acts as a reservoir for a variety of cytokines and growth factors, such as transforming growth factor- β (TGF- β) and fibroblast growth factors, acting as a communication liaison between cells. Therefore, ECM composition, biophysical properties and repertoire of ECM receptors on a cell can determine a variety of cellular responses. During pathologies, such as cancer, the ECM undergoes extensive remodelling, increasing ECM deposition and stiffness and changes to abundance and composition, helping to drive metastasis via the metastatic cascade through altered mechanosignaling and actin cytoskeleton reorganisation (reviewed by [114]). Importantly, ECM remodelling enzymes, such as matrix metalloproteinases (MMPs), are also often upregulated in cancer, providing a route for cell migration and invasion (reviewed by [115]).

1.4 Epithelial-mesenchymal transition (EMT)

Epithelial-mesenchymal transition (EMT), first described by Greenburg and Hay [116] as 'epithelial-mesenchymal transformations', is a biological process whereby basement membrane-bound, polarized epithelial cells trans-differentiate into migratory, invasive mesenchymal cells. This requires changes in cellular architecture, morphology, adhesion and migratory capabilities and is critical to biological processes including embryogenesis, wound healing and cancer metastasis [117]. It is a highly plastic, reversible process and mesenchymal-epithelial transition (MET) allows for the transition of motile, mesenchymal cells back into polarized epithelial cells (reviewed by [118]).

Epithelial cells typically demonstrate tight adhesions between neighbouring cells, formed by *zona occludens*, with regularly spaced cell-cell junctions, formed by E-cadherin, that allow for the formation of polarized sheets of cells and are characterised by their 'cobble-stone' appearance. Mesenchymal cells on the other hand do not demonstrate tight adhesions, are not uniform in composition or density and therefore do not demonstrate common shapes or structures. In general, compared to epithelial cells, they tend to demonstrate more extended and elongated shapes, front-back polarity and resistance to anoikis. During EMT, epithelial cells have their actin cytoskeleton re-structured with stress fibres and vimentin replaces cytokeratin intermediate filaments (reviewed by [119]). Weakened, transient adhesions between mesenchymal cells, due to the replacement of E-cadherin and epithelial-specific integrins with N-cadherin and $\alpha\beta6$ integrins, is what facilitates increased dynamic, migratory capacity [120]. When E-cadherin is cleaved and replaced during EMT, β -catenin, a protein

that binds to the cytoplasmic domain of E-cadherin, is released and accumulates in the cytoplasm and is ultimately translocated to the nucleus. Once in the nucleus, it can induce EMT-related gene expression, including the Snail transcription factor. E-cadherin is therefore often referred to as the “gatekeeper of the epithelial state” [121]. Mesenchymal cell movement is different to that of epithelial cells due to their ability to move individually, discussed further in ‘1.4.4 Cell migration’ (reviewed by [122]). There are a variety of protein markers that can be used to identify EMT, and these include the reduction of epithelial marker E-cadherin, and the increase of mesenchymal markers such as N-cadherin, vimentin and fibronectin. EMT also results in activation of key signalling pathways, such as the TGF- β pathway, and increased expression of EMT-inducing transcription factors, including TWIST, Snail, Slug and the ZEB family, discussed further in ‘1.4.3 Induction of EMT’. *In vitro* functional markers include elongation of cell shape and increased invasion and migration capability, see Figure 1.14. Whilst some cells transition fully into a mesenchymal phenotype, it is thought that some cells maintain their epithelial traits whilst also acquiring some mesenchymal ones. This is thought to be an intermediate state of EMT, referred to as partial EMT (pEMT), discussed further in ‘1.4.2 Partial EMT’.

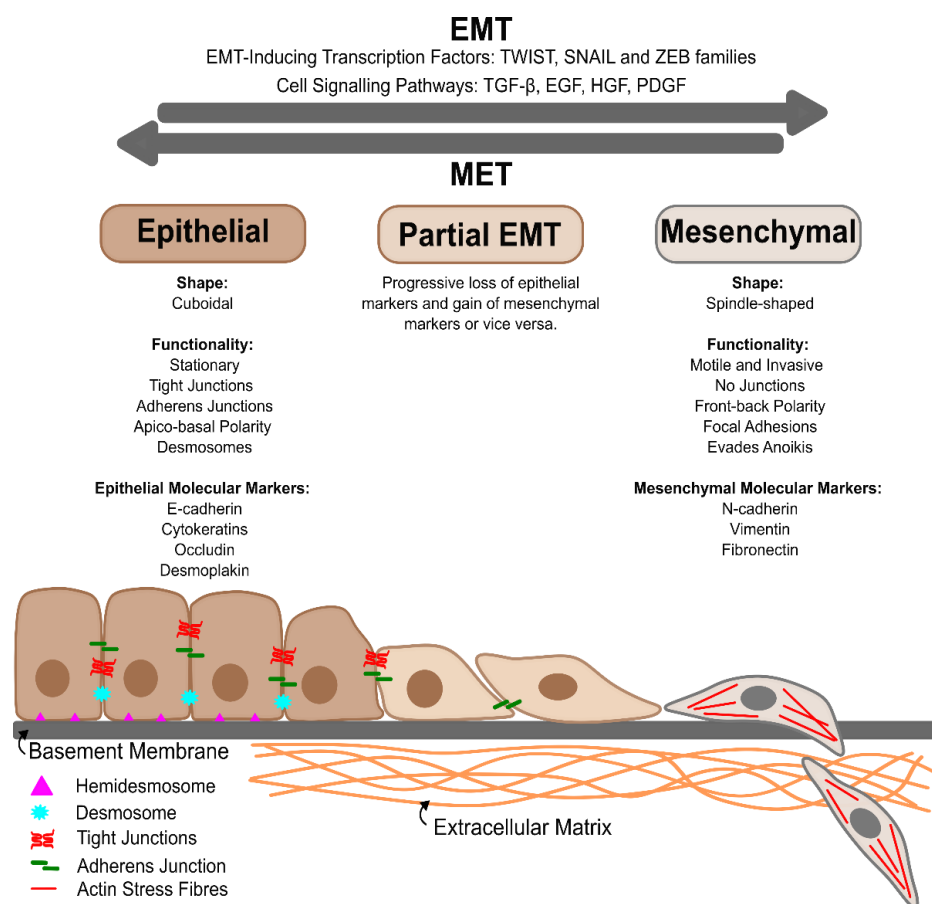


Figure 1.14 Overview of epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET). Key shape identifiers, functional identifiers and molecular markers of both epithelial and mesenchymal cell states. EMT-inducing transcription factors and cell signalling pathways involved in EMT are listed. Adapted from [123].

1.4.1 Types of EMT

EMT has been classified into three subtypes, as proposed by Kalluri and Weinberg [124], based on the three distinct biological settings in which it occurs and its functional consequences, see Figure 1.15. Despite this, the underlying genetic and biochemical elements that allow for phenotypic change contain considerable commonality. Type 1 EMT is associated with implantation, embryogenesis and organ development; its purpose is to allow tissue remodelling for morphogenesis and does not result in an invasive phenotype or fibrosis. Type 2 EMT is associated with wound healing, tissue regeneration and organ fibrosis and is the result of a repair-associated event and is linked to inflammation. Type 3 EMT is associated with cancer progression and metastasis in neoplastic cells, which is facilitated by mutational events affecting both oncogenes and tumour suppressor genes - for example, Ras GTPase - which activate downstream EMT inducers (for example, EMT-activating transcription factors) and results in an invasive phenotype (reviewed by [124]).

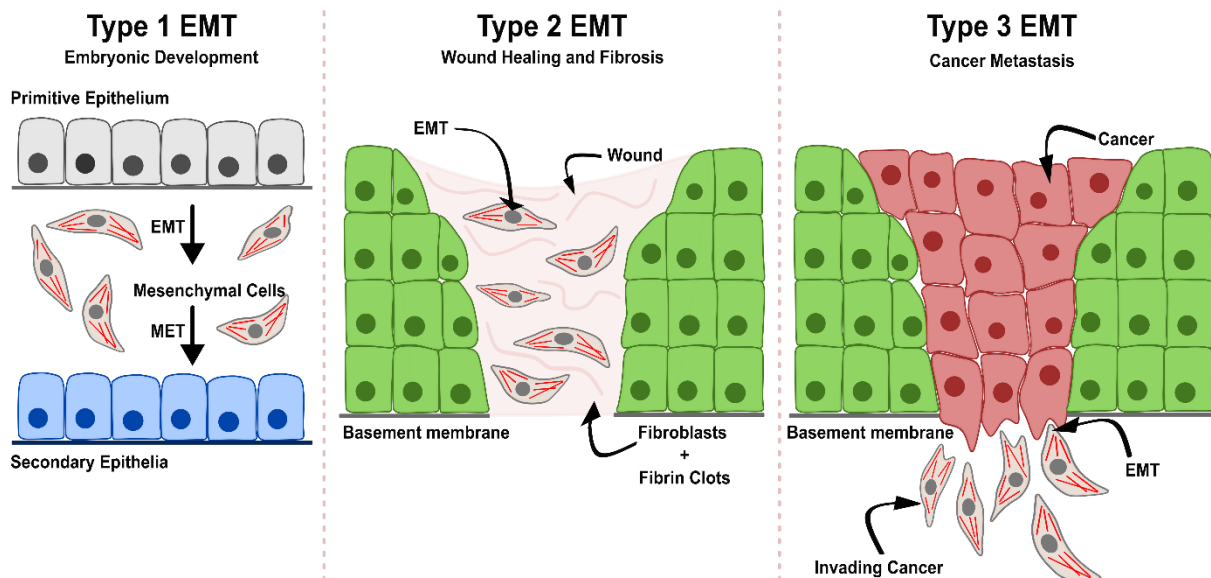


Figure 1.15 Three types of epithelial-mesenchymal transition (EMT). Type 1 EMT is associated with embryonic development, its purpose is to allow tissue remodelling for morphogenesis and does not result in an invasive phenotype or fibrosis. Type 2 EMT is associated with wound healing and fibrosis and is the result of a repair-associated event and is linked to inflammation. Type 3 EMT is associated with cancer metastasis. It is facilitated by mutational events affecting oncogenes and tumour suppressor genes, which activate downstream EMT inducers (for example, EMT-activating transcription factors) and result in an invasive phenotype. Adapted from [125].

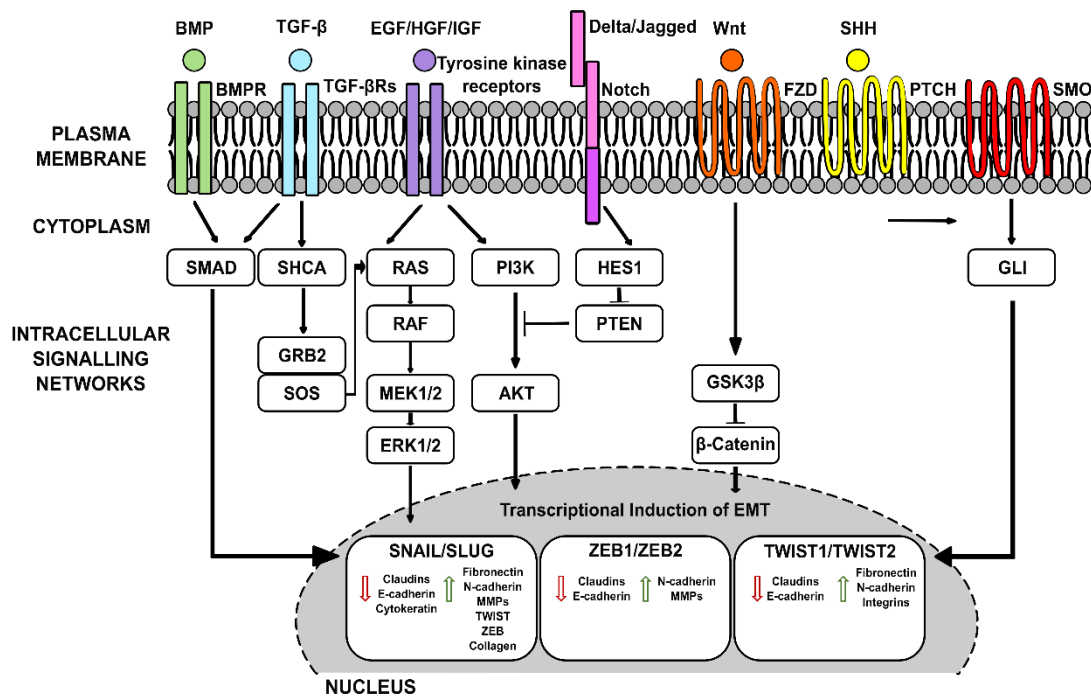
1.4.2 Partial EMT

There is increasing evidence that cancer cells undergoing type 3 EMT rarely transition into a full mesenchymal-cell state during cancer metastasis (reviewed by [123]). Instead, cancer cells are thought to exist in an array of partial, pEMT, states, in which cells can display a wide spectrum of and simultaneous expression of both epithelial and mesenchymal markers, allowing for a more dynamic and frequent reversion via MET [126]. Cells can also simply be induced into a pEMT state by the loss of epithelial markers alone [127]. This is important, as to successfully form secondary sites, cancer cells must be able to display versatile, dynamic phenotypes that allow them to acquire traits that meet specific needs required during each step of the metastatic cascade, as discussed in '1.3.4 Metastasis'. pEMT cancer cells are associated with more aggressive disease phenotypes, resistance to anoikis, immune evasion, and cancer stem cell-like features, therefore, making them more effective at metastasis (reviewed by [128]). pEMT is also referred to as intermediate EMT [129, 130], incomplete EMT [131, 132], hybrid epithelial/mesenchymal [133-138] and semi-mesenchymal states [139, 140].

1.4.3 Induction of EMT

EMT is thought to be induced by stromal signals including growth factors, cytokines and ECM proteins, secreted in both an autocrine and paracrine manner in response to a variety of environmental cues such as hypoxia and inflammation. These signals induce a variety of intracellular signalling networks and pathways and subsequently, transcription factors, mainly the SNAIL, TWIST and ZEB families, which in turn act pleiotropically to induce EMT. A variety of different signalling pathways can be involved in EMT including TGF- β , bone morphogenetic protein (BMP), receptor tyrosine kinase (RTK), Wnt/ β -catenin, Notch, Hedgehog, signal transducer and activator of transcription 3 (STAT3), ECM-mediated, and hypoxia signalling pathways (reviewed by [141]), see Figure 1.16 for an overview.

EMT INDUCING SIGNALS



Abbreviations

BMP - Bone morphogenic protein,
 BMPRs - Bone morphogenic protein receptors,
 TGF-β - Transforming growth factor-β,
 TGF-βRs - Transforming growth factor-β receptors,
 SHCA - SRC homology 2 domain-containing-transforming protein,
 GRB2 - Growth factor receptor-bound protein 2,
 SOS - Son of sevenless,
 EGF - Epidermal growth factor,
 HGF - Hepatocyte growth factor,
 IGF - Insulin-like growth factor,
 HES1 - Hairy and enhancer of split 1,

PTEN - Phosphatase and tensin homolog,
 FZD - Frizzled,
 GSK3β - Glycogen synthase kinase-3β,
 SHH - Sonic hedgehog,
 PTCH - Patched,
 SMO - Smoothened,
 GLI - Glioma-associated,
 ZEB - Zinc-finger E-box-binding,
 MMPs - Matrix metalloproteinases.

Figure 1.16 Overview of epithelial-mesenchymal transition (EMT) inducing signals, and associated transcriptional regulation. EMT can be induced by stromal signals including growth factors, cytokines and extracellular matrix (ECM) proteins, secreted in both an autocrine and paracrine manner in response to a variety of environmental cues, such as hypoxia and inflammation. These signals induce a myriad of intracellular signalling networks and pathways and subsequently, transcription factors, mainly the SNAIL, TWIST and ZEB families, which in turn act pleiotropically to induce EMT. Adapted from [54].

1.4.4 Cell migration

Cells can utilise a variety of modes of cell migration, which can be categorised into single-cell migration - either amoeboid or mesenchymal - or collective migration, see Figure 1.17 (reviewed by [142] and [143]). Typically, each cell type exhibits a preference for a particular mode of cell migration; for example, epithelial cells demonstrate a preference for collective cell migration, whereas, leukocytes demonstrate a preference for amoeboid [144].

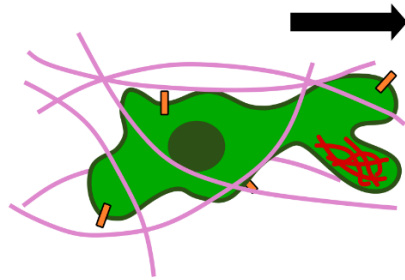
Amoeboid cell migration is characterised by high-speed cell migration of cells that are deformable and able to navigate through the ECM, exhibit 'bleb-like' protrusions, demonstrate a clear 'leading edge' and are independent of integrins, cadherins and metalloproteases and therefore do not demonstrate ECM remodelling [145-147].

Mesenchymal cell migration is characterised by cells that demonstrate an elongated morphology but, remain dependent upon integrins to contact the ECM, move via cytoskeletal contractility and require metalloproteases for ECM degradation and remodelling to generate a path for migration.

Mesenchymal cell migration is typically regarded as being slower than amoeboid. This is due to the relatively slow focal adhesion turnover required for migration [148]. Mesenchymal cell migration is categorised into five key steps that continuously cycle. This was first described by Lauffenburger and Horwitz [149] and includes 1) protrusion of the cell 'leading edge' resulting in the formation of actin-rich lamellipodia or filopodia, 2) occurrence of focal adhesion and clustering of integrins, 3) assembly of focal contacts and activation of proteolytic enzymes (MMPs) resulting in ECM remodelling and 'cell-matrix' interface, 4) cell contraction by actomyosin and 5) movement into newly formed defects in ECM by detachment of the 'trailing edge'. Amoeboid and mesenchymal single-cell migration are thought to be mutually interchangeable in a highly plastic process that is considered to be part of the spectrum of EMT, called mesenchymal-amoeboid transition (MAT) and amoeboid-mesenchymal transition (AMT). This change can result from responses to environmental and cellular changes, allowing enhanced migration in complex 3D environments [145, 150].

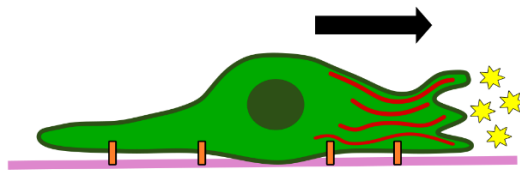
Finally, collective cell migration which is characterised by intact cell-cell junctions, and migration '*en bloc*' as a sheet or group of cells is dependent upon integrin contacts with the ECM and often demonstrates 'leader' cell(s) that use metalloproteases to degrade the ECM, and is followed by 'follower' cells [151, 152]. 'Leader' cell(s) have been shown to often display a mesenchymal-like phenotype and also demonstrate clear differences in gene expression [142]. In human breast tumours, three-dimensional analysis of the stromal border has shown a predominance of collective invasion and minimal single-cell migration [153].

① Amoeboid Migration



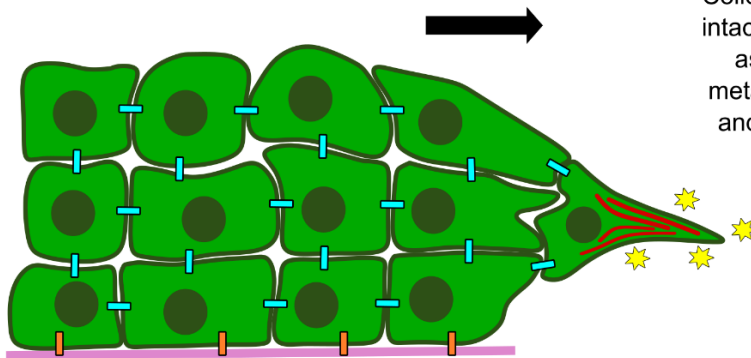
Amoeboid migration is characterised by individual cells that demonstrate a clear 'leading edge', can change shape and direction to migrate through gaps in the extracellular matrix (ECM) and are independent of integrins, cadherins and metalloproteases.

② Mesenchymal Migration



Mesenchymal migration is characterised by individual cells that demonstrate an elongated morphology but, remain dependent upon integrins to contact the ECM and metalloproteases.

③ Collective Migration



Collective migration is characterised by intact adherens, and migration '*en bloc*' as a sheet of cells with the use of metalloproteases to degrade the ECM and remain dependent upon integrin contact with the ECM.

- Integrin-mediated contact
- Adherens junction
- Degraded ECM
- ECM
- Actin Cytoskeleton

Figure 1.17 Cell migration methods. (1) Amoeboid migration is characterised by individual cells that demonstrate a clear leading edge, can change shape and direction to migrate through gaps in the extracellular matrix (ECM) and are independent of integrins, cadherins and metalloproteases. (2) Mesenchymal migration is characterised by individual cells that demonstrate an elongated morphology but, remain dependent upon integrins to contact the ECM and metalloproteases. (3) Collective migration is characterised by intact adherens, and migration "*en bloc*" as a sheet of cells with the use of metalloproteases to degrade the ECM and remain dependent upon integrin contact with the ECM. Adapted from [143].

1.4.5 Cancer stem cells

Cancer stem cells are a subpopulation of highly plastic tumour cells that express adult stem-cell-associated markers, such as cluster of differentiation 44 (CD44), an ability to self-renew, unlimited proliferative potential, and a high capacity for tumour generation *in vivo*. They were first described in breast cancers by Al-Hajj et al., [154] and have since been implicated in tumour initiation, progression, metastasis, recurrence and resistance to therapy (reviewed by [155]). There are two main theories for how cancer stem cells arise: firstly, that they arise from the deregulation of normal stem cell self-renewal pathways or secondly, that they arise from cells undergoing EMT (reviewed by [156], [157] and [158]). Cancer stem cells are thought to play an important role in cancer metastasis and EMT as they demonstrate increased cell motility, invasion and expression of genes that promote metastasis. Key pathways involved in cancer stem cells include Notch, Wnt/ β -catenin, Hedgehog, Hippo and STAT3 [159-161]. Cancer stem cells have also been documented as small populations in many cancer cell lines, including breast cancer cell line MCF-7, indicating that although cancer cell lines may be clonally derived they may contain a hierarchy of cellular differentiation stages [159, 162].

1.5 Project Aims

This project aims to investigate the functional significance of truncated O-GalNAc glycans in the adhesive and motile behaviours of breast epithelial cells, with a focus on the relationship between O-GalNAc glycosylation and EMT.

Chapter 2

Materials and methods

2. Materials and methods

For the catalogue numbers and supply companies of all reagents and products, please refer to Supplementary Table 1.

2.1 Cell lines and routine cell culture

2.1.1 Epithelial cell lines

The cell lines used in this project were chosen as a representation of a range of cellular phenotypes, some of which have been previously well described to stably synthesise heterogeneous profiles of O-GalNAc terminal glycoproteins, common with those seen in clinical samples [163-165]. This includes; non-malignant breast epithelial cells (hTERT-HME1), primary, non-metastatic breast cancer cells (BT-474) and breast cancer cells derived from metastases (MCF-7, T-47D and ZR-75-1). All breast epithelial cell lines are listed in Table 2.1 with corresponding information about their origin and characteristics, and details of culturing in Table 2.2. All breast cancer cell lines were obtained from the European Collection of Authenticated Cell Cultures (ECACC).

2.1.2 Primary endothelial cells

Primary, human umbilical vein endothelial cells (HUVECs), were obtained from PromoCell® (Heidelberg, Germany) (C-12203, Lot No: 479Z016). These cells are pooled from multiple donors (up to four different umbilical cords), and cultured as detailed in Table 2.2.

Table 2.1 Characteristics of epithelial cell lines. Adapted from [164].

	Cell Line				
	hTERT-HME1	BT-474	MCF-7	T-47D	ZR-75-1
Derivation	Breast epithelium of patient undergoing reduction mammoplasty surgery, no evidence of cancer	Invasive, primary ductal cancer	Metastatic site: pleural effusion	Metastatic site: pleural effusion	Metastatic site: malignant ascites
Type	Human epithelial breast cell line - immortalised using human telomerase reverse transcriptase (hTERT)	Human epithelial breast cancer cell line	Human epithelial breast cancer cell line	Human epithelial breast cancer cell line	Human epithelial breast cancer cell line
Molecular Subtype	None	Luminal B	Luminal A	Luminal A	Luminal A
Donor Age (Years)	53	60	69	54	63
Donor Gender	Female	Female	Female	Female	Female
Reference	[166]	[167]	[168]	[169]	[170]

Table 2.2 Details of cell line complete growth medium and supplements used in this work.

Cell Line	Basal Medium	Supplements
hTERT-HME1	Dulbecco's Modified Eagle Medium (DMEM) / Nutrient Mixture F12 (F12) / GlutaMAX™	10% v/v Heat inactivated foetal calf serum (FCS) 20ng/mL Epidermal growth factor (EGF) 10g/mL Insulin 100g/mL Hydrocortisone
BT-474	Roswell Park Memorial Institute Medium-1640 (RPMI-1640) with L-glutamine	10% v/v Heat inactivated FCS 2% v/v Insulin
MCF-7	DMEM / F12 / GlutaMAX™	10% v/v Heat inactivated FCS
T-47D	RPMI-1640 with L-glutamine	10% v/v Heat inactivated FCS
ZR-75-1	RPMI-1640 with L-glutamine	10% v/v Heat inactivated FCS 1mM Sodium pyruvate
HUVEC	Endothelial Cell Growth Medium 2	Supplement mix / endothelial cell growth medium 2 which results in final supplement concentrations of: 0.02mL/mL FCS 5ng/mL EGF 10ng/mL Basic fibroblast growth factor 20ng/mL Insulin-like growth factor 0.5ng/mL Vascular endothelial growth factor 165 1µg/mL Ascorbic acid 22.5µg/mL Heparin 0.2µg/mL Hydrocortisone

2.1.3 Re-culturing from frozen cell line stock

A 15mL centrifuge tube and T-25cm² cell culture flask were prepared with 5mL and 7mL of warmed preferred complete growth medium, respectively (see Table 2.2). A 1.2mL cryovial containing the frozen cell suspension was retrieved from liquid nitrogen long-term storage and was then thawed by rolling in the palm of the hand until only a small crystal of ice was observed. The thawed cell suspension was then quickly added into the 15mL centrifuge tube containing 5mL warmed preferred complete growth medium. This was then centrifuged at 1,300 revolutions per minute (RPM) (Megafuge™ Thermo Fisher, 75004230) for 5 minutes to pellet cells. The supernatant was discarded

and the pellet was then resuspended using a Pasteur pipette and some of the preferred complete growth medium from the T-25cm² cell culture flask. The cell suspension was then aspirated and dispensed into the prepared T-25cm² cell culture flask and placed into a humidified tissue culture incubator at 37°C and 5% carbon dioxide (CO₂) under humidified conditions. After 24 hours, spent medium was aspirated, discarded and replenished with 7mL of fresh, warmed preferred complete growth medium. Cells were then allowed to grow to confluence and were then passaged as described below.

2.1.4 Maintenance of cell culture

T-75cm² flasks were used to routinely culture cells and cell culture medium was changed every 48-72 hours. This was done by aspirating and discarding spent medium and replenishing it with 12mL of warmed preferred complete growth medium (see Table 2.2). Once 70% confluence was achieved cells were passaged as detailed below.

2.1.5 Passaging of cell cultures

Once cells had achieved 70% confluency they were passaged. This was done by firstly aspirating and discarding spent medium. Cells were then washed three times with warmed phosphate-buffered saline (PBS) (pH 7.4). 5mL of 0.5% w/v trypsin/ethylenediaminetetraacetic acid (EDTA) (pH 7.4) diluted 1:9 with PBS was then added to the cells and placed into a humidified tissue culture incubator at 37°C and 5% CO₂ under humidified conditions until cells were dissociated (usually two-three minutes). During this time, T-75cm² cell culture flasks were prepared with 10mL of warmed preferred complete growth medium. Once cells were detached they were resuspended with 10mL of fresh warmed preferred complete growth medium to deactivate the trypsin. Cells were then seeded from this suspension into the prepared T-75cm² cell culture flasks at a ratio of 1:5, 1:10 or 1:20, depending upon the downstream application and the growth characteristics of the individual cell line. T-75cm² cell culture flasks were then placed into a humidified tissue culture incubator at 37°C and 5% CO₂ under humidified conditions. Cells were used for a maximum of 20 passages before being discarded and replaced with new frozen stock to minimise the effects of phenotypic drift.

2.1.6 Preparation of frozen cell line stocks

Cells were allowed to grow to approximately 80% confluency and were then trypsinised as described above. Dissociated cells were resuspended in a 1:20 dimethyl sulfoxide (DMSO) mixture in their preferred complete growth medium (see Table 2.2). 1mL of the cell suspension was added per 1.2mL cryovial. Cryovials were then placed in a Mr Frosty™ freezing container at -80°C, this limited the temperature to drop by 1°C every minute. For long-term storage, they were then transferred to liquid nitrogen.

2.1.7 Routine mycoplasma testing

Cell lines were routinely tested for mycoplasma contamination using the MycoAlert® mycoplasma detection kit, according to manufacturer's instructions.

2.1.8 Cell counting

Cell medium was aspirated from T-75cm² flasks and discarded and cells were then washed three times with warmed PBS (pH 7.4). 5mL of 0.5% w/v trypsin/EDTA (pH 7.4) diluted 1:9 with PBS was then added to the cells and placed into a humidified tissue culture incubator at 37°C and 5% CO₂ under humidified conditions until cells were dissociated (usually two-three minutes). Once cells were detached they were resuspended with 10mL of fresh warmed preferred complete growth medium to deactivate the trypsin. The cell suspension was dissociated by pipetting up and down, and a 10µL aliquot was taken and transferred into a 1.5mL microcentrifuge tube and mixed well at a 1:1 ratio with trypan blue dye. 10µL of trypan blue dye and cell mixture was then added into a cell counting slide and cells were counted using a TC20 Automated Cell Counter (Bio-Rad), two replicates were performed per cell count.

2.2 Lectin labelling

2.2.1 Cell seeding and coverslip preparation

13mm diameter coverslips were sterilised in industrial methylated spirit (IMS) and transferred into the wells of a 24-well cell culture plate. Whilst they were left to air dry in the class II cell culture hood, a

T-75cm² flask of cells at around 70% confluency was dissociated using trypsin and counted, as described in '2.1.8 Cell counting'. A suspension of cells (MCF-7 - 50,000 cells/mL, hTERT-HME1 - 100,000 cells/mL, ZR-75-1 - 100,000 cells/mL, T-47D - 100,000 cells/mL, BT-474 - 150,000 cells/mL) was created in their respective preferred complete growth medium. 1mL of cell suspension was added into each well of the plate, the plate was swirled once and then incubated at 37°C and 5% CO₂ under humidified conditions for 72 hours to allow cells to adhere to the coverslips.

2.2.2 Biotinylated lectin cytochemistry with horseradish peroxidase (HRP)

Cell culture medium was aspirated and cells were washed three times with PBS. Cells were then fixed using 4% v/v paraformaldehyde (PFA)/PBS for 20 minutes at 4°C. Cells were then washed a further three times with PBS. For permeabilized samples, permeabilization was achieved using 0.1% v/v Triton X-100/PBS for 10 minutes, and then samples were washed three more times with PBS. For non-permeabilized samples, this step was skipped. Endogenous peroxidases were blocked by adding 3% v/v hydrogen peroxide (H₂O₂)/methanol for 20 minutes. Cells were then washed three times with Tris-buffered saline (TBS) (pH 7.6). Cells were further blocked using avidin D and biotin solutions (Avidin/Biotin blocking kit), each incubated for 10 minutes with three times TBS washes between. This step blocks all endogenous biotin and endogenous avidin binding sites. Cells were then incubated with the chosen lectin, as follows: 10µg/mL biotinylated-*Helix pomatia* agglutinin (HPA) or 5µg/mL biotinylated-*Vicia villosa* agglutinin (VVA) (predominance for VVA B₄ subunits), both in 3% w/v bovine serum albumin (BSA)/TBS for 5 minutes. Cells were then washed three times with PBS and incubated for 30 minutes in 5µg/mL avidin-peroxidase in 3% w/v BSA/TBS. Cells were then washed three times with PBS and then a solution of 3,3'-diaminobenzidine tetrahydrochloride hydrate (DAB)/H₂O₂ was prepared according to manufacturer's instructions and added to the cells for two minutes, see Figure 2.1A for schematic. Cells were then washed three times with distilled water (dH₂O) and counterstained using Harris haematoxylin for two minutes. Cells were then "blued" under running tap water for approximately two minutes. Coverslips were then dehydrated through 50%, 70%, and 90% IMS and then two washes in 100% IMS. Coverslips were then cleared in xylene, inverted and mounted using DPX mounting medium onto glass slides.

2.2.3 Biotinylated lectin cytochemistry with HRP image acquisition and quantification

An upright Zeiss Axioplan microscope (Carl Zeiss, Germany) with ProgResC3 colour camera (Jenoptik, Germany) and Plan-NEOFLUAR 10x/0,3 and Plan-NEOFLUAR 40x/0,75 objectives were

used to view slides. Cells on coverslips were manually counted as either HPA/VVA-positive or HPA/VVA-negative. HPA/VVA-positive cells were identified as any cell that displayed brown labelling of intensity judged equal to or greater than ++ (i.e. all those deemed ++, +++ or ++++). Those deemed to be + or less were considered to be HPA/VVA-negative. Nine distinct images of fields of view were collected per coverslip, starting in the middle and working outwards in an anti-clockwise spiral, see Figure 2.1B. All cells within the field of view were counted unless they did not display a visibly labelled nucleus (unlabelled by Harris haematoxylin) or straddled the edges of the field of view, see Figure 2.1C for HPA scoring example and Figure 2.1D for VVA scoring example. Each biological replicate value for HPA/VVA labelling consisted of an average of three coverslips and three biological replicates per cell line were performed.

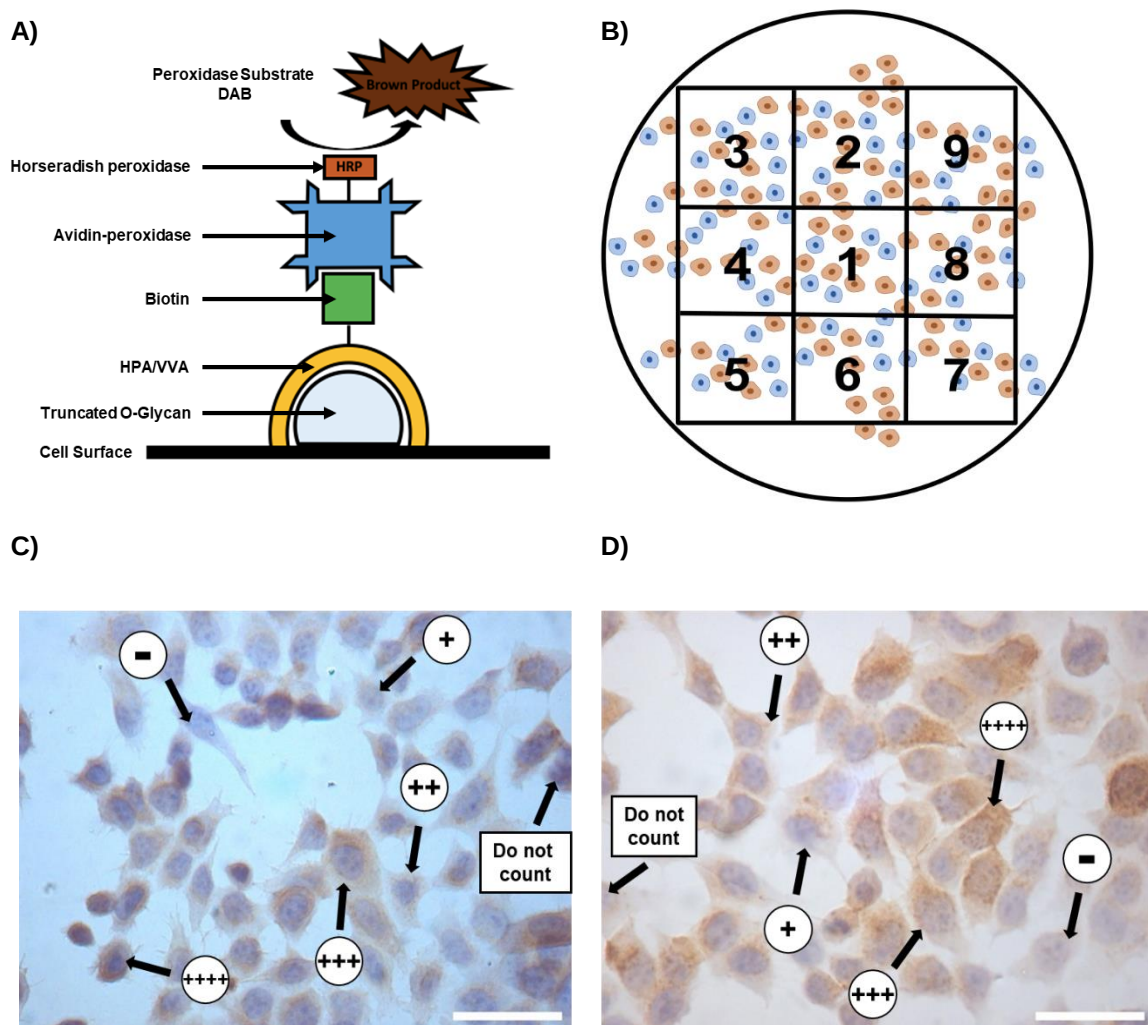


Figure 2.1 *Helix pomatia* agglutinin (HPA) and *Vicia villosa* agglutinin (VVA) lectin labelling. A) Lectin labelling protocol utilising the 'simple avidin-biotin method', adapted from [164]. B) Nine distinct images were taken per coverslip, starting in the middle of the coverslip and moving outward in an anti-clockwise spiral. C) Example of HPA scoring based on the level of displayed brown cytoplasmic intensity, ranging from – (no labelling), +, ++, +++ to +++++ (intense labelling) in MCF-7 cells. D) Example of VVA scoring based on the level of displayed brown cytoplasmic intensity, ranging from – (no labelling), +, ++, +++ to +++++ (intense labelling) in MCF-7 cells. Scale bar = 2µm.

2.2.4 Fluorescent lectin labelling

Cell culture medium was aspirated from the coverslips and cells were washed three times with PBS. Cells were then stained for 5 minutes at 37°C with CellMask™ Deep Red (1:1,000 in PBS), a cell membrane label, and incubated with 7.5µg/mL Alexa Fluor™ 488 HPA in 3% w/v BSA/TBS. This was then aspirated and coverslips were washed three times with PBS. Cells were then fixed using 4% v/v PFA/PBS for 20 minutes at 4°C. Cells were then washed a further three times with PBS. Coverslips were then mounted onto glass slides using ProLong Gold antifade mountant with DAPI and cured for 24 hours at room temperature in the dark.

2.2.5 Fluorescent lectin labelling image acquisition and quantification

Cells were imaged using an upright Zeiss confocal laser scanning microscope (LSM) 800 (Carl Zeiss, Germany) with a Plan-Apochromat 63x/1.4 oil DIC objective. Excitation and emission wavelengths were 493nm and 517nm for Alexa Fluor™ 488 HPA (0.2% laser power), 659nm and 676nm for CellMask™ deep red (0.1% laser power) and 353nm and 465nm for DAPI (1% laser power). Z-stacks of five positions randomly allocated working outwards from an initial image in an anti-clockwise spiral on the slide were acquired for each coverslip. Each technical replicate consisted of three coverslips and a total of three biological replicates were conducted.

Images were quantified in FIJI (NIH, Bethesda, MD) using an ImageJ macro script, see Supplementary Figure 1. This created a sum slice Z-stack projection and split the three channels. Using the CellMask™ deep red channel a MinError(I) threshold was applied, see Supplementary Figure 2, and a watershed was applied to identify individual cells creating regions of interest (ROIs). These ROIs were added to the ROI manager alongside one manually drawn ROI background measurement per cell. Corrected total cell fluorescence (CTCF) was calculated based on these measurements.

$$\text{CTCF} = \text{Integrated Density} - (\text{Area} \times \text{Mean Fluorescence of Background Measurement})$$

2.2.6 Lectin labelling controls

In biotinylated lectin cytochemistry experiments, three separate controls were employed to either confirm the specificity of HPA or VVA lectin binding or to act as negative controls. To control for non-specific labelling of endogenous biotin, the first control omitted the biotinylated lectin. To ensure there

was efficient blocking of endogenous peroxidases both the biotinylated lectin and avidin-peroxidase were omitted. Finally, to act as a specificity control, 0.1M α -N-acetyl-D-galactosamine (α -GalNAc) in 3% w/v BSA/TBS was added to the biotinylated lectin step. This results in competitive inhibition as both HPA and VVA lectin have an affinity for α -GalNAc monosaccharide which saturates the binding sites of the lectin, leaving it unable to bind to cells. For fluorescent lectin cytochemistry, the fluorescent lectin was omitted to check for autofluorescence. A sugar specificity control was included as described above, with 0.1M α -GalNAc, in 3% w/v BSA/TBS added to the fluorescent lectin step.

2.3 Lectin flow cytometry

2.3.1 HPA flow cytometry

Cells were grown to 70% confluency and spent cell medium was aspirated and discarded. Cells were then washed with 10mL PBS. Cells were then scraped, using a rubber scraper, into a fresh 10mL of PBS, sieved through a 70 μ m, then a 40 μ m Falcon™ cell-strainer, and then counted as described in '2.1.8 Cell counting'. A suspension of cells at 500,000 cells/mL was prepared in PBS. This was then centrifuged at 1,300 RPM (Megafuge™ Thermo Fisher, 75004230) for 5 minutes to pellet cells. The supernatant was then aspirated and 500 μ L of 4% v/v PFA/PBS was added and left for 20 minutes to fix cells. The sample was then centrifuged at 1,300 RPM for 5 minutes to pellet cells. The supernatant was aspirated and to wash cells, 1mL of PBS was added and cells were resuspended. The sample was then centrifuged at 1,300 RPM for 5 minutes to pellet cells. The supernatant was then aspirated and cell pellet was resuspended in 500 μ L of 3% w/v BSA/TBS. 7.5 μ g/mL Alexa Fluor™ 647 HPA in 3% w/v BSA/TBS was then added and cells were incubated in an end-over-end mixer at room temperature for 10 minutes. Cells were then pelleted by centrifuging at 1,300 RPM for 5 minutes. The supernatant was aspirated and cells were resuspended in 500 μ L of PBS. A total of 10,000 events at a medium flow rate (30 μ L/min) were collected, and three biological replicates were conducted. All analysis was conducted using a CytoFLEX S (Beckman Coulter) flow cytometer using CytExpert software version 2.4.

2.3.2 HPA flow cytometry controls

Three different controls were employed. The first was to confirm specificity of lectin binding. Here, 0.1M α -GalNAc in 3% w/v BSA/TBS was added to the cells at the same time as the Alexa Fluor™ 647 HPA to confirm competitive inhibition as previously described. The second was a negative control where no Alexa Fluor™ 647 HPA was added to the cells. Finally, another negative control was conducted by omitting the cells but still adding the Alexa Fluor™ 647 HPA.

2.4 Cell morphology analysis

Morphological analysis of cells labelled by lectin cytochemistry, as described in '2.2 Lectin labelling', was conducted as outlined in [171]. The freehand tool in FIJI (NIH, Bethesda, MD) was used to create ROIs on up to five HPA/VVA-positive and -negative cells per image, see Figure 2.2. Measurements were taken of the aspect ratio (length/width) and the circularity ($4\pi \cdot (\text{area}) / (\text{perimeter})^2$). A value of 1.0 indicates a perfect circle and a value that approaches 0.0 indicates an increasingly elongated shape. The major (x) and minor (y) aspects of a cell dictate the aspect ratio (x:y ratio). A value of >2:1 was classified as an elongated morphology and any value <2:1 is classified as having a rounded morphology [171]. Analysis was conducted across all nine images per coverslip, and three coverslips per technical replicate across all three biological replicates.

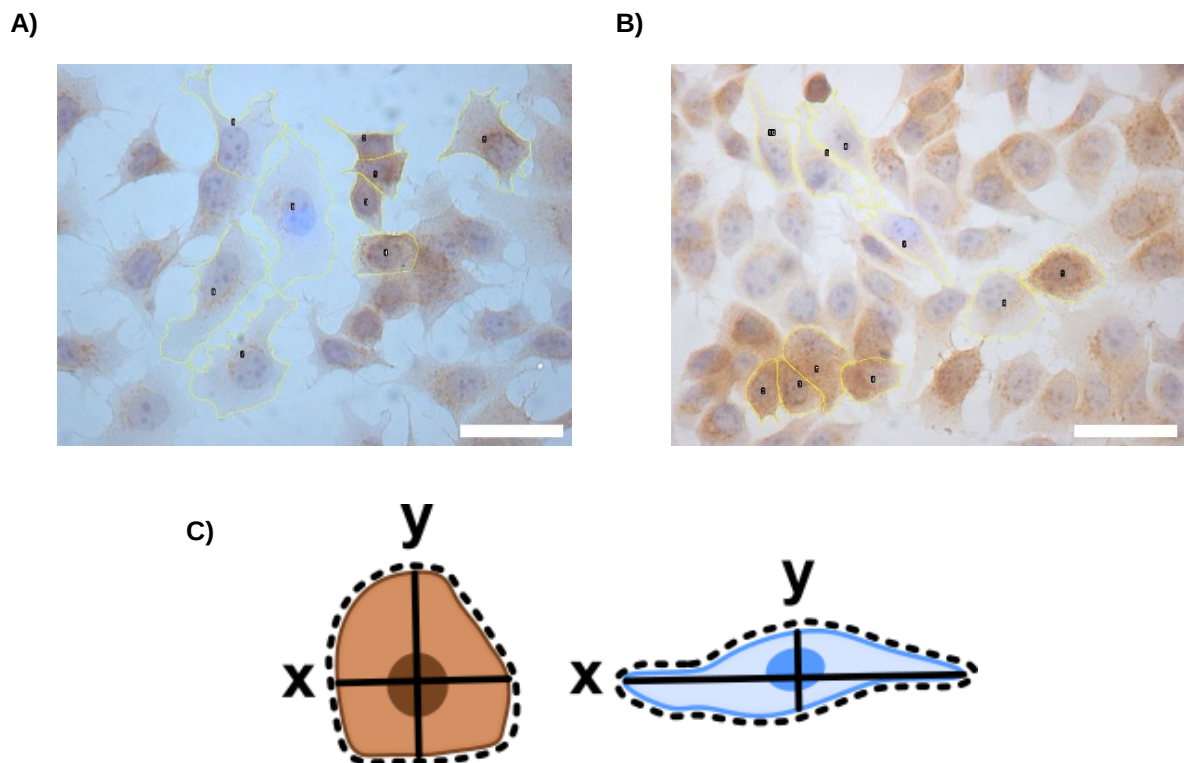


Figure 2.2 Cell morphology analysis. A) *Helix pomatia* agglutinin (HPA) morphology assessment example, MCF-7 cells. The freehand tool was used to create regions of interest (ROIs) on up to five HPA-positive and -negative cells per image. B) *Vicia villosa* agglutinin (VVA) morphology assessment example, MCF-7 cells. The freehand tool was used to create ROIs on up to five VVA-positive and -negative cells per image. Scale bar = 2 μ m. C) Representative figure of ROIs and the major (x) and minor (y) aspects of a cell.

2.5 Cell functional assays

2.5.1 Static adhesion assay

13mm diameter coverslips were sterilised in IMS, put into a 24-well cell culture plate and left to air dry in a class II cell culture hood. Once dry, coverslips were coated with 500µL of 0.2% w/v bovine gelatine/dH₂O and left to incubate for 30 minutes at 37°C. HUVEC cells were dissociated from culture flasks using trypsin and counted as described in '2.1.8 Cell counting', and a cell suspension of 100,000 cells/mL was created. Excess gelatine was aspirated and 1mL of cell suspension was added to wells, the plate was then swirled once and incubated at 37°C and 5% CO₂ under humidified conditions. HUVEC cells were grown until they were a confluent monolayer on the coverslips, with medium changed every other day. Immediately before the start of the assay, 10ng/mL tumour necrosis factor-alpha (TNF-α) in complete Endothelial Cell Growth Medium 2 was added to the confluent monolayer of HUVEC cells for two hours at 37°C and 5% CO₂ under humidified conditions, then aspirated. TNF-α inflammatory cytokines are used to activate the HUVEC monolayer by increasing expression of cell surface proteins.

10mg/mL 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS) in respective complete cell culture media, a fluorescent stain, was added into a T-75cm² cell culture flask of hTERT-HME1, MCF-7 or T-47D cells and incubated for two hours at 37°C and 5% CO₂ under humidified conditions. HPTS was then aspirated and cells were washed five times with PBS. 10mL of respective preferred cell media was added and cells scraped, using a rubber scraper, to make a suspension and counted, as described in '2.1.8 Cell counting'. A cell suspension of 50,000 cells/mL was prepared and introduced into the wells containing the HUVEC monolayers, which had been previously activated by TNF-α and incubated at 37°C and 5% CO₂ under humidified conditions for 10 minutes. Unbound breast cells and media were aspirated and plates were washed briefly three times with PBS. Cells were then fixed using 4% v/v PFA/PBS for 10 minutes at room temperature. This was then aspirated and cells were washed three times with PBS. Coverslips were removed, inverted and mounted using Fluoromount™ aqueous mounting medium onto slides. To determine the role of O-GalNAc glycans on epithelial breast cell adhesion to endothelial cells, O-GalNAc glycans were physically masked by pre-incubating breast cells with HPA. 10µg/mL of w/v HPA/TBS was incubated for 10 minutes using an end-over-end mixer with 50,000 cells/mL of breast cells in their preferred media before their addition to the HUVEC monolayer.

2.5.2 Static adhesion assay control

Since HPA is a large molecule likely to have a steric effect on the interaction of breast cells with the HUVEC monolayer, distinct from its masking of O-GalNAc glycans, a control was employed to allow for this. Using the same methodology, 10µg/mL of w/v Concanavalin A (ConA)/TBS, a lectin from

Canavalia ensiformis (jack beans) that recognises commonly present α -D-mannosyl and α -D-glucosyl residues [172] was used, see Figure 2.3A.

2.5.3 Static adhesion assay image acquisition and quantification

Once dried, slides were imaged using an upright Zeiss AxioImager widefield microscope (Carl Zeiss, Germany) with an Orca Flash 4.0 CMOS camera (Hamamatsu Photonics, Japan) and a Plan-Apochromat 10x/0.45 DICII objective, with an excitation wavelength of 488nm and emission wavelength of 509nm for enhanced green fluorescent protein (eGFP). A total of 144 tiles (12 x 12 tiles, each tile 14510.1 μ m x 14510.1 μ m) were acquired for each coverslip, which were then stitched together to form one image using ZEN 2.6 pro software, see Figure 2.3B. Image analysis was conducted using FIJI (NIH, Bethesda, MD). A threshold was applied to create a binary image and then the total number of adherent cells counted using particle analysis. Average particle count from three technical replicates was taken as total adherent cell count for each biological replicate, and three biological replicates were conducted.

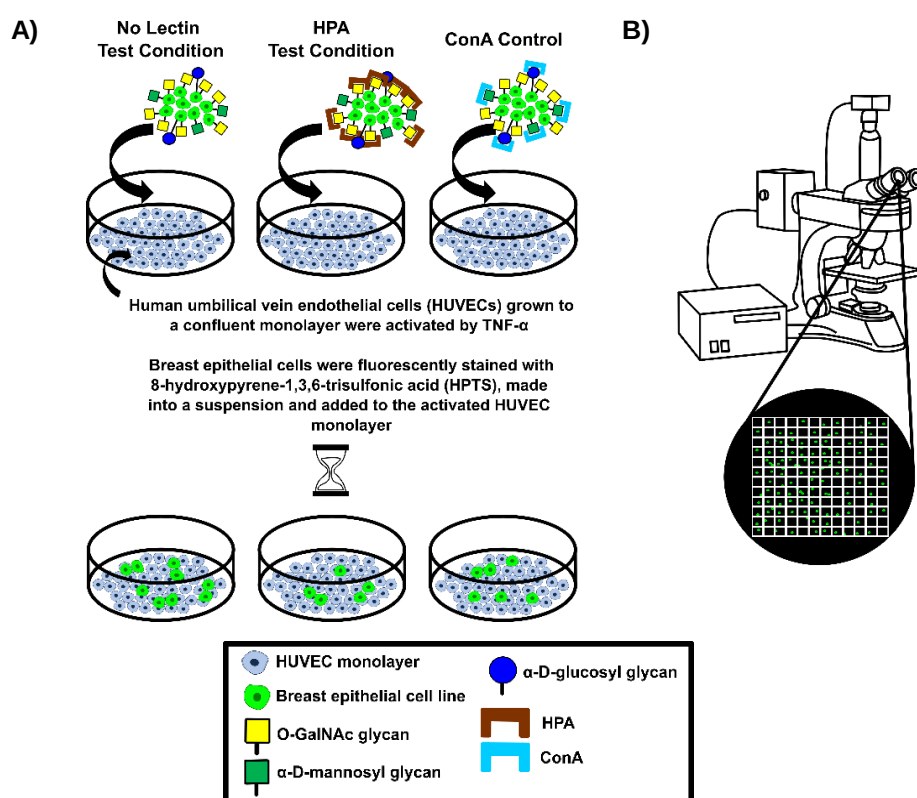


Figure 2.3 Static adhesion assay. To assess the adhesive potential of *Helix pomatia* agglutinin (HPA)-binding glycans a static adhesion assay was employed. This models cell adhesion to vascular endothelium under conditions of stasis. A) Conditions include; non-masked cells, masking cell surface glycans with the HPA or Concanavalin A (ConA) lectin. B) A total of 144 tiles (12 x 12 tiles, each tile 14510.1 μ m x 14510.1 μ m) were acquired for each coverslip, which were then stitched together to form one image using ZEN 2.6 pro software.

2.5.4 Wound healing assay

A cross comprised of two horizontal and two vertical lines was drawn on the bottom of each well of a 12-well plate using a permanent marker pen, with the centre of the cross acting as a navigation point for imaging later. Breast cells were seeded (MCF-7 - 50,000 cells/mL, T-47D and hTERT-HME1 - 90,000 cells/mL) into the plate and cultured at 37°C, 5% CO₂ under humidified conditions until they reached 90% confluency before being serum starved for 24 hours. A 'wound' was then inflicted in the cell monolayer using a p1000 pipette tip drawn down the centre of each well, and the plate was then washed twice with PBS to remove any floating cells and debris. Complete cell media was then added to each well onto the wounded cell monolayer.

A Zeiss Primovert inverted microscope with AxioCam IC (Carl Zeiss, Germany), Primo Plan-ACHROMAT 4x/0.10 objective, was used to image the cells; timepoint 0 hours. Plates were replaced into the cell culture incubator and then re-imaged at 4, 8, 24, 48 and 72-hour timepoints, see Figure 2.4. The percentage of 'wound' closure was analysed using the FIJI (NIH, Bethesda, MD) plugin "Wound_Healing_Size_Tool" [173]. To avoid bias due to unequal scratch widths, migration was calculated by subtracting from the area at timepoint 0 hours. An average of four technical replicates per biological replicate was calculated, and three biological replicates were conducted.

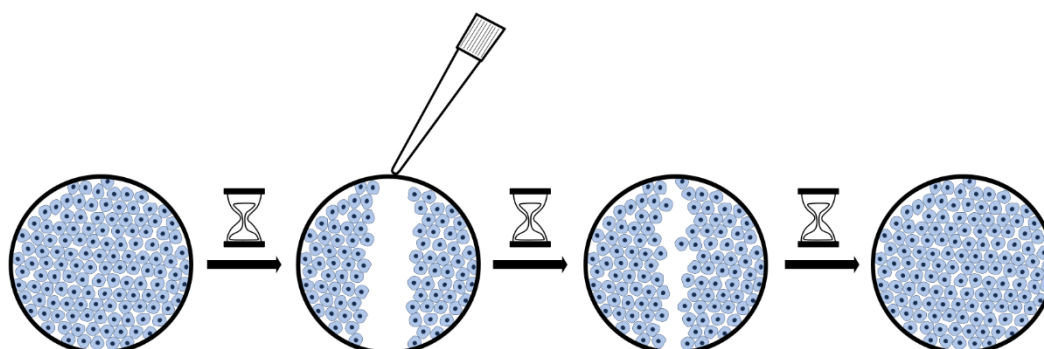


Figure 2.4 Wound healing assay. A wound healing assay was employed to investigate the motility of model epithelial cell lines in two-dimensions (2D). Cells were cultured to confluence, serum starved for 24 hours, and then a linear 'wound' was made using a sterile pipette tip. Cell migration and closure of the 'wound' were quantified over time.

2.5.5 HPA labelling of wound healing assay

13mm diameter coverslips were sterilised in IMS, put into a 24-well cell culture plate and left to air dry in a class II cell culture hood. Once dry, cells were seeded (MCF-7 - 50,000 cells/mL, T-47D and hTERT-HME1 - 90,000 cells/mL) onto the coverslips in the 24-well cell culture plate and cultured at 37°C, 5% CO₂ under humidified conditions until they reached 90% confluency before being serum

starved for 24 hours. A 'wound' was then inflicted in the cell monolayer as described in '2.5.4 Wound healing assay'. At each timepoint investigated (0, 4, 8, 24, 48 and 72 hours) the cell culture medium was aspirated and cells were washed three times with PBS. Cells were then stained for 5 minutes at 37°C with CellMask™ deep red (1:1,000 in PBS) and 7.5µg/mL Alexa Fluor™ 488 HPA in 3% w/v BSA/TBS. This was then aspirated and cells were washed three times with PBS. Cells were then fixed using 4% v/v PFA/PBS for 20 minutes at 4°C. Cells were then washed a further three times with PBS. Coverslips were then mounted onto glass slides using ProLong Gold antifade mountant with DAPI and cured for 24 hours at room temperature in the dark.

Coverslips were then imaged using a Zeiss confocal LSM 800 (Carl Zeiss, Germany) using both EC Plan-Neofluar 10x/0.3 objective for stake-tiled images of the 'wounds' and Plan-Apochromat 20x/0.8 objective for quantification images. Excitation and emission wavelengths were 493nm and 517nm for Alexa Fluor™ 488 HPA (2% laser power), 659nm and 676nm for CellMask™ deep red (0.1% laser power) and 353nm and 465nm for DAPI (2% laser power). For quantification, Z-stacks of two positions randomly allocated along the 'wound' edge and two positions randomly allocated in the main body of the 'wound' were imaged and acquired for the technical replicate, and a total of three biological replicates were conducted. Images were quantified in FIJI (NIH, Bethesda, MD) by creating a sum slice Z-stack projection and the three channels split. Using the CellMask™ deep red channel a manual threshold was applied to identify the cells from the background, this created an ROI. This ROI was then added to the ROI manager. The Alexa Fluor™ 488 HPA channel was then selected, the ROI was transposed onto it and the integrated density (AU) of ROI was measured.

2.5.6 E-cadherin, vimentin and HPA labelling of wound healing assay

Cells were seeded onto coverslips as described in '2.5.5 HPA labelling of wound healing assay' and 'wound' inflicted in the cell monolayer as described in '2.5.4 Wound healing assay'. At each timepoint investigated (0, 4, 8, 24, 48 and 72 hours) the cell culture medium was aspirated and cells were washed with PBS three times. They were then fixed using 4% v/v PFA/PBS for 20 minutes at 4°C. Cells were then permeabilized using 0.1% v/v Triton X-100/PBS for 10 minutes and then washed a further three times with PBS. They were then blocked for one hour at room temperature using 3% w/v BSA/TBS to prevent non-specific binding. Following this they were then washed three times with PBS and then incubated with the primary antibody for one hour at room temperature, either mouse anti-E-cadherin monoclonal antibody (1:100) or mouse anti-vimentin monoclonal antibody (1:500) diluted in 3% w/v BSA/TBS. Cells were washed three times with PBS and then a secondary antibody Alexa Fluor™ 594-conjugated goat polyclonal anti-mouse IgG (1:400) diluted in 3% w/v BSA/TBS was added and incubated for one hour at room temperature. This was then aspirated after the incubation period and cells were washed three times with PBS. 7.5µg/mL Alexa Fluor™ 488 HPA in 3% w/v BSA/TBS was then added to cells and incubated for 30 minutes at room temperature. This was then aspirated and cells were then washed a further three times with PBS. Coverslips were then mounted

onto glass slides using ProLong Gold antifade mountant with DAPI and cured for 24 hours at room temperature in the dark.

Coverslips were then imaged on an upright Zeiss confocal LSM 800 (Carl Zeiss, Germany) using both EC Plan-Neofluar 10x/0.3 objective for stake-tiled images of the 'wounds' and Plan-Apochromat 20x/0.8 objective for the E-cadherin quantification images. Excitation and emission wavelengths were 493nm and 517nm for Alexa Fluor™ 488 HPA (0.2% laser power), 280nm and 618nm for Alexa Fluor™ 594-conjugated goat polyclonal anti-mouse IgG (0.2% laser power) and 353nm and 465nm for DAPI (2% laser power). For quantification of E-cadherin, Z-stacks of two positions randomly allocated along the 'wound' edge and two positions randomly allocated in the main body of the 'wound' were imaged and acquired per replicate, and a total of three biological replicates were conducted. Images were quantified in FIJI (NIH, Bethesda, MD) by creating a sum slice Z-stack projection and the three channels split. Selecting only the DAPI channel the brightness of the image was elevated until the outline of the cells was visible. Using this a manual threshold was then applied to identify cells from the background, and an ROI was created and added to the ROI manager. The unaltered Alexa Fluor™ 594-conjugated goat polyclonal anti-mouse IgG antibody channel (E-cadherin channel) was then selected and the ROI was transposed onto it and the integrated density (AU) of the ROI was measured.

2.5.7 MTT assay

Cells were seeded at two seeding densities per cell line investigated (10,000 cells/mL and 20,000 cells/mL) into a 96-well plate, with 6 technical replicates per seeding condition, and three biological replicates per timepoint. Control wells used to remove background absorbance at the endpoint were also set up, these contained no cells and only complete growth media. These wells were treated the same as experimental wells throughout. The 3-(4,5-dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide (MTT) assay was performed at 24, 48 and 72-hour timepoints. MTT was dissolved in PBS at 4mg/mL in complete cell media and filtered through a 0.22µm filter. This was then added to each well at a working concentration of 2mg/mL for a 4-hour incubation at 37°C, 5% CO₂ under humidified conditions. Following the incubation, complete cell media with a working concentration of 2mg/mL MTT was carefully aspirated without disturbing the purple precipitate. The precipitate was then dissolved with 200µL of MTT solvent (4mM hydrochloric acid (HCl), 0.1% IGEPAL CA-630® in isopropanol) and absorbance was read using a SpectraMax i3x plate reader (Molecular Devices) spectrophotometer at a wavelength of 570nm, see Figure 2.5. To remove background absorbance from the media, the absorbance of wells containing only media and no cells was subtracted from that of the experimental wells.

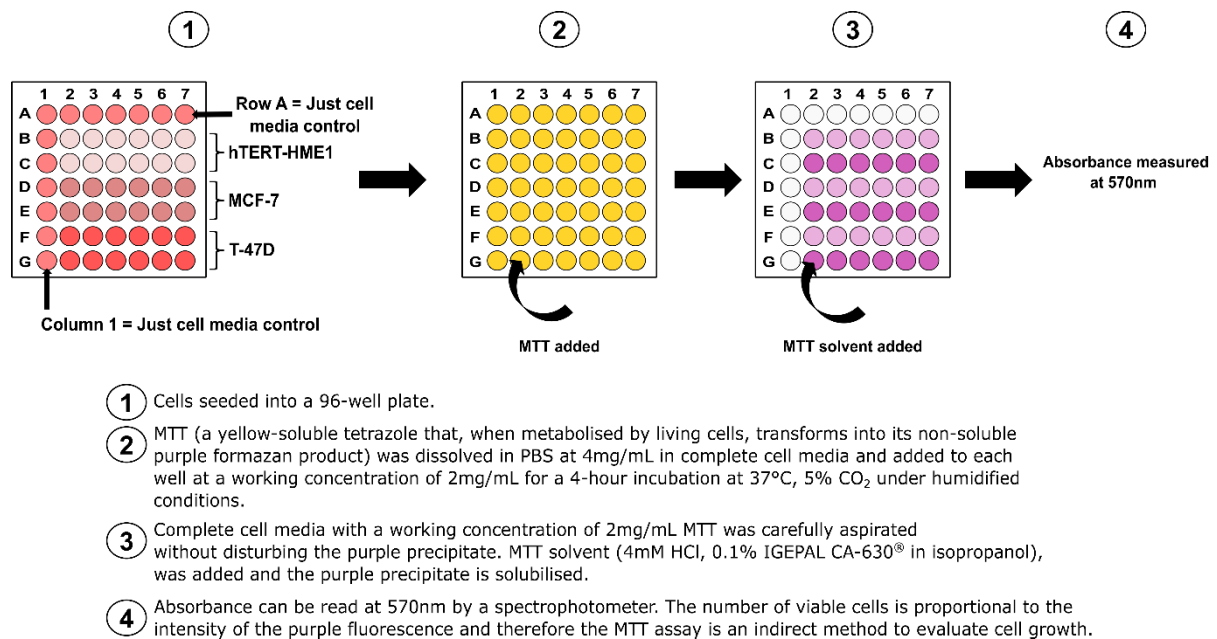


Figure 2.5 MTT (3-(4,5-dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide) assay. The rate of cell proliferation was assessed using an MTT assay.

2.6 EMT marker immunofluorescence

2.6.1 E-cadherin, N-cadherin and vimentin immunofluorescence

Cells were seeded as described in '2.2.1 Cell seeding and coverslip preparation'. Cell culture medium was aspirated and cells were washed with PBS three times. They were then fixed using 4% v/v PFA/PBS for 20 minutes at 4°C. Cells were then washed a further three times with PBS. Cells were then permeabilized using 0.1% v/v Triton X-100/PBS for 10 minutes and then washed a further three times with PBS. They were then blocked for one hour at room temperature using 3% w/v BSA/TBS to prevent non-specific antibody binding. Following this they were then washed three times with PBS and then incubated with the primary antibody for one hour at room temperature, either mouse monoclonal anti-E-cadherin (1:100), mouse monoclonal anti-vimentin (1:500) or mouse monoclonal anti-N-cadherin (1:200), all diluted in 3% w/v BSA/TBS. Primary antibodies were aspirated and cells were washed three times with PBS. Then a secondary antibody, Alexa Fluor™ 594-conjugated goat polyclonal anti-mouse IgG (1:400) diluted in 3% w/v BSA/TBS, was added and incubated for one hour at room temperature. This was then aspirated and cells were washed three times with PBS. 7.5µg/mL Alexa Fluor™ 488 HPA in 3% w/v BSA/TBS was then added to cells and incubated for 30 minutes at room temperature. This was then aspirated and cells were then washed a further three times with PBS. Coverslips were then mounted onto glass slides using ProLong Gold antifade mountant with DAPI and cured for 24 hours at room temperature in the dark.

2.6.2 E-cadherin, N-cadherin and vimentin immunofluorescence image acquisition and analysis

Coverslips were then imaged using an upright Zeiss confocal LSM 800 (Carl Zeiss, Germany), and Plan-Apochromat 63x/1.4 oil DIC objective. Excitation and emission wavelengths were 493nm and 517nm for Alexa Fluor™ 488 HPA (0.2% laser power), 280nm and 618nm for Alexa Fluor™ 594-conjugated goat polyclonal anti-mouse IgG (0.2% laser power) and 353nm and 465nm for DAPI (2% laser power). Three images were captured per coverslip, each technical replicate consisted of three coverslips and three biological replicates were conducted.

Images were quantified in FIJI (NIH, Bethesda, MD) by creating a sum slice Z-stack projection and the three channels split. Selecting only the DAPI channel the brightness of the image was elevated until the outline of the cells was visible. The freehand selection tool in FIJI (NIH, Bethesda, MD) was used to identify cells from the background, an ROI was created per cell in the image and then added to the ROI manager. The unaltered Alexa Fluor™ 488 HPA channel was then selected, ROIs transposed and the integrated density (AU) of each ROI was measured. For analysis of HPA-binding populations, HPA-positive cells were identified as cells that demonstrated >2,000 AU in the Alexa Fluor™ 488 HPA channel and HPA-negative cells were identified as <1,800 AU in the Alexa Fluor™ 488 HPA channel. These thresholds were chosen based on results in this thesis that classified hTERT-HME1 as an HPA-negative cell line with an average fluorescence of 1,900.4 (± 270.0) AU per cell. Then the unaltered Alexa Fluor™ 594-conjugated goat polyclonal anti-mouse IgG antibody channel (EMT marker channel) was selected ROIs were transposed and the integrated density of each ROI was measured to determine EMT marker immunolabelling.

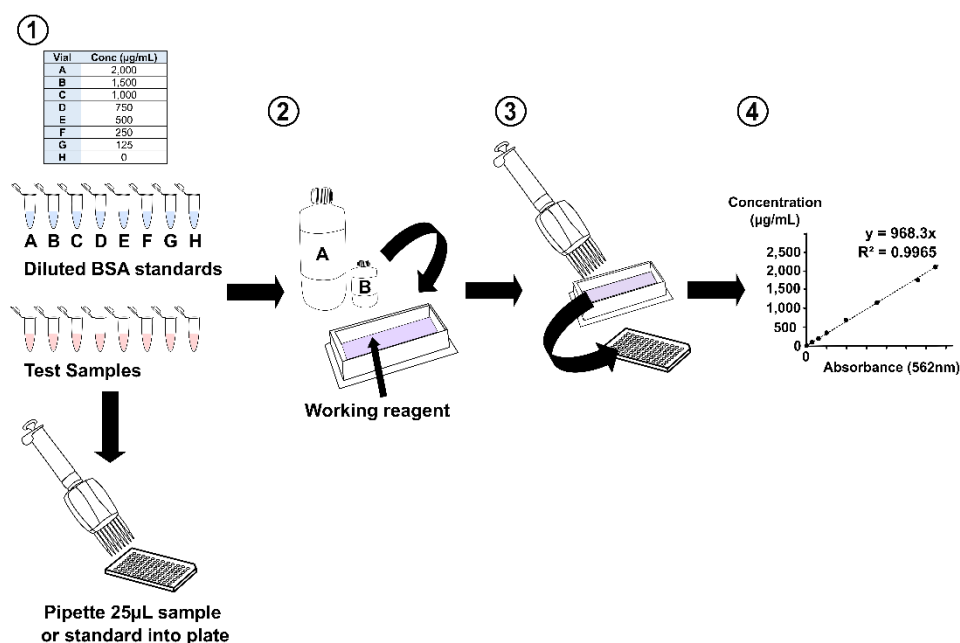
2.6.3 E-cadherin, N-cadherin and vimentin immunofluorescence controls

Two separate controls were employed, the first was a no primary antibody control (omitting either mouse monoclonal anti-E-cadherin (1:100), mouse monoclonal anti-vimentin (1:500) or mouse monoclonal anti-N-cadherin (1:200)). This was to check for any non-specific binding of the secondary antibody, Alexa Fluor™ 594-conjugated goat polyclonal anti-mouse IgG (1:400). The second control was an unlabelled sample control, which involved omitting both primary and secondary antibodies, to check for autofluorescence in the samples.

2.7 Western Blotting

2.7.1 Protein extraction and quantification from breast epithelial cell lines

One T-75cm² flask of cells was washed twice with 10mL ice-cold PBS. 1mL of 1X radioimmunoprecipitation assay (RIPA) buffer supplemented with a protease inhibitor cocktail was added to the flask of cells. Once cells were dislodged they were agitated constantly on a digital three-dimensional rocker (Sciogex, SK-D1807-S) in a 1.5mL microcentrifuge tube for 20 minutes at 4°C. Cell debris was then pelleted by centrifuging samples at 16,000xg for 20 minutes (Megafuge™ Thermo Fisher, 75004230) at 4°C. The supernatant was transferred into a new 1.5mL microcentrifuge tube and the protein concentration of cell lysates was determined by Pierce bicinchoninic acid (BCA) protein assay according to the manufacturer's protocol, see Figure 2.6 for an overview. Absorbance was read using a SpectraMax i3x plate reader (Molecular Devices) at 562nm.



- ① Bovine serum albumin (BSA) standard was diluted into a range of known concentrations. 25µL of known diluted BCA standards and test samples were added to a plate, three replicates per sample.
- ② Reagents A and B were mixed at a ratio of 50:1, respectively, to create the working reagent.
- ③ 200µL of working reagent was added to samples and incubated for 30 minutes at 37°C.
- ④ Following the incubation, the plate was equilibrated to room temperature and absorbance was read on a SpectraMax i3x plate reader (Molecular Devices) at 562nm. A standard curve was prepared by plotting the known concentration of BSA standards in µg/mL vs blank-corrected 562nm measurement. This standard curve was then used to determine the protein concentration of the unknown test samples.

Figure 2.6 Overview of Pierce bicinchoninic acid (BCA) protein assay. A colorimetric detection and quantification of total protein.

2.7.2 SDS-PAGE and Western Blotting for EMT markers

A final concentration of 20µg/µL of protein from the cell lysates was reduced in 0.1M dithiothreitol (DTT) in 4x Laemmli sample buffer at 95°C using a dry block heater (Techne) for 5 minutes, see Figure 2.7 for overview. 20µg of samples were then loaded into 8-16% Mini-PROTEAN® TGX Stain-Free™ Precast Gels alongside 5µL of Precision Plus Protein Dual Colour Standards and subjected to sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to separate proteins based on molecular weight. To achieve this, gels were immersed in 1X Tris-glycine-SDS buffer (TGS) (10X stock: 25mM Tris base, 190mM glycine, 0.1% w/v SDS in dH₂O), placed in a Bio-Rad Mini-PROTEAN® Tetra System and subjected to electrophoresis for 10 minutes at 50 volts (V) before increasing to 100V for an additional one hour.

Proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane, Trans-Blot® Turbo™ Mini-size low fluorescence PVDF Membrane (0.45µm pore) using a semi-dry Trans-Blot® Turbo™ Transfer System, Trans-Blot® Turbo™ Mini-size Transfer Stacks in Trans-Blot® Turbo™ 1X Transfer Buffer (200mL 5X Transfer Buffer, 600mL dH₂O, and 200mL ethanol). Trans-Blot® Turbo™ Mini-size Transfer Stacks were wet and equilibrated in Trans-Blot® Turbo™ 1X Transfer Buffer. Trans-Blot® Turbo™ Mini-size low fluorescence PVDF Membrane was immersed in 100% ethanol until translucent before also being equilibrated in Trans-Blot® Turbo™ 1X Transfer Buffer. A Trans-Blot® Turbo™ Mini-size Transfer Stack was then placed in the bottom (+) cassette, and the Trans-Blot® Turbo™ Mini-size low fluorescence PVDF Membrane was placed on top, followed by the proteins separated on the 8-16% Mini-PROTEAN® TGX Stain-Free™ Precast Gel and finally followed by another Trans-Blot® Turbo™ Mini-size Transfer Stack. A blot roller was used to remove any air bubbles in the assembled transfer stack and provide consistent contact between the layers. The top (-) cassette was added to complete the transfer pack and this was placed inside the semi-dry Trans-Blot® Turbo™ Transfer System. The proteins separated on the gel were transferred onto the membrane using the 'mixed molecular weight protocol' setting (2.5 amperes (A), up to 25 V for 7 minutes).

Membranes were then cut as required and unsaturated protein-binding capacity was blocked in 5% w/v skimmed milk powder in TBST (1% v/v Tween® 20) for two hours at room temperature. Primary antibodies were then incubated with the membranes for 24 hours at 4°C. This included: mouse monoclonal anti-β-actin (1:10,000), mouse monoclonal anti-E-cadherin (1:1,000), mouse monoclonal anti-vimentin (1:2,000) and mouse monoclonal anti-N-cadherin (1:500), all diluted in 5% w/v skimmed milk powder in TBST. Membranes were then washed three times for 5 minutes in TBST and then incubated with secondary antibody, goat polyclonal anti-mouse IgG HRP-conjugate (at a dilution of 1:10,000 for vimentin and β-actin labelling and 1:5,000 for E-cadherin and N-cadherin labelling), diluted in 5% w/v skimmed milk powder in TBST for one hour at room temperature. Membranes were then washed three times for 5 minutes each in TBST and then incubated with Clarity™ Western Enhanced Chemiluminescence (ECL) substrate and imaged using a ChemiDoc™ imaging system

(Bio-Rad). Band intensity was quantified using Bio-Rad Image Lab software (version 6.1) and was normalised to β -actin labelling, three biological replicates were conducted.

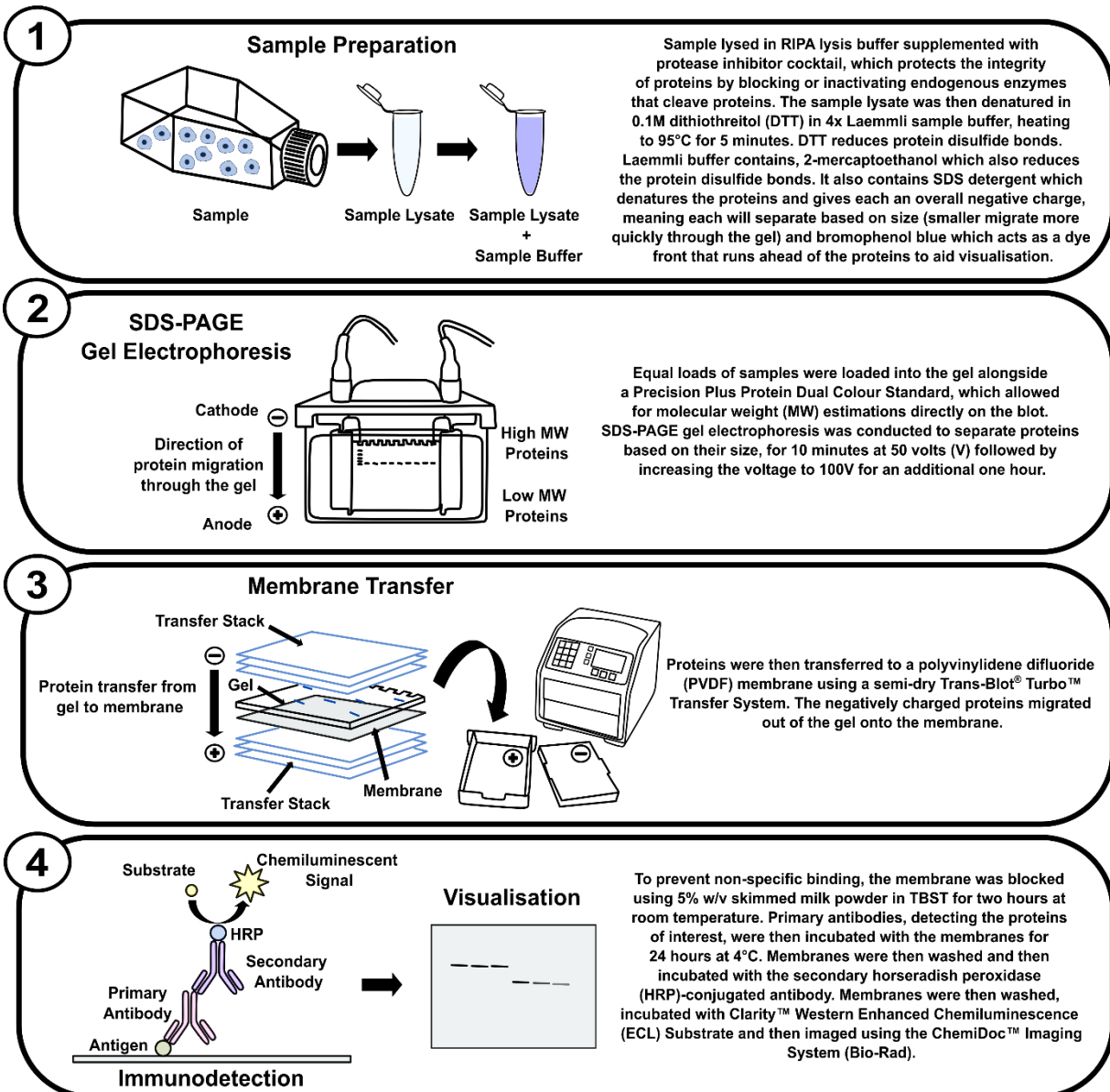


Figure 2.7 Western Blotting overview. (1) The sample was lysed in RIPA lysis buffer supplemented with a protease inhibitor cocktail. The sample lysate was then denatured in 0.1M dithiothreitol (DTT) in 4x Laemmli sample buffer, heating to 95°C for 5 minutes. (2) Equal loads of samples were loaded into the gel alongside a Precision Plus Protein Dual Colour Standard, which allowed for molecular weight (MW) estimations directly on the blot. SDS-PAGE gel electrophoresis was conducted to separate proteins based on their size, for 10 minutes at 50 volts (V) followed by increasing the voltage to 100V for an additional one hour. (3) Proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane using a semi-dry Trans-Blot® Turbo™ Transfer System. (4) To prevent non-specific binding, the membrane was blocked using 5% w/v skimmed milk powder in TBST for two hours at room temperature. Primary antibodies, detecting the proteins of interest, were then incubated with the membranes for 24 hours at 4°C. Membranes were then washed and then incubated with the secondary horseradish peroxidase (HRP)-conjugated antibody. Membranes were then washed, incubated with Clarity™ Western Enhanced Chemiluminescence (ECL) Substrate and then imaged using the ChemiDoc™ Imaging System (Bio-Rad).

2.8 TGF- β 1 induced EMT

2.8.1 TGF- β 1 treatment

Cells were seeded onto coverslips as described in '2.2.1 Cell seeding and coverslip preparation' (MCF-7 - 8,000 cells/mL and T-47D - 18,000 cells/mL). Complete cell culture media was aspirated and replaced with 1% v/v FCS preferred cell culture media for 24 hours. Following this, 1% v/v FCS preferred cell culture media containing 10ng/mL TGF- β 1 was added and replenished every other day. Control cells were cultured in 1% v/v FCS preferred cell culture media, and replenished at the same time as those subject to treatment conditions.

2.8.2 TGF- β 1 treatment and Western Blotting

For Western Blotting, cells were seeded into a 6-well cell culture plate (MCF-7 - 10,000 cells/mL and T-47D - 18,000 cells/mL). Cell culture media was aspirated and replaced with 1% v/v FCS preferred cell culture media for 24 hours. Following this, 1% v/v FCS preferred cell culture media containing 10ng/mL TGF- β 1 was added and replenished every other day. Control cells were cultured in 1% v/v FCS preferred cell culture media, and replenished at the same time as those subject to treatment conditions. Cells at the 72-hour timepoint were analysed in the Western Blot analysis. Cells were washed twice with 2mL ice-cold PBS. 500 μ L of 1X RIPA buffer supplemented with a protease inhibitor cocktail was added to the flask of cells, once cells were dislodged they were agitated constantly on a digital three-dimensional rocker (Sciogex, SK-D1807-S) in a 1.5mL microcentrifuge tube for 20 minutes at 4°C. They were then pelleted by centrifuging samples at 16,000xg (Megafuge™ Thermo Fisher, 75004230) for 20 minutes at 4°C. The supernatant was transferred into a fresh 1.5mL microcentrifuge tube and protein concentration of the cell lysates was determined by Pierce BCA protein assay according to the manufacturer's protocol. Absorbance was read on a SpectraMax i3x plate reader (Molecular Devices) at 562nm. Western Blotting was then conducted as described in '2.7.2 SDS-PAGE and Western Blotting for EMT markers'.

2.8.3 Identifying effects of TGF- β 1 treatment on cell morphology, HPA labelling and EMT markers

Timepoints of 24, 48, 72 hours and 7 days post initial TGF- β 1 treatment were analysed for MCF-7 cells. Timepoints of 72 hours and 7 days post initial TGF- β 1 treatments were analysed for T-47D cells. Images were captured under phase contrast using a Zeiss Primovert inverted microscope with

AxioCam IC (Carl Zeiss, Germany) and a Primo Plan-ACHROMAT 10x/0.25 Ph1 objective. Morphological analysis of cells was conducted as previously described in '2.4 Cell morphology analysis'. This time, the freehand tool in FIJI (NIH, Bethesda, MD) was used to create ROIs of 300 or all the whole cells (if less than 300), in an image, see Figure 2.8. This was done across three images per technical replicate, three biological replicates were conducted. HPA labelling was performed using the methods described in '2.2.4 Fluorescent lectin labelling'. Changes to EMT markers were analysed using the methods as described in '2.6.1 E-cadherin, N-cadherin and vimentin immunofluorescence' and '2.7.2 SDS-PAGE and Western Blotting for EMT markers'.

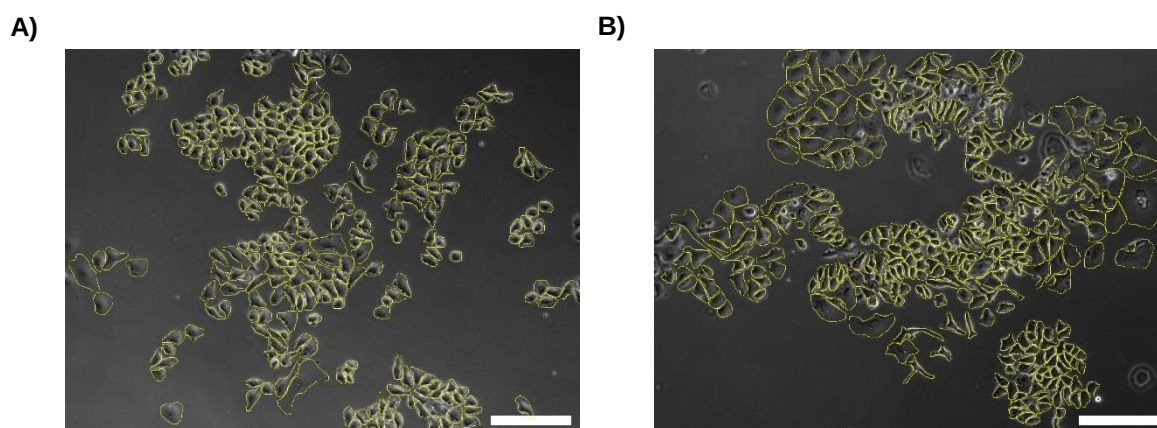


Figure 2.8 Representative phase contrast images of freehand tool region of interest (ROI) selection overlay of control and transforming growth factor- β 1 (TGF- β 1) treatment conditions in MCF-7 cells. A) Control MCF-7 cells at 24 hours. B) TGF- β 1 treated MCF-7 cells at 24 hours. Scale bar = 50 μ m.

2.9 Lectin-coated magnetic bead isolation of glycosylation-specific cell populations

The method, employing lectin-coated magnetic beads, described by [174] was employed under sterile conditions to separate cells into distinct populations based on their cell surface lectin-binding profiles.

2.9.1 Preparation of cells

MCF-7 and T-47D cells were grown to 70% confluency in T-75cm² cell culture flasks and were washed with PBS. This was then aspirated and discarded and cells were scraped from the culture flask or plate into fresh PBS using a sterile rubber cell scraper. To yield a single-cell suspension cells were sieved through a 70 μ m then a 40 μ m Falcon™ cell-strainer. Cells were then pelleted by centrifugation at 1,300 RPM (Megafuge™ Thermo Fisher, 75004230) for 5 minutes and the supernatant was aspirated and discarded.

2.9.2 Preparation of HPA-coated magnetic beads

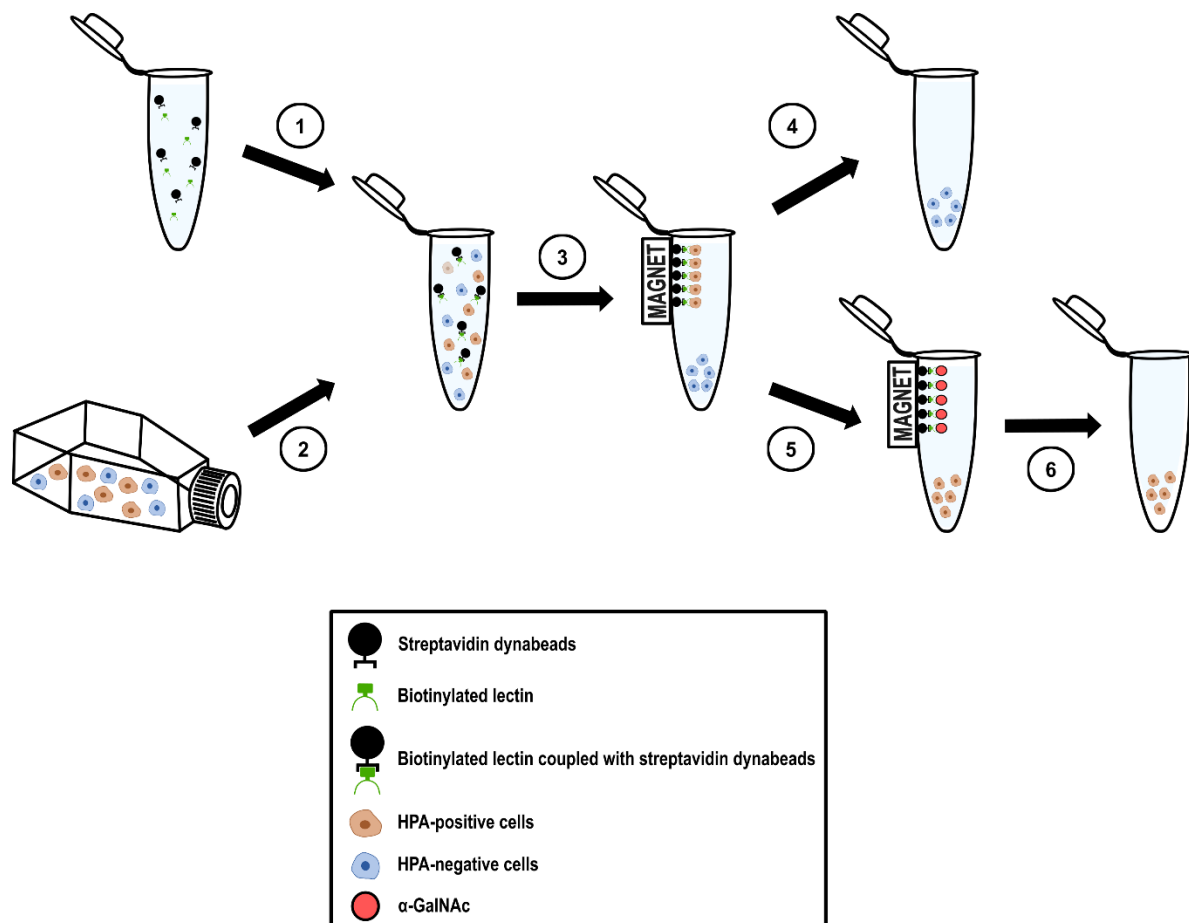
100µL of streptavidin-Dynabeads® (pH 7.4, at a concentration of 10mg/mL PBS containing 0.1% BSA and 0.02% sodium azide (NaN₃)) was vortexed to mix and aliquoted into a 1.5mL microcentrifuge tube, then beads pooled to the side of the tube using a magnet for two minutes. The supernatant was aspirated and discarded then the magnet was removed and 500µL of 3% w/v BSA/TBS was added to wash the beads. After mixing by pipetting, the magnet was replaced onto the side of the tube for two minutes to pool the beads and the supernatant was aspirated and discarded. This wash stage was repeated a further two times. Then, the magnet was removed and 10µL of 10µg/mL biotinylated-HPA/TBS was added to the beads and incubated for 60 minutes at room temperature on an end-over-end mixer. After this incubation period, the HPA-coated beads were pooled using a magnet, as before, for two minutes. The unbound-lectin/buffer supernatant was then aspirated and discarded and the magnet was removed from the tube. The beads were then washed with 500µL 3% w/v BSA/TBS, pipetting to mix, and the magnet was replaced onto the tube for two minutes to pool the beads. The supernatant was aspirated and discarded. This wash stage was repeated a further two times and once completed the HPA-coated beads were stored in 1mL of 3% w/v BSA/TBS.

2.9.3 Separating glycosylation-specific cell populations

The pelleted cells, prepared as described in '2.9.1 Preparation of cells', were combined with 1mL of HPA-coated beads in a 1.5mL microcentrifuge tube and placed on an end-over-end mixer for 5 minutes at room temperature. A magnet was then placed on one side of the tube for two minutes at room temperature to pool the beads. The supernatant and unbound cells were gently aspirated into a new 1.5mL microcentrifuge tube (this is the negative population i.e. the cells without the targeted HPA-binding glycosylation profile). Once the supernatant had been aspirated the magnet was removed from the tube and the positive population (i.e. the cells that have the targeted HPA-binding glycosylation profile) was resuspended in 1mL 3% w/v BSA/TBS.

2.9.4 Dissociation of cells from HPA-coated magnetic beads using competitive inhibition of lectin-binding

To dissociate the cells from the HPA-coated magnetic beads, beads were re-suspended in 500µL of 0.1M α-GalNAc in 3% w/v BSA/TBS and placed on an end-over-end mixer for 5 minutes at room temperature. A magnet was then placed on the side of the tube for two minutes to pool the beads. The dissociated HPA-binding cells in the supernatant were then aspirated into a new 1.5mL microcentrifuge tube, see Figure 2.9.



- 1 Streptavidin Dynabeads[®] in 3% w/v BSA/TBS and biotinylated HPA are combined and mixed in an end-over-end mixer for 60 minutes to create HPA-coated magnetic beads.
- 2 Once grown to 70% confluency, cells are removed from the cell culture flask and mixed with the HPA-coated magnetic beads for 5 minutes.
- 3 A magnet is placed on the side of the tube and the HPA-binding cells are isolated as they collect on the side of the tube, bound to the HPA-coated magnetic beads.
- 4 The cells that did not bind to the HPA-coated beads are left in the supernatant and are removed into a new 1.5mL microcentrifuge tube.
- 5 0.1M α -GalNAc in 3% w/v BSA/TBS can be added to disassociate the cells from the HPA-coated beads via competitive inhibition. The application of a magnet means that the HPA-coated beads collect on the side of the tube and release cells are left in the supernatant.
- 6 Cells are removed and placed into a new 1.5mL microcentrifuge tube.

Figure 2.9 Workflow of *Helix pomatia* agglutinin (HPA)-coated magnetic bead isolation of glycosylation-specific cell populations.

2.10 Ribonucleic acid (RNA) sequencing

2.10.1 Total RNA isolation and DNase treatment

The following experimental groups were: MCF-7 HPA-positive, MCF-7 HPA-negative, T-47D HPA-positive, T-47D HPA-negative and 72-hour MCF-7 TGF- β 1 treated cells. For the MCF-7 TGF- β 1 treated cells, cells were washed and extracted from a 6-well cell culture plate. MCF-7 and T-47D glycosylation-specific cell populations were isolated as described in '2.9 *Lectin-coated magnetic bead isolation of glycosylation-specific cell populations*'. RNA extraction was then performed on these experimental groups using RNeasy® Mini Kit (50) and RNase-free DNase set (50), conducted according to the manufacturer's instructions 'Quick-start protocol RNeasy® Mini Kit, parts 1 and 2', see Figure 2.10 for workflow. Four biological replicates were conducted per experimental group.

Sequencing was carried out by an external provider, Novogene, Hong Kong. Their sample guidelines were: (1) RNA concentration $\geq 20\text{ng}/\mu\text{L}$ in a minimum of $10\mu\text{L}$, (2) total RNA content $\geq 200\text{ng}$ per sample, (3) RNA integrity number (RIN) value ≥ 4.0 and (4) A_{260}/A_{280} and A_{260}/A_{230} value of ≥ 2.0 .

The RNA concentration and RNA integrity number equivalent (RINe) were measured using Agilent Technologies 4150 TapeStation using RNA ScreenTape, RNA ScreenTape sample buffer and RNA ScreenTape ladder, see Figure 2.11. Briefly, RNA ScreenTape sample buffer was allowed to equilibrate to room temperature for 30 minutes whilst RNA ScreenTape ladder was thawed on ice. RNA samples were tested straight after isolation and kept on ice until use. The RNA ScreenTape was then inserted into the TapeStation instrument and sample positions were selected. Reagents and samples were vortexed for one minute and spun down for one minute at 6,000 RPM using a Prism™ mini centrifuge (Labnet™, C1801-230V-UK) before use. $5\mu\text{L}$ of RNA sample buffer and either $1\mu\text{L}$ of RNA sample or RNA ladder (one ladder per run) were added into a tube strip and caps were added. Samples and ladder were then mixed by vortexing for one minute at 2,000 RPM using an IKA® MS3 Vortexer (Merck). Samples and ladder were then spun down for one minute (6,000 RPM, Prism™ mini centrifuge). Following this, they were then denatured by heating to 72°C for three minutes. They were then put on ice for two minutes and spun down (6,000 RPM, Prism™ mini centrifuge) for one minute. The ladder was then placed at position A1 and samples were loaded into the TapeStation, caps were removed and sample RNA concentration and RINe were measured. Samples with RINe values ≥ 8.0 were used for downstream sequencing.

The RNA concentration was also assessed along with the RNA purity using a Thermo Fisher Scientific NanoDrop™ One spectrophotometer using ratios of absorbance at wavelengths of 260nm and 280nm (A_{260}/A_{280}) and at 260nm and 230nm (A_{260}/A_{230}). A $1\mu\text{L}$ sample of Ambion™ nuclease-free water was used as a blank before loading samples, and $1\mu\text{L}$ of the sample was loaded and measured. Samples with ratio values of around 2.0 for both A_{260}/A_{280} and A_{260}/A_{230} were used.

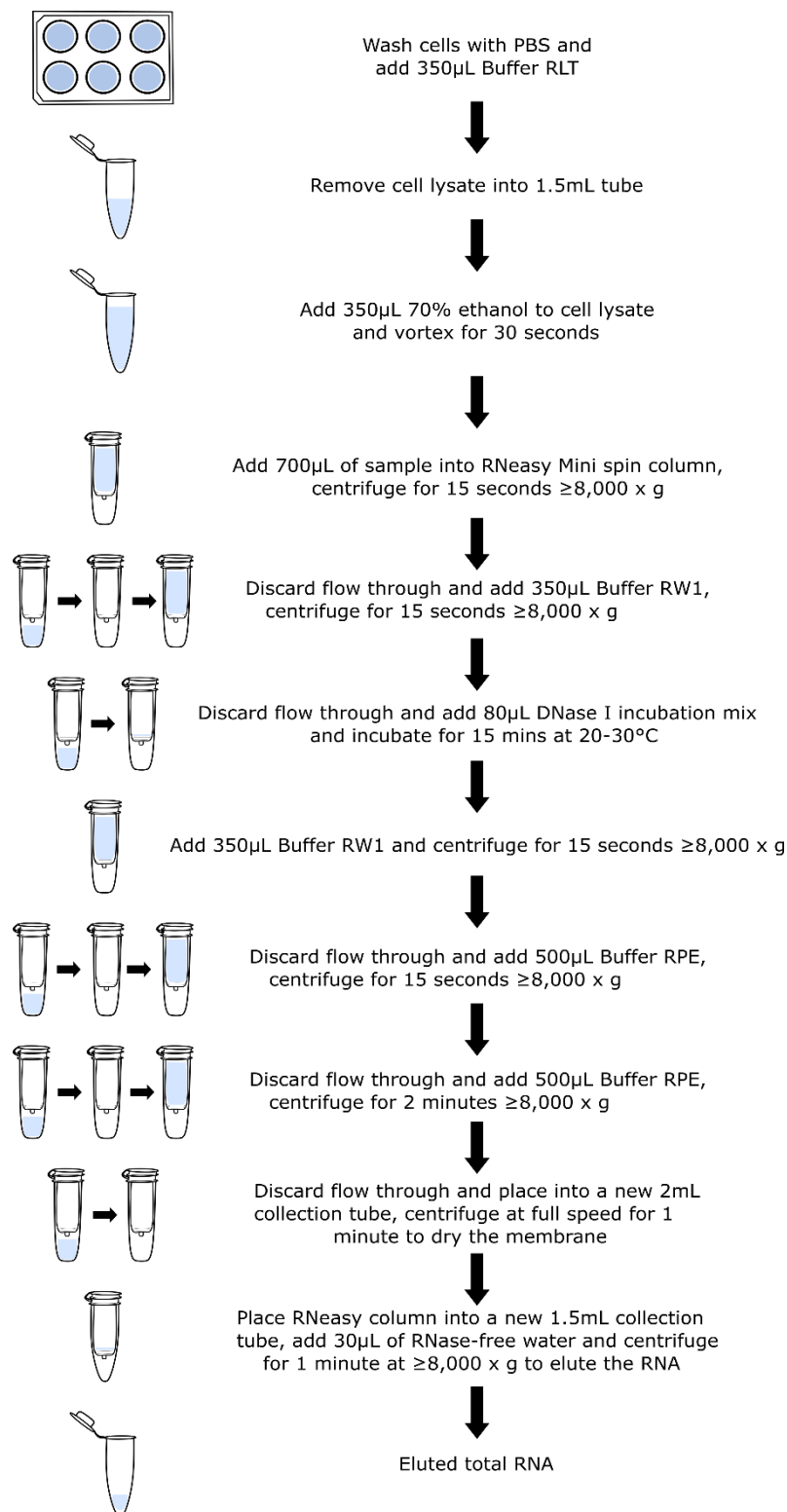


Figure 2.10 Total RNA isolation and DNase treatment workflow. RNeasy® Mini Kit (50) (QIAGEN GmbH) was used for the RNA extraction and was conducted according to the 'Quick-start protocol RNeasy® Mini Kit, parts 1 and 2'.

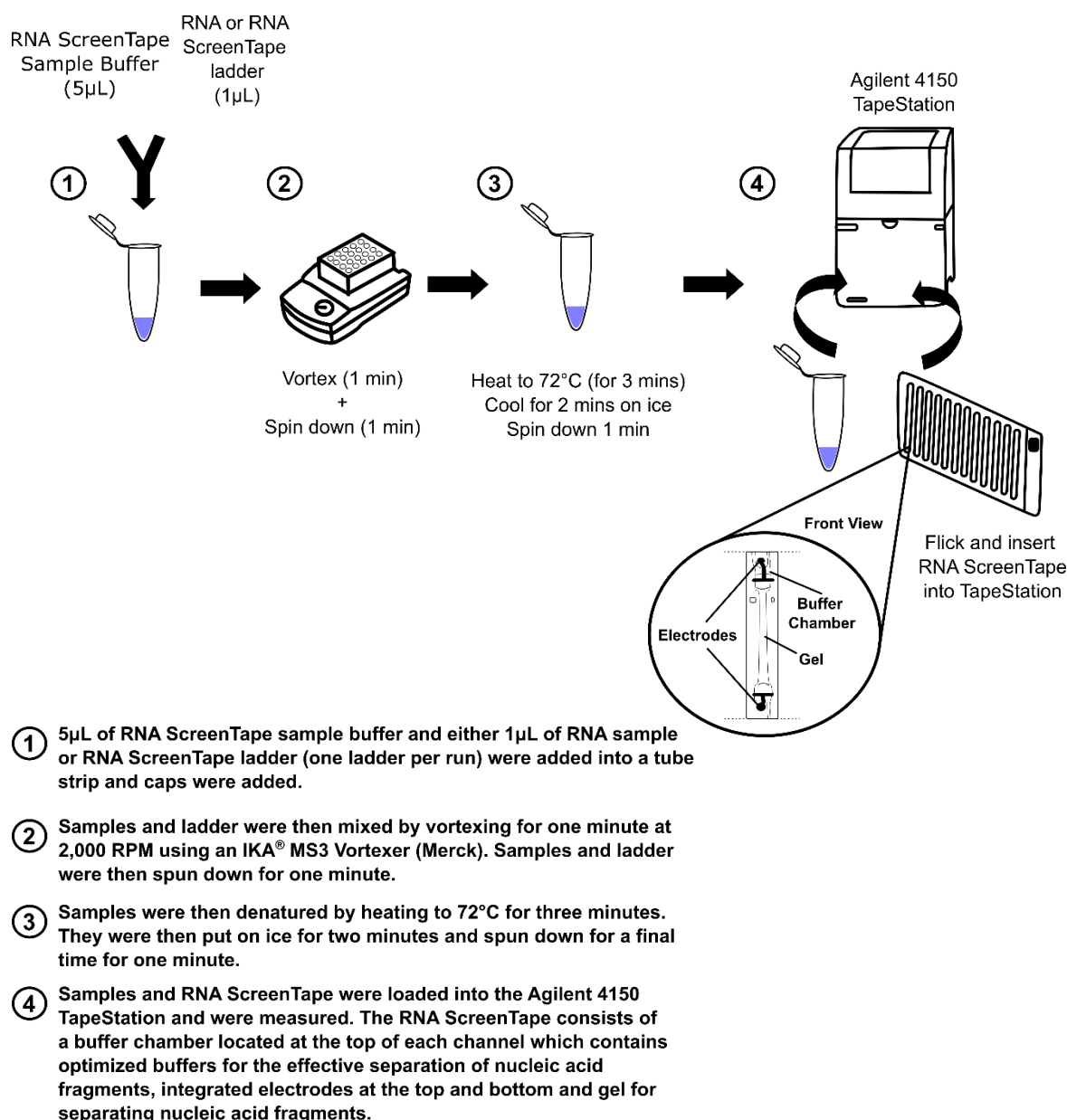


Figure 2.11 Workflow of Agilent Technologies 4150 TapeStation measurement of RNA concentration and RNA integrity number equivalent (RINe) value.

2.10.2 Novogene RNA quality control

Chosen samples that met the sequencing sample guidelines above were sent to Novogene, where upon receipt they underwent post-shipping quality control checks to ensure that they were still of sufficient quality and quantity. Sample quantification of concentration (ng/µL), volume (µL), total amount (µg), integrity and purity (RIN scores) were analysed using an Agilent 5400 fragment analyser system (Agilent Technologies).

2.10.3 RNA sequencing by Novogene

Samples were processed for library preparation, see Figure 2.12 for an overview of the workflow performed by Novogene. The following methods described in the remainder of section 2.10 were performed by Novogene: the RNA library constructed was quality-checked using Qubit (Thermo Fisher) and real-time polymerase chain reaction (PCR). Sequencing was performed using an Illumina (San Diego, California) platform and paired-end 150 base pair (bp) (PE150) reads were generated into a FASTQ file, with ≥ 20 million read pairs per sample. The sequenced data were then subjected to data quality control to create 'clean data'. Novogene then trimmed sequences to remove reads containing the adaptor sequences and low-quality reads. Phred scores, or quality or Q scores, were assigned for each nucleotide base call which indicates the probability the base call is incorrect. A Q30 score indicates a 1 in 1,000 (0.1%) probability that the base is being called incorrectly and a Q20 score indicates a 1 in 100 probability (1%). Reads with uncertainty levels of $>10\%$ ($N>10\%$) were also removed.

2.10.4 Analysis of sequenced data by Novogene

'Clean data' created from the data quality control was then used by Novogene to map onto the reference genome: the human genome (hg38 from NCBI/UCSC/Ensembl) using HISAT2 (version 2.0.5). Gene expression quantification was then conducted by Novogene using the software FeatureCounts (version 1.5.0-p3), and reads per kilobase per million (RPKM) values were calculated for genes. Raw read count data was used to assess differential expression analysis using edgeR (version 3.22.5) and DESeq2 (version 1.20.0) software to identify differentially expressed genes (DEGs). The DEGs were then subjected to function enrichment analysis by Novogene which included Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway enrichment analysis using the ClusterProfiler software (version 3.8.1).

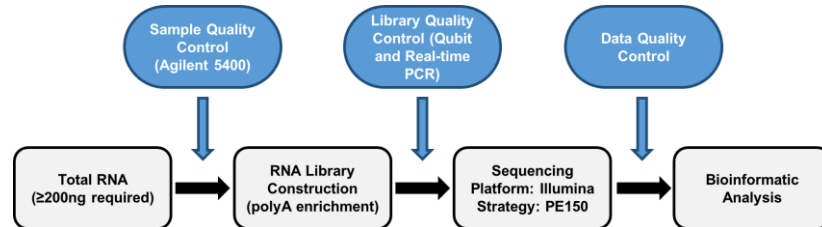
2.11 Statistical analysis

All statistical analysis was conducted in GraphPad Prism software version 10.2.1. Each data set was tested for normality using the Shapiro-Wilk test and Quantile-Quantile (Q-Q) plots. Subsequently, the recommended parametric or non-parametric alternative with multiple comparison tests was conducted. If data were normally distributed a Student's t-test was conducted. If the sample was not normally distributed a Mann-Whitney U test was conducted for unpaired samples, and a Wilcoxon matched-pair signed rank test was conducted for paired samples. All statistical tests are referred to in

the figure legends, and probability (p)-values are indicated using asterisks denoting the following:

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, and **** $p \leq 0.0001$.

A)



B)

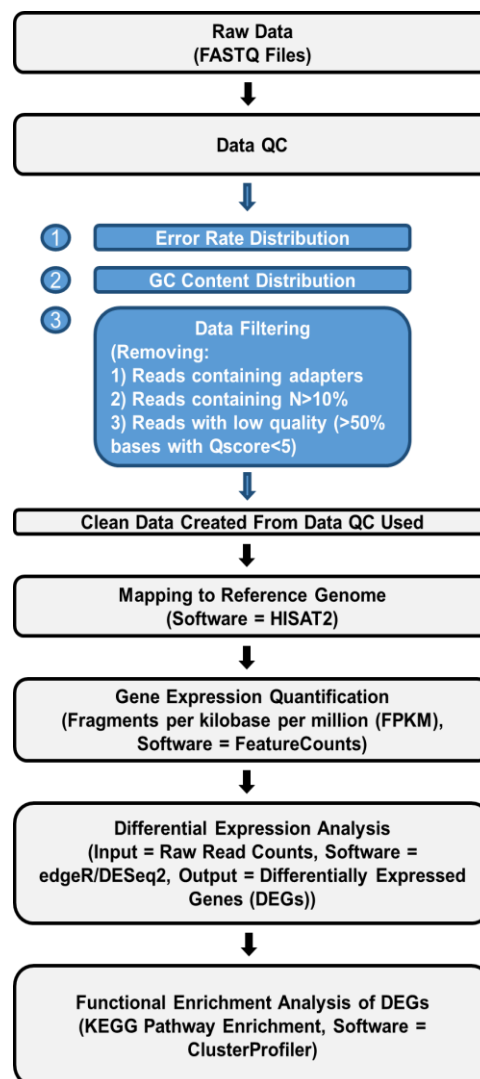


Figure 2.12 Workflow overview of RNA Sequencing performed by Novogene. A) RNA library construction, sequencing and bioinformatic analysis. Quality control steps were performed for sample, library and data quality – highlighted in blue. B) General overview of bioinformatic analysis workflow performed by Novogene. Data quality control steps are highlighted in blue.

Chapter 3

Characterisation of model breast epithelial cell lines

3. Characterisation of model breast epithelial cell lines

3.1 Background

3.1.1 Altered glycosylation in cancer

It is well-established that in pathologies such as cancer, there are frequent changes in glycosylation (reviewed by [175], [16], [176], [177] and [178]). The two main archetypes of altered glycosylation that are seen in cancers are; i) incomplete synthesis or truncation resulting in exposure of usually cryptic glycans (e.g. Tn antigen, sTn antigen), and ii) neo-synthesis of *de novo* glycans (e.g. sialyl Lewis a (SLe^a) and sialyl Lewis x (SLe^x)) [179]. Model cancer cell lines that display these altered glycans are therefore of great utility when investigating the effects of changes in glycosylation. Lectins, as described in '1.2 Lectins', are useful tools for glycobiologists to help identify these changes in glycosylation and can be utilised in a variety of applications, including both lectin cytochemistry and flow cytometry (reviewed by [95]).

3.1.2 *Helix pomatia* agglutinin (HPA) binding as a prognostic marker in cancer

HPA is a lectin that is derived from the albumin gland of the Roman or edible snail and has nominal specificity for terminal α -GalNAc monosaccharide residues. HPA also binds with varying affinity to a wide variety of other structures, see Table 3.1. Elucidating the crystal structure of HPA has demonstrated that a β -sandwich lectin fold is the reason for a preference for binding α -GalNAc structures, like Tn antigen [180, 181]. Leathem and Brooks [182] first described that the presence of HPA-binding glycans in paraffin wax-embedded sections of primary breast cancer was associated with poor long-term patient prognosis, aggressive biological behaviour and metastasis. Since the initial study, other studies have confirmed that HPA labelling is associated with the prognostic result of shorter survival times and association with lymph node metastasis in breast cancer [183-188]. Studies into other cancers such as gastric [189], prostate [190], colorectal [191, 192], oesophageal [193], thyroid [194] and lung adenocarcinomas [195] have also demonstrated the prognostic utility of HPA labelling.

Breast cancer cell lines have been previously characterised for their levels of HPA labelling. Within our research group, MCF-7, T-47D, BT-474 and ZR-75-1 breast cancer cell lines have been documented as stably synthesising a heterogenous profile of HPA-binding glycoproteins which are also representative of those seen in clinical breast cancer samples [39, 164, 165]. Brooks et al., [164]

demonstrated that MCF-7 cells displayed intense (+++++) labelling in HPA lectin cytochemistry and 98.66% of cells displayed HPA labelling in flow cytometry. BT-474 cells demonstrated weak (+) labelling and 31.00% of cells were labelled with HPA in flow cytometry. ZR-75-1 cells demonstrated a +++ intensity score in lectin cytochemistry and 73.56% of cells were labelled for HPA in flow cytometry analysis. These results demonstrated similar HPA-binding profiles compared to cells from clinical samples of the same origin.

In the literature, lectin histochemistry and cytochemistry results have often been scored, by eye, by two independent observers. A scoring method commonly used is to evaluate the intensity of the labelling combined with an estimate of the percentage of cells labelling positively. Leathem and Brooks [182] first reported a “binary” HPA scoring system in paraffin wax-embedded primary breast cancers, ‘negative’ (+/-) determined by little to no HPA labelling and ‘positive’ (+ to +++++), which demonstrated definite, strong HPA labelling. This reflected the two main HPA labelling patterns seen, non or occasional and most or all. Brooks et al., [196] used (+/-) to denote very weak, (+) weak through to (+++++) very strong HPA labelling of clinical breast cancer samples, combined with an estimate of the percentage of cancer cells that were labelling e.g. 80% +++, 15%+/- . Ultimately, a simple binary of “positive stainers” (defined as 5% or more of cells labelling with an intensity of + or more, or more than 50% with an intensity of +/-) or negative “non-stainers” was used for the analysis of the relationship of lectin labelling with prognosis [196], the same methodology employed previously by Brooks and Leathem [186]. There was no relationship between the level of HPA labelling and prognosis, rather simply its presence or absence was sufficient hence, a binary has often been used [182, 184-186, 196].

3.1.3 *Vicia villosa* agglutinin (VVA) binding as a prognostic marker in cancer

VVA, a lectin isolated from hairy vetch seeds, comprises two subunits, an A₄ and a B₄ subunit. The B₄ subunit has nominal specificity for terminal α -N-acetyl-D-galactosamine (α -GalNAc) monosaccharide residues whereas, the A₄ subunit will agglutinate A₁ erythrocytes as well [197, 198], see Table 3.1. Kawaguchi et al., [199] demonstrated that VVA-binding carbohydrates are associated with tumour stage, lymphatic invasion and lymph node metastasis in primary breast cancers. They were not, however, associated with disease recurrence or a 5-year survival rate. The authors also used an anti-Tn monoclonal antibody, HB-Tn1, and failed to find a relationship with tumour stage, lymphatic invasion or lymph node metastasis. The difference between HB-Tn1 and VVA is that HB-Tn1 is thought to recognise a different carbohydrate epitope(s). VVA was subsequently demonstrated to bind to both single Tn antigen structures and also clusters of Tn, whereas HB-Tn1 may only bind to clusters of Tn, identified using affinity chromatography [199, 200]. Similar to HPA, VVA binding has been significantly associated with aggressive biological behaviour and metastasis in a wide range of

cancer types. For example, it has been associated with histological grade and disease recurrence in bladder cancer [201], metastatic potential and poor prognosis in cervical cancer [202, 203], lymph node metastasis in gastric cancer [204], and lymphatic invasion, lymph node metastasis in colorectal cancer [205].

Breast tissues have also been investigated for their levels of VVA labelling using subjective scoring systems. Konska et al., [206] demonstrated that normal breast tissue was entirely negative for VVA labelling, or in one case weakly labelled (+) in the cytoplasm. Primary breast cancers were moderately labelled (++) and metastases were labelled intensely (+++). Korourian et al., [207] used a four-point scoring system (0 for negative, 1,2 to 3 for strongly positive) to demonstrate that labelling was associated with tumour grade. They also showed that aggressive breast cancer cell line MDA-MB-231 was labelled strongly for VVA. Huang et al., [208] demonstrated that breast cancer cell line MCF-7 labels strongly for VVA using fluorescent lectin cytochemistry. These results indicate that similar to HPA, VVA labelling detects aberrant O-GalNAc glycans, like Tn antigen, and is associated with a prognostic significance in cancer.

Table 3.1 Lectin binding specificities.

Lectin	Binding Specificity	Reference
<i>Helix pomatia</i> agglutinin (HPA)	GalNAc α 1 $\rightarrow$ 3 GalNAc β 1 $\rightarrow$ 3 Gal α 1 $\rightarrow$ 4 Gal β 1 $\rightarrow$ 4 Glc (Forssman specific pentasaccharide) > GalNAc α 1 $\rightarrow$ 3 [Fuc α 1 $\rightarrow$ 2] Gal β -(Blood group A substance) > GalNAc α 1 $\rightarrow$ 3 Gal β 1 $\rightarrow$ 3 GlcNAc > α -GalNAc (Tn antigen, GalNAc- α -Ser/Thr) > GlcNAc	[209]
<i>Vicia Villosa</i> agglutinin (VVA)	A ₄ subunit – GalNAc α 1 $\rightarrow$ 3 Gal = α -GalNAc (Tn antigen, GalNAc- α -Ser/Thr) B ₄ subunit – α -GalNAc (Tn antigen, GalNAc- α -Ser/Thr) > β -GalNAc	[198], [210]

3.2 Aims and Objectives

The experiments described in this chapter aimed to optimise, confirm and characterise HPA and VVA lectin labelling profiles of a range of model breast epithelial cell lines as a means of their validation for use in further investigations.

The objectives of this chapter are to:

- 1) Optimise and characterise HPA and VVA lectin labelling profiles of hTERT-HME1, BT-474, MCF-7, T-47D and ZR-75-1 model breast epithelial cell lines.
- 2) Investigate if the presence of differing glycosylation profiles is associated with different cell morphology.

3.3 Methods

3.3.1 Biotinylated lectin labelling

To characterise HPA lectin binding to model breast epithelial cell lines, biotinylated HPA lectin labelling was conducted as described in '*2.2 Lectin labelling*', with or without the 0.1% v/v Triton X-100 permeabilization step. This method has been adapted by members of our research group (unpublished work), based on the protocol described by Brooks and Hall [211] which was used to label fixed, paraffin wax-embedded histology sections to optimally label cultured cells.

To assess VVA lectin binding, the labelling protocol was also optimised using a range of lectin concentrations: 2.5µg/mL, 5µg/mL and 10µg/mL, all for 5-minute incubations based on the previously optimised HPA labelling protocol. Controls were employed as described in '*2.2.6 Lectin labelling controls*' and quantification was conducted as described in '*2.2.3 Biotinylated lectin cytochemistry with HRP image acquisition and quantification*'.

3.3.2 Fluorescent HPA lectin labelling

Fluorescent HPA lectin labelling was conducted and analysed using the methodology described in '*2.2.4 Fluorescent lectin labelling*'. Controls included a no lectin control and a competitive inhibition control, see Supplementary Figure 3.

3.3.3 Lectin flow cytometry

HPA lectin flow cytometry and controls were conducted as described in '*2.3 Lectin flow cytometry*' and '*2.3.2 HPA flow cytometry controls*'.

3.3.4 Cell morphology analysis

Morphological analysis of cells that had been subjected to lectin cytochemistry was conducted as described in '2.4 Cell morphology analysis'.

3.4 Results

3.4.1 Biotinylated HPA lectin cytochemistry

Figure 3.1 and 3.2B demonstrate clear visual differences in HPA labelling between the model breast epithelial cell lines, hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D. hTERT-HME1 cells showed little to no HPA labelling and if they did label, they labelled in the perinuclear region. BT-474 demonstrated very minimal HPA labelling and labelling was concentrated at the 'leading edge' of the growing cell colony. MCF-7, T-47D and ZR-75-1 demonstrated very strong HPA labelling across most cells. Progressing from the non-malignant epithelial cells (hTERT-HME1) through to the primary, non-metastatic breast cancer cells (BT-474) and then finally to the breast cancer cells derived from metastatic sources (T-47D, ZR-75-1 and MCF-7) there was a visual increase in HPA labelling, see Figure 3.2B.

A lack of labelling in the no lectin control demonstrated that there was no non-specific binding due to either endogenous biotin or endogenous peroxidase. No labelling in the DAB control confirmed effective blocking of endogenous peroxidases. The sugar specificity control resulted in a marked diminution, if not complete abolition, of lectin labelling through competitive inhibition, thus confirming sugar specificity of lectin binding to cells.

Each biological replicate value for percentage HPA labelling consisted of the average of three coverslips and three biological replicates per cell line were performed. hTERT-HME1 cells demonstrated an average percentage HPA lectin labelling of 5.0% (SD 4.7%), BT-474 21.4% (SD 8.5%), ZR-75-1 90.3% (SD 2.8%), MCF-7 92.8% (SD 2.8%), T-47D 95.5% (SD 3.4%), see Figure 3.2A. All breast cancer epithelial cell lines stably synthesised statistically significantly more O-GalNAc glycans than the normal breast epithelial cells, hTERT-HME1, as detected by the HPA lectin (BT-474 $p < 0.05$, MCF-7 $p < 0.0001$, ZR-75-1 $p < 0.0001$, T-47D $p < 0.0001$). The primary breast cancer cell line BT-474 also synthesised significantly less O-GalNAc glycans than the other metastatic breast cancer cell lines, MCF-7 ($p < 0.001$), T-47D ($p < 0.001$) and ZR-75-1 ($p < 0.001$). There was no significant difference in levels of O-GalNAc glycans between the breast cancer cell lines derived from metastases (MCF-7, T-47D and ZR-75-1).

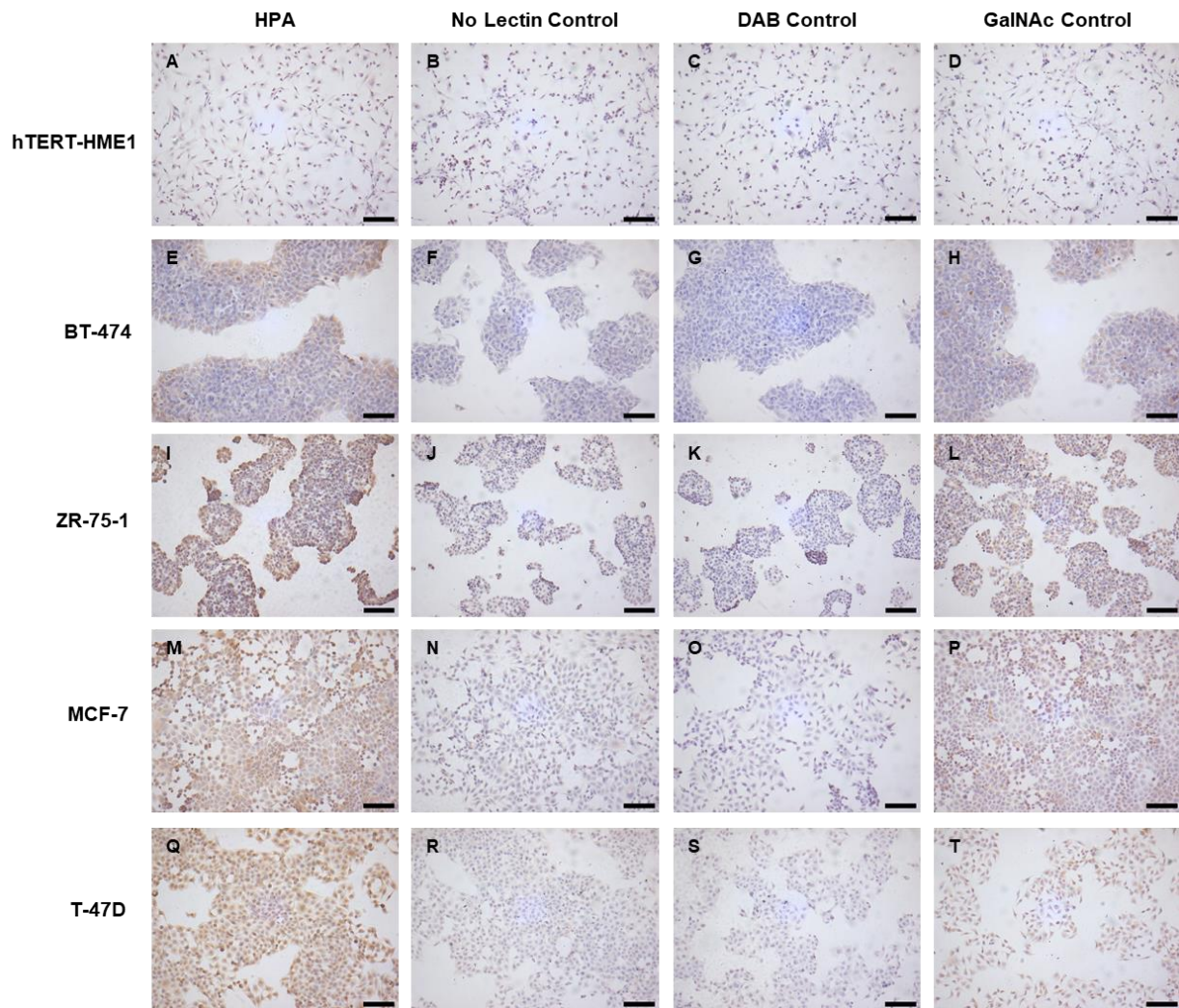


Figure 3.1 *Helix pomatia* agglutinin (HPA) lectin cytochemistry. hTERT-HME1 (A-D), BT-474 (E-H), ZR-75-1 (I-L), MCF-7 (M-P), T-47D (Q-T) cells incubated with 10µg/mL HPA for 5 minutes. Images are representatives of three biological replicates, each replicate consisting of three technical replicates per cell line. HPA lectin was omitted in the no lectin control. Both HPA and avidin peroxidase were omitted in the DAB control. Finally, a N-acetylgalactosamine (GalNAc) control was conducted as a specificity control. Scale bar = 10µm.

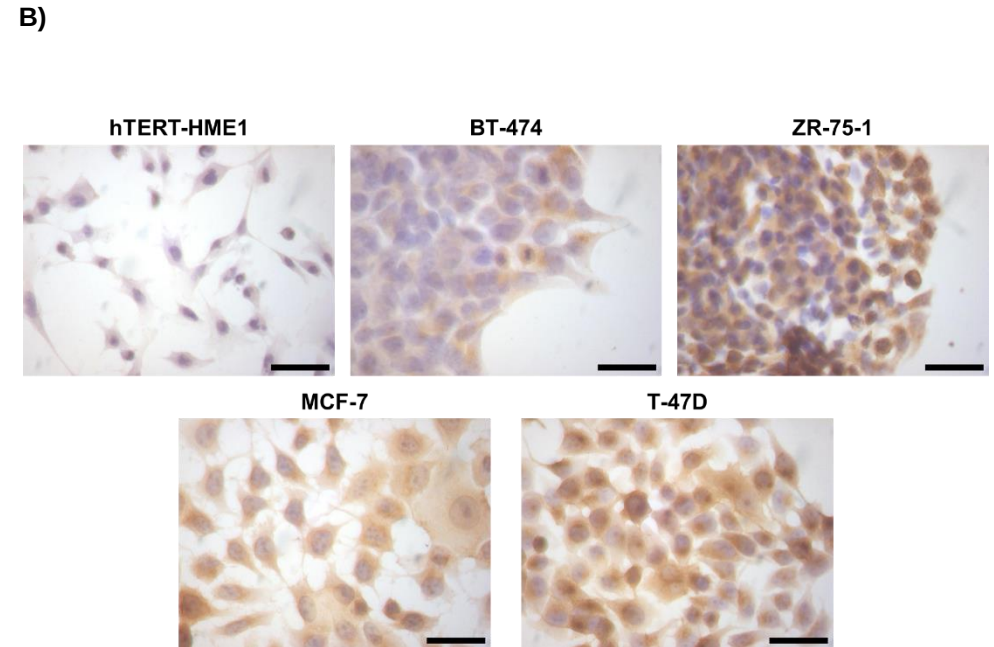
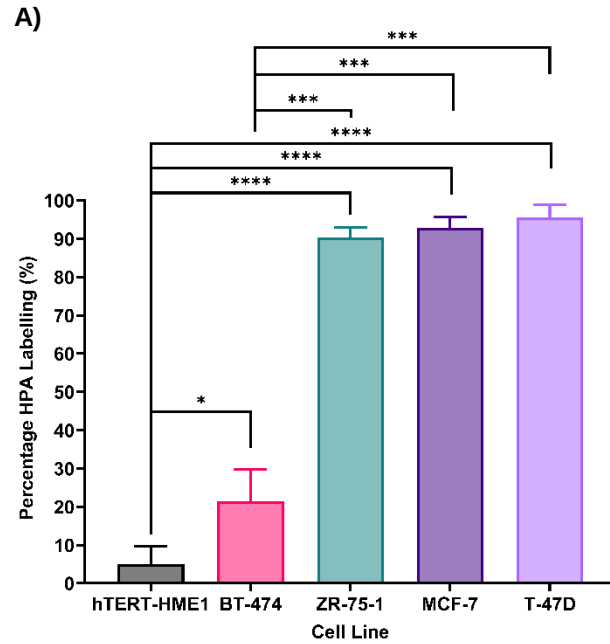


Figure 3.2 *Helix pomatia* agglutinin (HPA) lectin cytochemistry quantification. A) hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D cell lines were investigated. Each biological replicate value for percentage HPA labelling consisted of the average of three coverslips and three biological replicates per cell line were performed. hTERT-HME1 5.0% (SD 4.7%), BT-474 21.4% (SD 8.5%), ZR-75-1 90.3% (SD 2.8%), MCF-7 92.8% (SD 2.8%), T-47D 95.5% (SD 3.4%). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. B) Representative images of hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D HPA lectin cytochemistry. Scale bar = $2\mu\text{m}$.

3.4.2 Fluorescent HPA lectin labelling

Normal breast epithelial cell line, hTERT-HME1, demonstrated significantly less O-GalNAc glycans as detected by HPA labelling compared to all breast cancer cell lines (MCF-7 $p<0.01$, BT-474 $p<0.001$, T-47D $p<0.05$, ZR-75-1 $p<0.001$), see Figure 3.3 and 3.4A. The primary breast cancer-derived cell line, BT-474, displayed significantly less labelling than the metastatically derived T-47D ($p<0.05$) and ZR-75-1 ($p<0.01$) but, not the MCF-7 cell line. The MCF-7 cell line also displayed significantly less HPA labelling than the T-47D ($p<0.05$) and ZR-75-1 ($p<0.01$) cell lines. All cell lines, apart from hTERT-HME1, demonstrated significantly more labelling than their corresponding no lectin control and significantly more labelling than their corresponding GalNAc control, Figure 3.4B.

MCF-7 cells demonstrated two main labelling patterns: cells with diffuse punctate labelling across the cell surface or cells with highly clustered perinuclear labelling. T-47D cells that are labelled for HPA display high-density diffuse punctate labelling across the cell surface, often concentrated at the cell edge. BT-474 and ZR-75-1 cells that label for HPA show quite evenly diffuse punctate labelling across the cell surface, see Figure 3.3C.

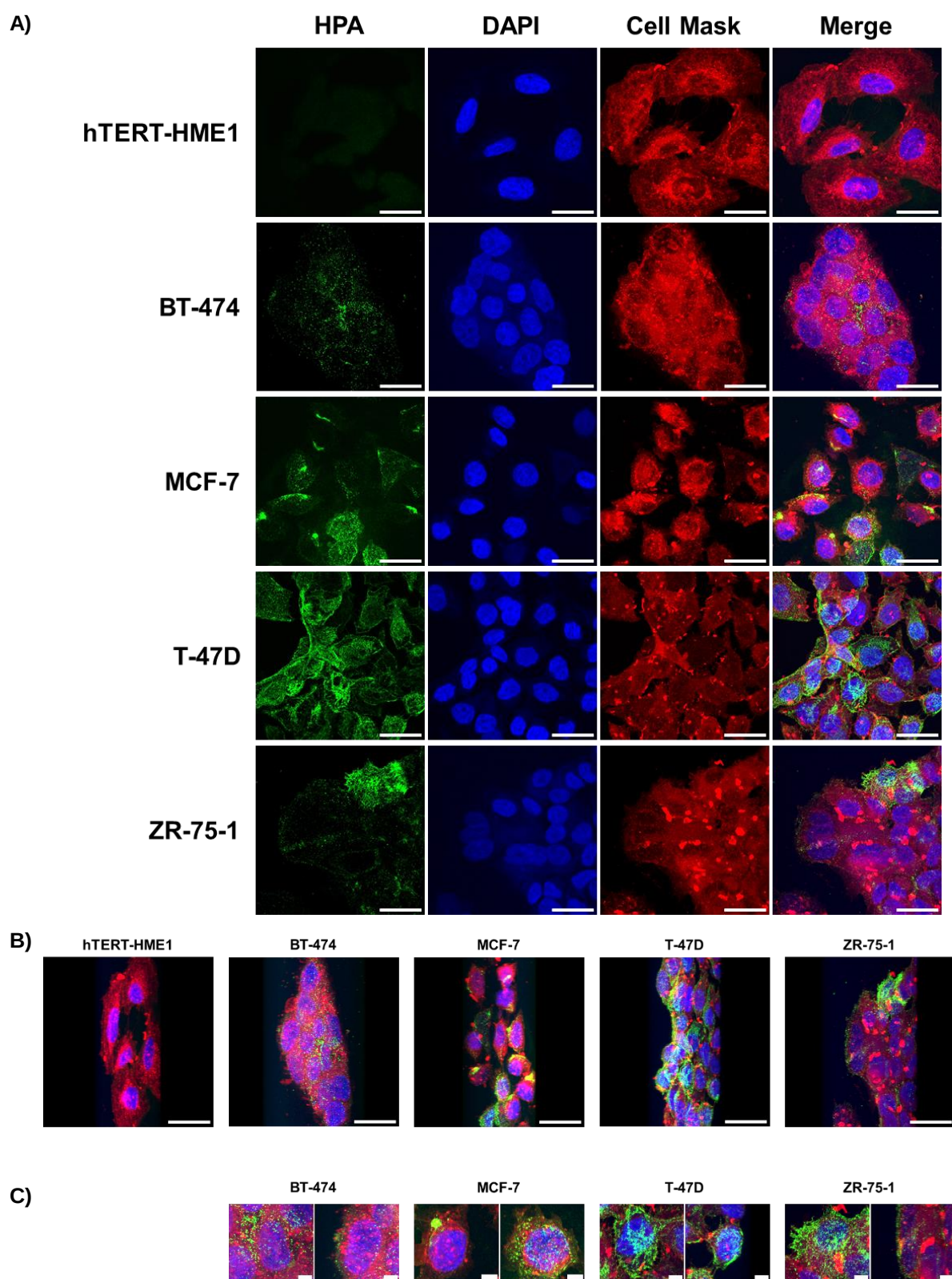


Figure 3.3 Fluorescent *Helix pomatia* agglutinin (HPA) lectin labelling. A) Representative Max Intensity Z-stack projections of model breast epithelial cell lines. HPA (green), CellMask™ (red) and DAPI (blue). Scale bar = 25µm. B) Representative 3D projects of model breast epithelial cell lines. Scale bar = 25µm. C) Region of interest from 3D project to demonstrate the pattern of HPA labelling in model breast epithelial cell lines. Scale bar = 5µm.

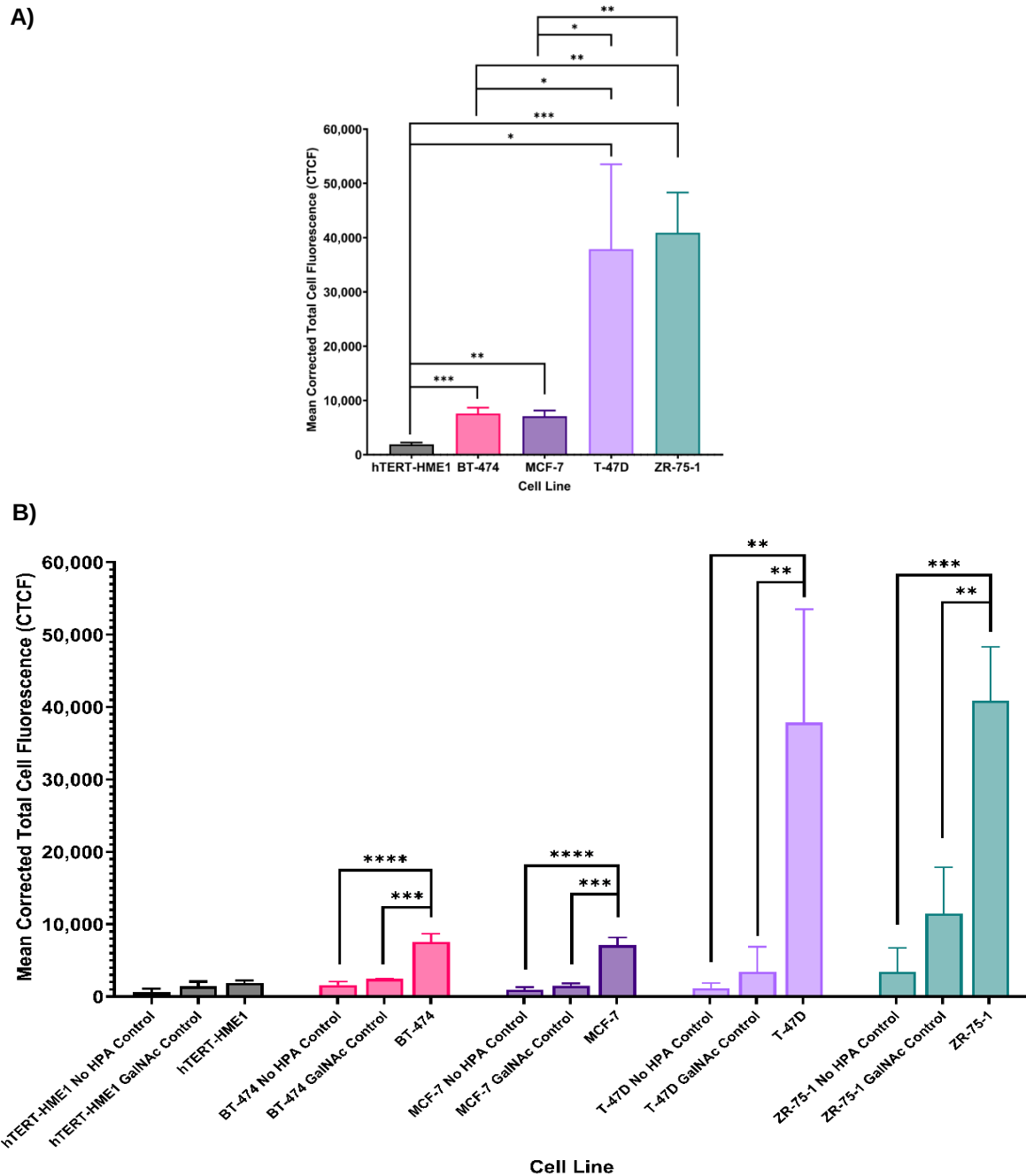


Figure 3.4 Mean corrected total cell fluorescence (CTCF) of *Helix pomatia* agglutinin (HPA) labelling. A) Mean CTCF of fluorescent HPA labelling in hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D cell lines, three biological replicates ($\pm$ SD). Each biological replicate value consists of an average of three coverslips, for which 5 images were acquired per coverslip. Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. B) Mean CTCF of fluorescent HPA labelling of hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D cell lines, three biological replicates ($\pm$ SD). HPA lectin was omitted in the no lectin control to assess levels of autofluorescence. An N-acetylgalactosamine (GaINAc) control was conducted as a specificity control. Each biological replicate value consists of an average of three coverslips, for which 5 images were acquired per coverslip. An ANOVA with Tukey's multiple comparisons test was used to test for significance within each cell line condition, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

3.4.3 VVA lectin cytochemistry

The VVA lectin was optimised over a range of concentrations (2.5µg/mL, 5µg/mL and 10µg/mL) at a 5-minute incubation period, see Supplementary Figure 4. A concentration-dependent increase in VVA labelling was observed, with positive labelling at the lowest concentration of 2.5µg/mL and 10µg/mL demonstrating the strongest labelling patterns. The chosen optimised concentration was 5µg/mL at a 5-minute incubation period.

Figure 3.5 and 3.6B demonstrate the clear visual differences in VVA labelling between the model breast epithelial cell lines, hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D. The cell lines demonstrated similar labelling patterns for both HPA and VVA lectins. hTERT-HME1 demonstrated little to no VVA labelling and if they did label, they labelled in the perinuclear region. Both BT-474 and ZR-75-1 cells demonstrated VVA labelling that is concentrated on the edge of the growing cell colony. MCF-7, T-47D and ZR-75-1 demonstrated strong VVA labelling across most cells. The same three controls were utilised in the VVA lectin cytochemistry as in the HPA lectin cytochemistry. These controls resulted in the abolition of the majority of labelling and therefore confirmed that there was competitive inhibition of VVA binding through the addition of GalNAc monosaccharides and specific binding, see Figure 3.5.

hTERT-HME1 cells demonstrated an average percentage VVA lectin labelling of 1.7% (SD 0.7%), BT-474 65.6% (SD 8.7%), MCF-7 90.5% (SD 1.5%), ZR-75-1 94.3% (SD 1.3%), T-47D 96.1% (SD 0.9%), see Figure 3.6A. Consistent with the results of HPA lectin cytochemistry, all the breast cancer epithelial cell lines stably synthesised statistically significantly more O-GalNAc glycans than the normal breast epithelial cells, hTERT-HME1, as detected by the VVA lectin (BT-474 $p<0.001$, MCF-7 $p<0.0001$, ZR-751 $p<0.0001$, T-47D $p<0.0001$).

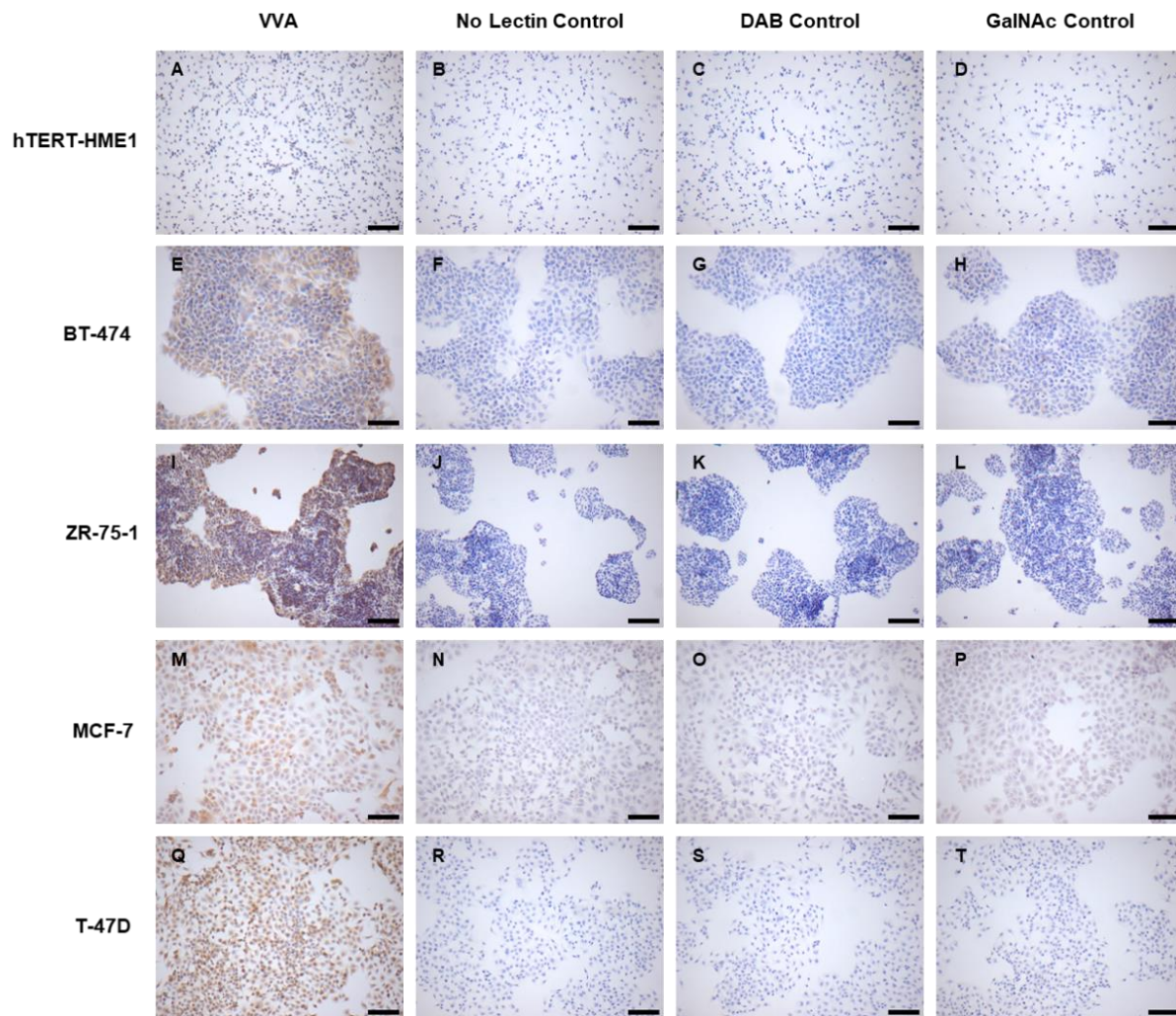
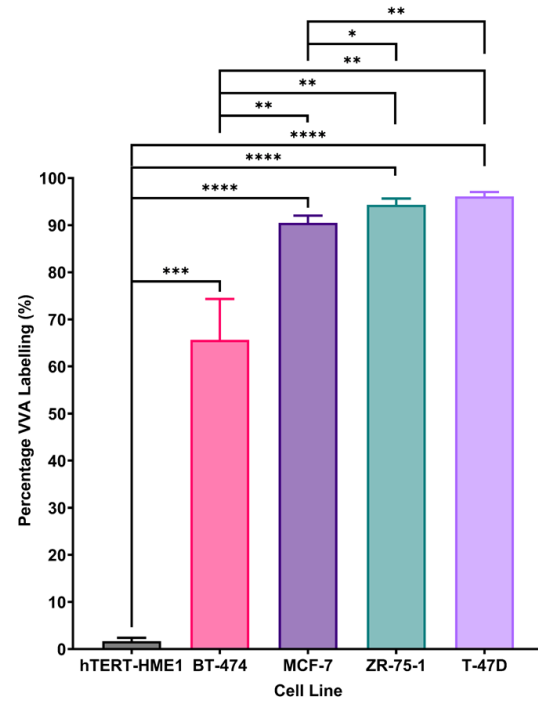


Figure 3.5 Optimised *Vicia villosa* agglutinin (VVA) lectin cytochemistry of 5µg/mL VVA for 5 minutes. hTERT-HME1 (A-D), BT-474 (E-H), ZR-75-1 (I-L), MCF-7 (M-P), T-47D (Q-T) cells incubated with 5µg/mL VVA for 5 minutes. Images are representatives of three biological replicates, each replicate consisting of three technical replicates per cell line. VVA lectin was omitted in the no lectin control. Both VVA and avidin peroxidase were omitted in the DAB control. Finally, a N-acetylgalactosamine (GalNAc) control was conducted as a specificity control. Scale bar = 10µm.

A)



B)

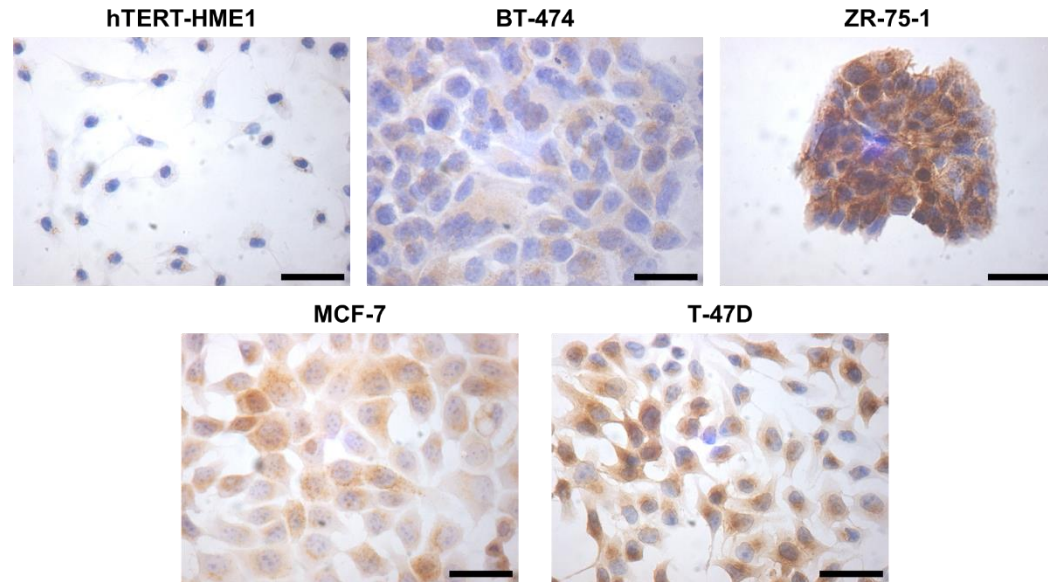


Figure 3.6 Optimised *Vicia villosa* agglutinin (VVA) lectin cytochemistry quantification. A) Average percentage of VVA lectin labelling of hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D cell lines, three biological replicates ($\pm$ SD). Each biological replicate value for percentage VVA labelling consisted of the average of three coverslips. hTERT-HME1 1.7% (SD 0.7%), BT-474 65.6% (SD 8.7%), MCF-7 90.5% (SD 1.5%), ZR-75-1 94.3% (SD 1.3%), T-47D 96.1% (SD 0.9%). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. B) Representative images of hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D VVA lectin cytochemistry. Scale bar = $2\mu\text{m}$.

3.4.4 Cell morphology investigations

When assessing the images of the lectin cytochemistry it was strikingly apparent that there were obvious observable differences in the morphology of the lectin-positive and lectin-negative binding populations within individual cell lines, see Figure 3.7 for examples of HPA and VVA-positive and -negative cells. Across all breast epithelial cell lines, the average aspect ratios of both the HPA and VVA-negative cells are significantly more elongated than HPA and VVA-positive cells, see Figure 3.8A and 3.9A.

MCF-7 (x:y ratio 2.1:1), T-47D (x:y ratio 2.0:1), hTERT-HME1 (x:y ratio 2.2:1) and BT-474 (x:y ratio 2.1:1) HPA-negative cells all breached >2:1 aspect ratio. ZR-75-1 HPA-negative cells were close at an x:y ratio 1.9:1. MCF-7 (x:y ratio 2.2:1), T-47D (x:y ratio 2.0:1), hTERT-HME1 (x:y ratio 2.3:1) and BT-474 (x:y ratio 2.2:1) VVA-negative cells all breached the >2:1 aspect ratio. Again, ZR-75-1 VVA-negative cells were close at an x:y ratio of 1.9:1. Across all the breast epithelial cell lines both HPA and VVA-positive cells were significantly more circular than HPA-negative cells, see Figure 3.8B and 3.9B. Both HPA and VVA-positive cells featured a more rounded morphology, suggestive of an epithelial-like phenotype. HPA and VVA-negative cells demonstrated an elongated morphology, suggestive of a mesenchymal-like phenotype.

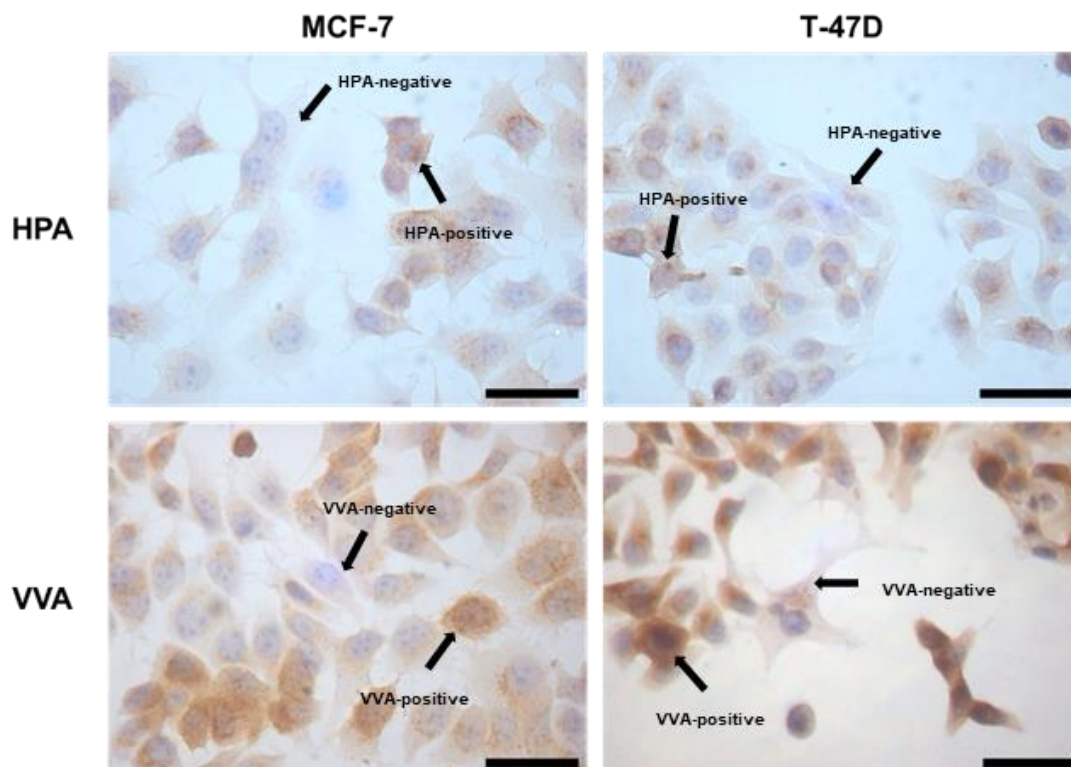


Figure 3.7 Morphology examples of MCF-7 and T-47D lectin positive and negative populations. *Helix pomatia* agglutinin (HPA) and *Vicia villosa* agglutinin (VVA) labelled cells demonstrate a different morphology compared to weakly or unlabelled cells. Scale bar = 2µm.

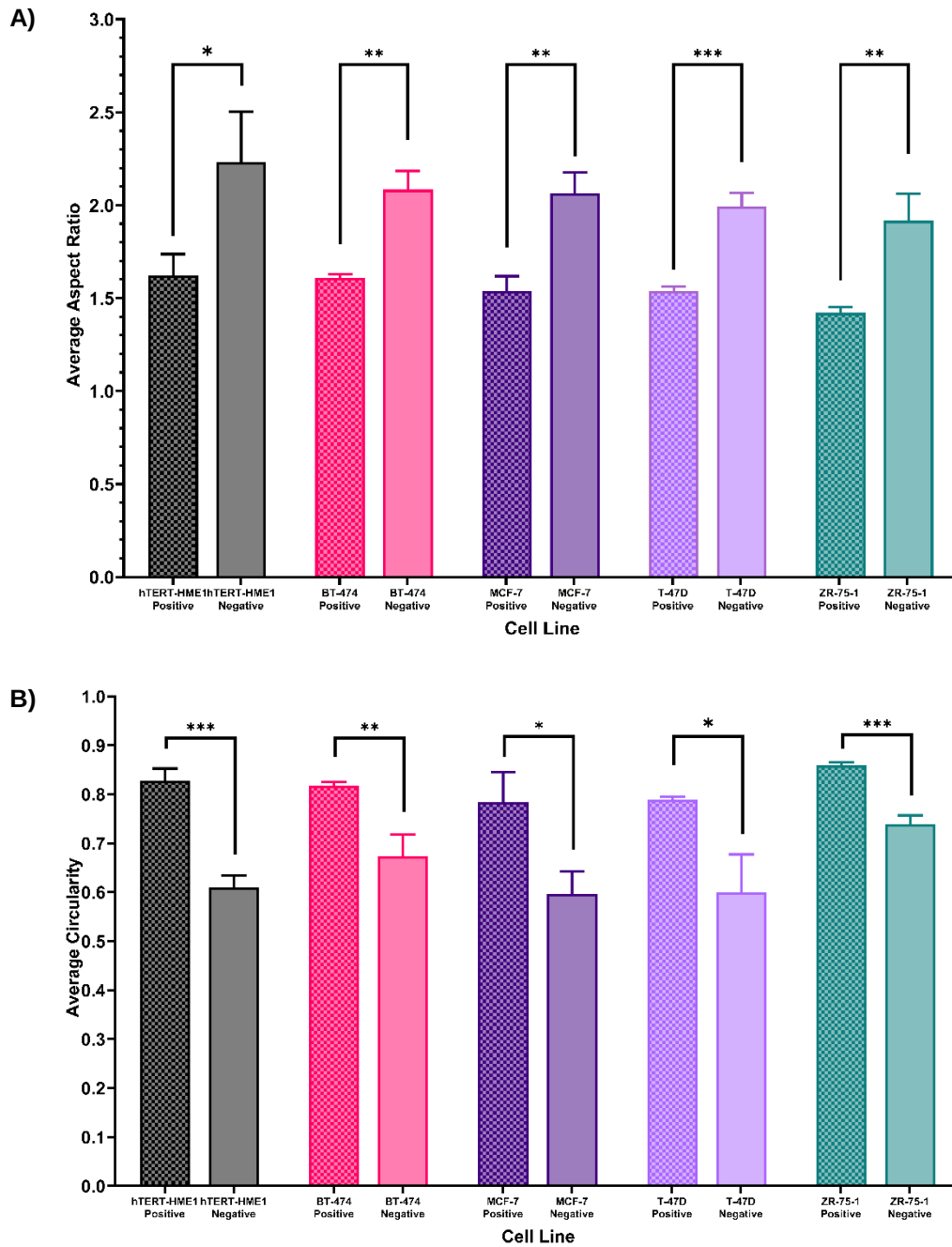


Figure 3.8 *Helix pomatia* agglutinin (HPA) morphology assessments. Each biological replicate value consists of an average of three coverslips and nine images are used per coverslip, three biological replicates per cell line were performed. Unpaired Students t-test was used to test for significance between cell line conditions, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. A) Average aspect ratio of HPA-positive cells and HPA-negative cells across hTERT-HME1, BT-474, MCF-7, T-47D and ZR-75-1 cell lines ($\pm$ SD). B) Average circularity of HPA-positive cells and HPA-negative cells across hTERT-HME1, BT-474, MCF-7, T-47D and ZR-75-1 cell lines ($\pm$ SD).

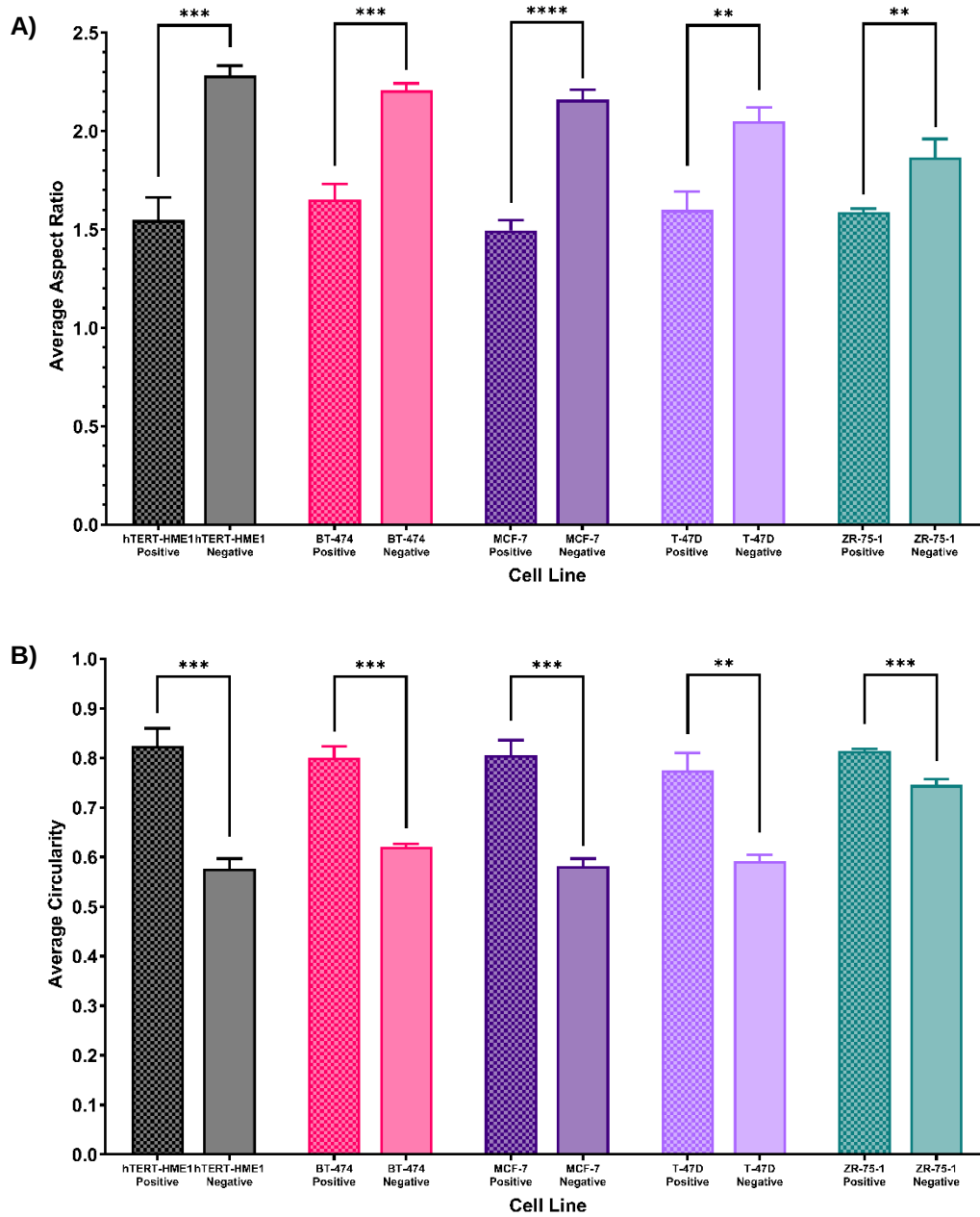


Figure 3.9 *Vicia villosa* agglutinin (VVA) morphology assessments. Each biological replicate value consists of an average of three coverslips and nine images are used per coverslip, three biological replicates per cell line were performed. Unpaired Students t-test was used to test for significance between cell line conditions, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. A) Average aspect ratio of VVA-positive cells and VVA-negative cells across hTERT-HME1, BT-474, MCF-7, T-47D and ZR-75-1 cell lines ($\pm$ SD). B) Average circularity of VVA-positive cells and VVA-negative cells across hTERT-HME1, BT-474, MCF-7, T-47D and ZR-75-1 cell lines ($\pm$ SD).

3.5 Discussion

3.5.1 Breast cancer epithelial cell lines synthesize significantly more O-GalNAc glycans than normal breast epithelial cell lines

Model breast epithelial cell lines, hTERT-HME1, BT-474, MCF-7, T-47D and ZR-75-1 were assessed for their HPA and VVA binding using biotinylated lectin cytochemistry, fluorescent lectin cytochemistry and flow cytometry. The HPA and VVA lectins were employed to characterise the model epithelial cell lines for their levels of truncated O-GalNAc glycans due to their binding specificities. HPA demonstrates a range of binding specificity including the Forssman antigen (GalNAc- α 1,3 GalNAc), blood group A substance (GalNAc- α 1,3 [Fuc- α 1,2] Gal- β), Tn antigen (GalNAc- α -Ser/Thr) and GlcNAc ([212] and reviewed by [104]), as described in Table 3.1. VVA is reported to have a more limited range of binding specificity, notably the B₄ subunit, which binds to Tn antigen (GalNAc- α -Ser/Thr) and β -GalNAc. The VVA used in this thesis is primarily comprised of B₄ subunits, and so labelling is more likely to be representative of binding to Tn antigen structures [199, 213, 214]. Whereas, the HPA lectin labelling is probably representative of a range of O-GalNAc structures, which includes the Tn antigen but, has still been linked to poor prognosis and aggressive disease, as discussed previously in '3.1.2 *Helix pomatia* agglutinin (HPA) binding as a prognostic marker in cancer'.

Biotinylated HPA was initially employed in lectin cytochemistry in this thesis, which utilised a classical chromogenic method via an enzyme peroxidase reporter-based system. Leathem and Brooks [182] first demonstrated the relationship between HPA binding and poor long-term prognosis and subsequently HPA-binding association with the presence of lymph node metastasis in 373 primary breast cancers [186]. Analysis of this method of lectin cytochemistry is laborious and relatively subjective however, it does mirror what is conducted in a clinical practice. For example, pathologists score breast histopathology specimens labelled for hormone receptor proteins [215]. Like in clinical practice, lectin scoring techniques in the literature often use two independent observers to score the cells [194-196].

Biotinylated HPA lectin labelling was assessed at a light microscope level, by eye, and HPA-positive cells were identified as any cell that displayed brown cytoplasmic labelling that was of intensity equal to or greater than ++ (i.e. all those deemed ++, +++ or ++++). Those deemed to be + or less were considered to be HPA-negative. The percentage of HPA-positive cells was calculated by dividing by the total number of cells in the image. The imposed binary scoring of 'positive' (cells scored $\geq$ ++) or 'negative' (cells scored $\leq$ +) labelling was utilized as the literature has previously described that simply the presence of HPA labelling alone was sufficient for the prognostic implications, as discussed previously in '3.1.2 *Helix pomatia* agglutinin (HPA) binding as a prognostic marker in cancer' [182, 184-186, 196]. Results showed that the normal breast epithelial cell line hTERT-HME1 demonstrates

an average percentage of HPA lectin labelling of 5.0% (SD 4.7%). The breast cancer epithelial cell lines demonstrated an average percentage labelling of 21.4% (SD 8.5%) for BT-474, 90.3% (SD 2.8%) for ZR-75-1, 92.8% (SD 2.8%) for MCF-7 and 95.5% (SD 3.4%) for T-47D. All the breast cancer epithelial cell lines synthesised significantly more O-GalNAc glycans than the normal breast epithelial cells, hTERT-HME1, as detected by the HPA lectin (BT-474 $p < 0.05$, MCF-7 $p < 0.0001$, ZR-751 $p < 0.0001$, T-47D $p < 0.0001$). The primary breast cancer cell line BT-474 also synthesised significantly less O-GalNAc glycans than the other metastatic breast cancer cell lines, MCF-7 ($p < 0.001$), T-47D ($p < 0.001$) and ZR-75-1 ($p < 0.001$), indicating an increase in O-GalNAc glycans with progressive disease stage cell line derivation.

These results are consistent with those reported for the same cell lines previously in the literature by Brooks et al., [164], Schumacher et al., [165], and unpublished work from our laboratory group. MCF-7 have been described as having intense labelling (++++), BT-474 as weakly labelled (+) and ZR-75-1 as strongly labelling (+++) [164]. T-47D cells have also been described in the literature as being intensely labelled with HPA [165]. The epithelial breast cancer cell lines also demonstrated comparable labelling patterns to the clinical samples, whereby there is granular cytoplasmic labelling with mainly cell surface localisation [182, 186, 216]. The hTERT-HME1 normal breast epithelial cell line HPA labelling characteristics have not been described in the literature before however, another normal breast epithelial cell line, HMT 3522, has been shown to exhibit negligible HPA binding which is consistent with these results [164]. Normal breast tissue has been shown to bind HPA however labelling is highly localised to the luminal, apical surface of the epithelial cells [216].

Utilizing the same methodology, the VVA lectin was optimised for labelling the cultured cells. The chosen optimised concentration allowed for optimal labelling of all the breast epithelial cell lines that allowed for clear identification of VVA-positive and -negative cells. Lectin cytochemistry was then repeated using biotinylated VVA to investigate if using a lectin with a more distinct nominal specificity for terminal α -GalNAc monosaccharide residues resulted in a difference in labelling patterns compared to the more diverse binding profile of the HPA lectin. hTERT-HME1 cells demonstrated an average percentage VVA lectin labelling of 1.7% (SD 0.7%), BT-474 65.6% (SD 8.7%), MCF-7 90.5% (SD 1.5%), ZR-75-1 94.3% (SD 1.3%), T-47D 96.1% (SD 0.9%). Similarly, to HPA lectin cytochemistry, all the breast cancer epithelial cell lines displayed statistically significantly more labelling for O-GalNAc glycans than the normal breast epithelial cells, hTERT-HME1, as detected by the VVA lectin (BT-474 $p < 0.001$, MCF-7 $p < 0.0001$, ZR-751 $p < 0.0001$, T-47D $p < 0.0001$). The results in cell lines here demonstrated similar labelling profiles to those of clinical samples representative of their origins [206, 207] and results of the same cell line, MCF-7, reported previously in the literature [208].

VVA and HPA biotinylated lectin cytochemistry results were generally comparable, except for BT-474 which displayed larger differences. BT-474 cells are known in the literature to grow in clusters that form multilayered colonies [167, 217], and lectin labelling has been optimized for cells that grow in monolayers. These multilayered colonies could aggregate and physically trap reagents, and make

effective washes between steps technically challenging, causing artefactual labelling and could explain the variations seen. It also made scoring by eye more difficult and resulted in bias to scoring areas of images that were more accessible to the scoring approach, which incidentally was the edge of the colonies, which often demonstrated more labelling. Iskratsch et al., [214] have demonstrated the binding profiles of HPA and VVA using dot blots with varying glycoproteins and monosaccharide samples to screen for lectin binding. The VVA lectin used in this thesis has a predominance of B₄ subunits but still may contain A₄ subunits which, like HPA, has also been shown to be able to bind to the T antigen, demonstrating that the lectins show broad specificity for similar structures [198, 210, 218].

Across the permeabilization conditions results were also comparable between the breast epithelial cell lines, see Supplementary Figure 5-8. Removing the permeabilization step made no significant difference to labelling scores or visual labelling intensities, indicating the majority of terminal O-GalNAc glycoproteins were predominantly cell-surface in breast cancer epithelial cell lines, see Supplementary Figure 9. This observation is consistent with those made by Leathem and Brooks [182] and Brooks and Leathem [186] who observed that HPA binding demonstrated marked cell surface localisation in clinical samples.

Fluorescently labelled HPA in flow cytometry, provided an additional automated method with a less subjective quantification. Results confirmed the levels of cell-surface O-GalNAc glycans, demonstrating comparable results to the other methods employed in this study. See Supplementary Figure 10 for representative images of the gating strategy and the resulting histograms and Supplementary Figure 11 for results. Normal breast epithelial cell line, hTERT-HME1, demonstrated a significantly lower mean Allophycocyanin-Area (APC-A) value, indicating less HPA labelling and therefore the presence of O-GalNAc glycans, than all the breast cancer cell lines (BT-474 $p < 0.01$, MCF-7 $p < 0.01$, T-47D $p < 0.0001$ and ZR-75-1 $p < 0.01$). These results are consistent with flow cytometry results previously reported for BT-474, MCF-7 and ZR-75-1 cell lines in the literature [164].

Finally, fluorescently labelled HPA was employed using confocal microscopy. In contrast to light microscopy, which illuminates samples uniformly, confocal microscopy uses a laser for excitation light in a defined area at a defined depth. The use of a pinhole also allows for the rejection of out-of-focus light, resulting in higher-resolution imaging [219]. This allowed for higher-resolution imaging of the localisation of labelling and revealed cell surface and highly punctate labelling. This is clear in comparison to the more diffuse, amplified labelling achieved from the chromogenic methods. Results were consistent with all other methods previously described, hTERT-HME1 demonstrated significantly less O-GalNAc glycans, detected by HPA labelling, compared to all breast cancer cell lines (MCF-7 $p < 0.01$, BT-474 $p < 0.001$, T-47D $p < 0.05$, ZR-75-1 $p < 0.001$). hTERT-HME1 cells did not demonstrate significantly more labelling than the no lectin or GalNAc specificity control, this could be due to the lack of O-GalNAc present on the cell line, to begin with.

Across all methodologies employed, normal breast epithelial cell line hTERT-HME1 was demonstrated to synthesise significantly less truncated O-GalNAc glycans than the breast cancer

epithelial cell lines regardless of disease stage derivation, as detected by HPA and VVA. The consistency in these results, between the different approaches employed in this study and with the literature, therefore warrants a strong argument for the utility of employing these cell lines as model breast epithelial cell lines that display consistent and heterogeneous glycosylation profiles. Acknowledging the usual caveats of using cell lines as a model system, they could provide a strong tool for investigating the effects of altered glycosylation, particularly for truncated O-glycosylation associated with cancer.

3.5.2 Truncated O-GalNAc glycans are associated with cell morphology

When observing the model breast epithelial cell lines using microscopy, it was clear that there were observable differences in the morphology of lectin-positive and lectin-negative binding populations. The result that a mixed population of both lectin-positive and lectin-negative cells exists within a cell line is consistent with observations of heterogeneous binding populations that exist in clinical samples. Subpopulations of HPA-positive and -negative cells in single primary tumours were described in the first studies by Leathem and Brooks [182] and Brooks and Leathem [186] investigating HPA labelling in clinical samples. Both Schumacher et al., [220] and Brooks and Leathem [221] have demonstrated that metastases originating from HPA-positive breast cancers exhibit both populations of HPA-binding and HPA-nonbinding cells. Metastases taken at post-mortem from originally HPA-positive primary breast cancers demonstrated that 79% of metastases were HPA-positive and 21% were negative, demonstrating the existence of HPA-negative clonal cell populations. Similarly, Bresalier et al., [222] demonstrated that the metastases of colon cancers had reduced VVA labelling compared to the primary tumours they arose from and acknowledged that intra-tumoral heterogeneity of VVA labelling exists within these populations, assessed by lectin cytochemistry. Yamashita et al., [204] demonstrated that VVA labelling was reduced in tumour cells in lymph node metastases of VVA-positive gastric tumours, indicating the existence of a mixed population of VVA-positive and -negative cells. Model cell lines may therefore be of utility when trying to understand the dynamics of glycosylation, and are reflective of the mixed populations of lectin-positive and lectin-negative cells that exist in clinical tumours and their metastases.

Changes in cell shape are characteristic of EMT, cells change from an epithelial-rounded morphology to a mesenchymal-elongated morphology [223]. A simple, and commonly used method to quantify this is through taking measurements of aspect ratio (length/width) which indicates elongation of a cell and circularity ($4\pi \cdot (\text{area}) / (\text{perimeter})^2$) using ImageJ [171]. Across all the epithelial cell lines studied here, both HPA and VVA-positive cells were significantly more circular than HPA and VVA-negative cells. The average aspect ratio of HPA and VVA-negative cells was significantly more elongated than HPA and VVA-positive cells from the cell lines, MCF-7, T-47D, hTERT-HME1 and BT-474, all breaching the >2:1 aspect ratio which indicates an elongated morphology. Both ZR-75-1 HPA and VVA-negative cells were close to the elongation value at a value of 1.9:1 for both. However, they were still

significantly more elongated than their positive binding counterparts. HPA/VVA-positive cells demonstrated a more rounded morphology, suggestive of an epithelial-like phenotype. HPA/VVA-negative cells demonstrated an elongated morphology, suggestive of a mesenchymal-like phenotype. This, intriguingly, indicates a potential association of truncated O-GalNAc glycans to cell morphology, implicated in EMT and therefore warranted further investigation.

3.6 Key findings

- Normal breast epithelial cell line hTERT-HME1 displayed significantly less HPA/VVA lectin binding, and therefore reduced terminal O-GalNAc glycans, than breast cancer epithelial cell lines BT-474, MCF-7, T-47D and ZR-75-1.
- Permeabilizing cells before HPA and VVA biotinylated lectin cytochemistry resulted in no significant difference in cell labelling, indicating the majority of terminal O-GalNAc glycoproteins detected by the HPA and VVA lectins were cell-surface in breast cancer epithelial cell lines. This localisation was also visualised in the 3D Z-stack confocal microscopy images.
- Sub-populations of HPA/VVA-positive and -negative cells exist in the cell lines which display distinct cell morphologies.
- HPA/VVA-positive cells were significantly more rounded whereas, HPA/VVA-negative cells demonstrated a significantly more elongated phenotype. These phenotypes are reminiscent of those displayed in EMT, suggesting a potential relationship between terminal O-GalNAc glycans and EMT.

Chapter 4

Epithelial-mesenchymal transition and O-GalNAc glycans

4. Epithelial-mesenchymal transition and O-GalNAc glycans

4.1 Background

Results described in Chapter 3 demonstrated that HPA-negative cells demonstrated a mesenchymal-like, elongated phenotype and HPA-positive cells demonstrated an epithelial-like, rounded phenotype. As HPA labelling in cancer has been associated with metastasis as described in '3.1.2 *Helix pomatia* agglutinin (HPA) binding as a prognostic marker in cancer' and with this evidence of phenotypical changes related to HPA labelling it seemed prudent to investigate the effect, role or potential changes in glycosylation during EMT.

4.1.1 EMT

EMT, first described by Greenburg and Hay [116] as 'epithelial-mesenchymal transformations', is a biological process whereby basement membrane-bound, polarised epithelial cells trans-differentiate into migratory, invasive mesenchymal cells. The change from 'transformations' to 'transition' was introduced to reflect the reversibility of the process as exemplified through mesenchymal-epithelial transition (MET), and to distinguish the process from neoplastic transformation, reviewed by [124]. For an overview of EMT, see '1.4 *Epithelial-mesenchymal transition (EMT)*', and see Table 4.1 for common EMT markers.

Table 4.1 Common epithelial-mesenchymal transition (EMT) markers.

Epithelial		Mesenchymal	
Marker	Description	Marker	Description
E-cadherin	Cell surface adhesion protein	N-cadherin	Cell surface adhesion protein
β -catenin (cell membrane location)	Cell membrane localised	β -catenin (cytoplasmic/nuclear location)	Cytoplasmic/nuclear location where it can act as transcription factor cofactor
Zonula occludens-1	Cell surface adhesion protein	Fibronectin	Glycoprotein of the extracellular matrix
Cytokeratins	Cytoskeletal protein	Vimentin	Intermediate filament protein

Three subtypes of EMT have been described in the literature, based on the biological context in which they were observed [124], see '1.4.1 Types of EMT' for an overview. Growth factors, cytokines and ECM proteins in the primary tumour microenvironment produced as a result of inflammation and hypoxia can induce EMT by binding to receptors on the cell surface and inducing intracellular signalling networks [224], see '1.4.3 Induction of EMT' for an overview. It is also rare for an epithelial cancer cell to fully progress in a full binary switch into a mesenchymal state, with a spectrum of pEMT states being more commonly found, which can frequently revert via MET back to an epithelial state [126], discussed in '1.4.2 Partial EMT'. These versatile, dynamic phenotypes are important for cells to successfully navigate the metastatic cascade previously discussed in '1.3.4 Metastasis'. pEMT also allows cancer cells to move via collective migration, as a result of retaining some epithelial traits whilst being 'led' by a pEMT cell, and benefit from enhanced attachment to the ECM through gaining some mesenchymal traits. Single-cell mesenchymal migration relies upon integrin-mediated interactions with the ECM and is characterised by a fibroblast-like morphology, as previously described in '1.4.4 Cell migration' [225].

4.1.2 Functional assays to investigate EMT

Alongside investigating for presence of key EMT markers at the protein or mRNA transcript level and the morphological characterisation of cells, functional assays give critical insight into the functional effect of EMT in a cell's migratory, invasion and adhesive capabilities. Commonly used functional assays that model key stages in EMT and MET include the wound healing assay (scratch assay), transwell chamber migration assay, transwell invasion assay and adhesion assays (static, rocking and under flow conditions), see Table 4.2 (described by Pantazi et al., [171]). Increased cell migration is a key functional marker of EMT whereas, increased adhesive capabilities are key for success in late stages of metastasis for example, in adhesion to vascular endothelium in hematogenous dissemination, and the reverse process of EMT, MET.

As the focus of this study was two-dimensional (2D) single-cell and collective cell migration, a wound healing assay was utilised. This assay also allowed for fluorescent labelling of specific markers on cells in the healing 'wound', which allowed for investigations of both the glycosylation status and EMT marker status of cells. Adhesive capability is also critically important during EMT and MET, notably cell adhesion to vascular endothelium during hematogenous dissemination, which ultimately allows the formation of metastatic tumour sites [226]. A static adhesion assay was therefore utilised, which mimics conditions whereby cancer cells come to arrest in small capillaries and adhere to an endothelial monolayer, to investigate the adhesive capacity of cell lines with differing O-GalNAc glycan profiles. However, due to the lack of shear flow forces some adhesive events may be missed using this assay [227], Lawrence et al., [228] documented that fluid shear-induced torque is required for maintained leukocyte rolling, a biological phenomenon that is often compared to cancer cell adhesion in blood vessels.

Table 4.2 Evaluation of some functional assays commonly used to investigate epithelial-mesenchymal transition (EMT). Adapted from [171].

Functional Assays to Investigate Epithelial-Mesenchymal Transition (EMT)				
Assay Type	Description	Advantages	Limitations	References
Wound healing assay	2D space filling collective and single migration, 'wound' can be inflicted using a sterile pipette tip	Inexpensive, no specialised equipment required, easy setup, standardised technique with easy analysis	Reproducibility issues ('wound' thickness, folded cells at 'wound' edge which can reattach), optimisation is required, can cause damage to cells and >24hrs migration cannot be distinguished from cell proliferation	[229], [230]
Transwell chamber migration assay	2D single-cell migration along vertical chemotaxis	Commercially available, relatively high throughput	Relatively high-cost, end-point-only assay, quantification can affect results	[231], [230]
Transwell invasion assay	3D single-cell migration along vertical chemotaxis	Commercially available, relatively high throughput	High-cost, distance of invasion not considered	[232], [230]
Static adhesion assay	Adhesion of cells onto an endothelial monolayer, models arrest and adhesion of cells in a capillary bed	Relatively inexpensive, no specialised equipment required, easy setup, multiple replicates possible	Does not represent <i>in vivo</i> shear forces	[233]
Rocking adhesion assay	Adhesion of cells onto an endothelial monolayer under sweeping flow conditions	Some <i>in vivo</i> physiological characteristics mimicked and multiple replicates possible	Does not represent <i>in vivo</i> shear forces, increased technical complexity, time-consuming, requires specific equipment	[234]

Table 4.2 (Continued).

Flow adhesion assay	Adhesion of cells onto an endothelial monolayer under shear flow conditions	Representative of <i>in vivo</i> shear forces	Requires specialist equipment, technically complex, time-consuming, high cost	[235], [236]
---------------------	---	---	---	--------------

4.1.3 Glycosylation and EMT

Several types of glycosylation have been documented as being associated with EMT, including O-GlcNAcylation and N-linked glycosylation events however, the mechanisms linking EMT and glycosylation have not been fully elucidated (reviewed by [237]). The relationship could be due to alterations in protein trafficking, binding properties, activity, stability or protein half-life as a result of altered glycosylation. For example, Snail-induced EMT resulted in differential N- and O-linked glycosylation of CRB3, a transmembrane protein required for apical membrane and tight junctions, which resulted in a reduction in its membrane half-life and therefore increased the depolarisation of the cells [238]. Glycosylation has been also previously linked to controlling the adhesion and functionality of cadherins, either through control of trafficking via O-GlcNAcylation, resulting in retention in the ER, or through dictating the spatial cadherin arrangement at the cell membrane therefore impacting the adhesive capability of cells and ultimately migratory capabilities [239].

Alterations in mucin-type O-linked glycosylation have also been associated with regulating EMT, key glycans implicated in EMT are the Tn antigen, and sialylated glycans like sialyl-Tn and sialyl-core 1 (reviewed by [54] and [240]). For example, the stability and subsequent degradation of fibronectin have been demonstrated to be associated with GALNT6. O-GalNAc glycosylation of fibronectin, identified by VVA lectin binding, resulted in increased protein stabilisation in breast cell lines [241]. Reduced O-GalNAc glycosylation of TGF- β type I and II receptors, by GALNT4 knockout, has been demonstrated to result in increased TGF- β receptor dimerization and activity, which would increase signal transduction of a key EMT signalling pathway [242]. Site-specific changes to O-linked glycosylation could result in a highly dynamic and specific means of regulating protein function and proteolysis (reviewed by [40]). For example, site-specific O-GalNAc glycosylation of low-density lipoprotein receptors catalysed by GALNT11 was shown to support low-density lipoprotein ligand binding [243]. Together these findings demonstrate that further exploration and understanding are required and suggest that there are layers of complexity to the regulation of EMT by mucin-type O-linked glycosylation and the GALNTs, that could vary throughout each step of the metastatic cascade.

4.2 Aims and Objectives

Breast epithelial cell lines hTERT-HME1, MCF-7 and T-47D were used as models to investigate mucin-type O-linked glycosylation in EMT. The experiments described in this chapter aimed to investigate how O-GalNAc glycans are implicated in the migratory and adhesive capabilities of cells, which are both key functional components for the success of the metastatic cascade. It also sought to investigate if key EMT markers, E-cadherin, N-cadherin and vimentin are associated with the glycosylation status of cells. The objectives of this chapter are to:

- 1) Characterise model breast epithelial cell lines for their levels of candidate EMT markers and investigate if levels of O-GalNAc glycans are correlated with levels of EMT markers.
- 2) Investigate how O-GalNAc glycans are correlated with the migratory capability of cells.
- 3) Investigate how E-cadherin is correlated with the migratory capability of cells.
- 4) Investigate the potential involvement of O-GalNAc glycans in epithelial breast cell adhesion to endothelial monolayers under static conditions.

4.3 Methods

For the work described in this chapter, only hTERT-HME1, MCF-7 and T-47D breast epithelial cell lines were utilised in experiments. BT-474 and ZR-75-1 cell lines were not included in these experiments because they grow in multi-layered clusters and were therefore unsuitable for use in the wound healing assays where a confluent monolayer of cells is required. This growth pattern has also been documented in the literature where they are reported to rarely become confluent [244].

4.3.1 EMT marker identification

To investigate levels of EMT markers E-cadherin (epithelial marker), vimentin and N-cadherin (mesenchymal markers) immunofluorescence was utilised as described in '2.6 EMT marker immunofluorescence'. To confirm these results, Western Blotting was also utilised as described in '2.7 Western Blotting'.

4.3.2 Wound healing assay and MTT assay

To examine the migratory capacity of model breast epithelial cell lines a wound healing assay was employed as outlined in '2.5.4 *Wound healing assay*'. An MTT assay was also used as an indirect method to investigate the proliferation rates of the model breast epithelial cell lines, as outlined in '2.5.7 *MTT assay*'.

4.3.3 Labelling O-GalNAc glycans in the wound healing assay

To determine if there is a relationship between O-GalNAc glycans and cell migration, HPA labelling of cells in a wound healing assay was investigated as outlined in '2.5.5 *HPA labelling of wound healing assay*'.

4.3.4 Labelling E-cadherin and vimentin in the wound healing assay

The cells in the wound healing assay were labelled for E-cadherin or vimentin using immunofluorescence to investigate the relationship between migratory cells and the expression of EMT markers as outlined in '2.5.6 *E-cadherin, vimentin and HPA labelling of wound healing assay*'.

4.3.5 Static adhesion assay

To investigate if O-GalNAc glycans have a functional role in breast epithelial cell adhesion to endothelial monolayers under conditions of stasis, mimicking a cancer cell coming to arrest in a capillary bed, a static adhesion assay was utilised as outlined in '2.5.1 *Static adhesion assay*' with control as outlined in '2.5.2 *Static adhesion assay control*' and results analysed as outlined in '2.5.3 *Static adhesion assay image acquisition and quantification*'.

4.4 Results

4.4.1 EMT marker identification and relationship to O-GalNAc glycans

Primary antibody concentrations were optimised based on those used previously in the literature, see Supplementary Figure 12-14 for representative images of optimisation. Alongside the fluorescently labelled monoclonal antibodies directed against an EMT marker (either E-cadherin, vimentin or N-cadherin) DAPI was used to visualise the nucleus and HPA was used to label truncated O-GalNAc glycans, see Figure 4.1B, 4.2B and 4.3A.

Omitting the primary antibody acted as a control for non-specific binding, and omitting both primary and secondary antibodies controlled for autofluorescence, see Supplementary Figure 15 and 16A. HUVEC cells, known to label for N-cadherin [245], were used as a positive control for N-cadherin as the hTERT-HME1, MCF-7 and T-47D cell lines did not demonstrate labelling, see Supplementary Figure 16B.

Results demonstrate that MCF-7 and T-47D cells have significantly more E-cadherin than hTERT-HME1, see Figure 4.1A. This was confirmed by Western Blot analysis, see Figure 4.1C and D. hTERT-HME1 cells demonstrate significantly more vimentin than MCF-7 or T-47D cells, see Figure 4.2A. This was also confirmed by Western Blot analysis, see Figure 4.2C and D. hTERT-HME1, MCF-7 and T-47D did not show immunolabelling of N-cadherin, see Figure 4.3A and was confirmed by Western Blot analysis Figure 4.3B.

HPA-positive cells in MCF-7 and T-47D cell lines showed significantly more E-cadherin than HPA-negative cells, $p < 0.01$. There was no significant difference in E-cadherin labelling between hTERT-HME1 HPA-positive and HPA-negative cells, see Figure 4.4A. There was no significant difference in vimentin labelling of hTERT-HME1 between the HPA-positive and HPA-negative populations, see Figure 4.4B.

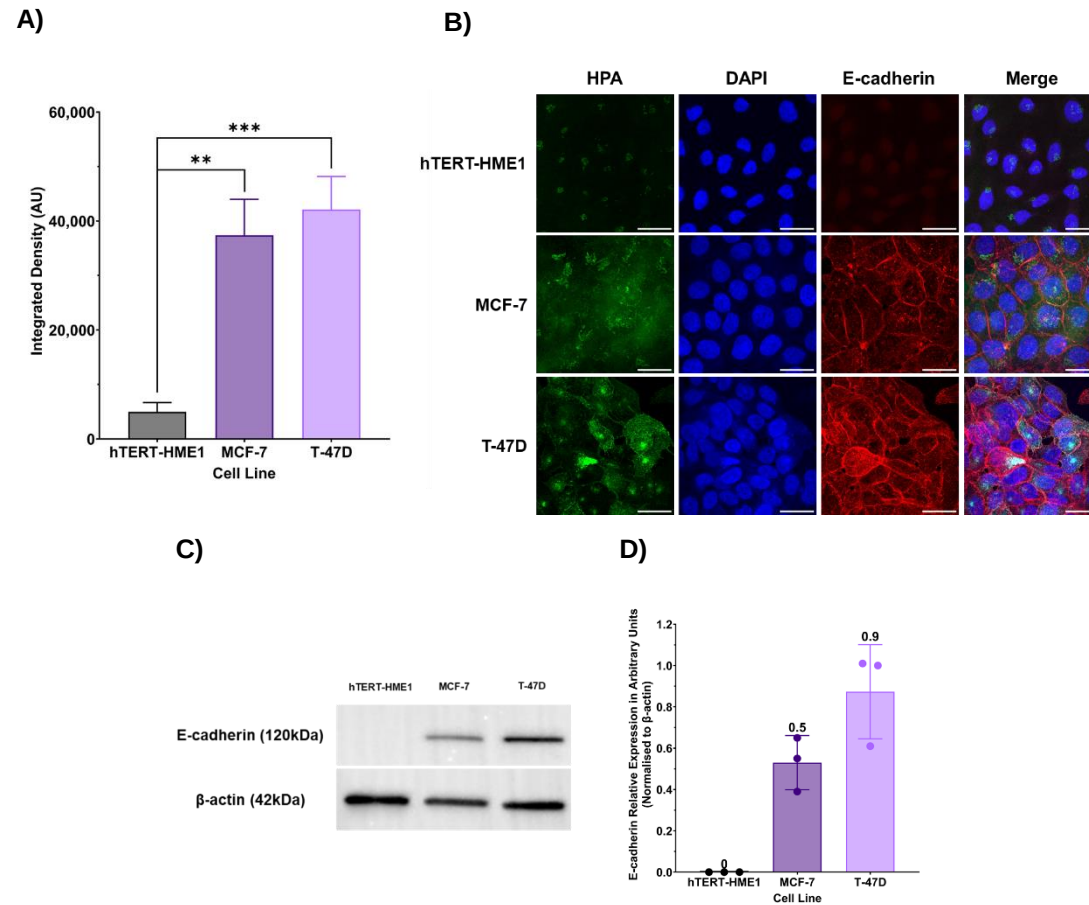


Figure 4.1 Levels of E-cadherin in hTERT-HME1, MCF-7 and T-47D cells. A) E-cadherin immunofluorescence. Average integrated density (AU) of E-cadherin labelling per cell of hTERT-HME1, MCF-7 and T-47D cell lines, three biological replicates ($\pm$ SD). Each biological replicate value consists of an average of three coverslips, for which three images were acquired. Unpaired Students t-test was used to test for significance, ** $p \leq 0.01$, *** $p \leq 0.001$. B) Representative Max Intensity Z-stack projections of E-cadherin immunofluorescence of hTERT-HME1, MCF-7 and T-47D cells. *Helix pomatia* agglutinin (HPA) (green), E-cadherin (red) and DAPI (blue). Scale bar = 25 μ m. C) Representative Western Blots for E-cadherin (120kDa) and loading control β -actin (42kDa) in hTERT-HME1, MCF-7 and T-47D cells. D) E-cadherin relative protein expression normalised to β -actin in hTERT-HME1, MCF-7 and T-47D cells in Western Blots, three biological replicates ($\pm$ SD).

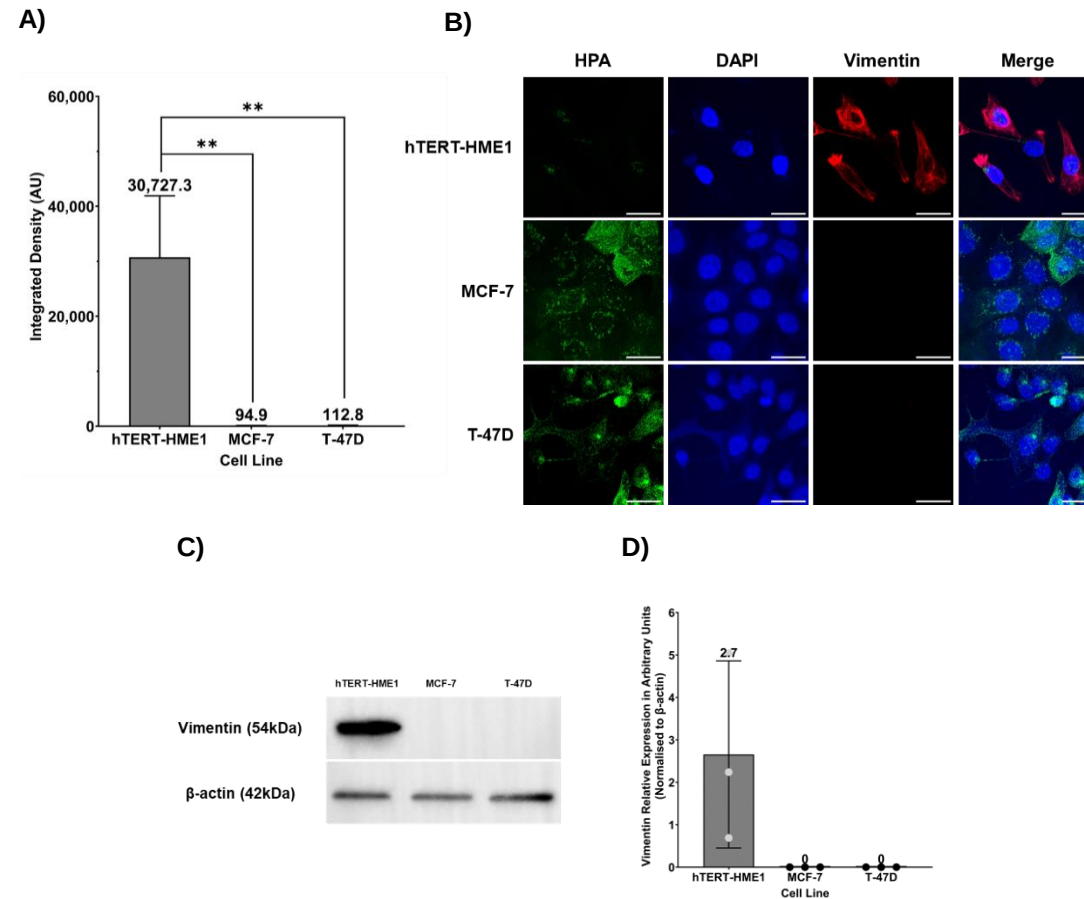


Figure 4.2 Levels of vimentin in hTERT-HME1, MCF-7 and T-47D cells. A) Vimentin immunofluorescence. Average integrated density (AU) of vimentin labelling per cell of hTERT-HME1, MCF-7 and T-47D cell lines, three biological replicates ($\pm$ SD). Each biological replicate value consists of an average of three coverslips, for which three images were acquired. Unpaired Students t-test was used to test for significance, ** $p \leq 0.01$. B) Representative Max Intensity Z-stack projections of vimentin immunofluorescence of hTERT-HME1, MCF-7 and T-47D cells. *Helix pomatia* agglutinin (HPA) (green), vimentin (red) and DAPI (blue). Scale bar = 25 μ m. C) Representative Western Blots for vimentin (54kDa) and loading control β -actin (42kDa) in hTERT-HME1, MCF-7 and T-47D cells. D) Vimentin relative protein expression normalised to β -actin in hTERT-HME1, MCF-7 and T-47D cells in Western Blots, three biological replicates ($\pm$ SD).

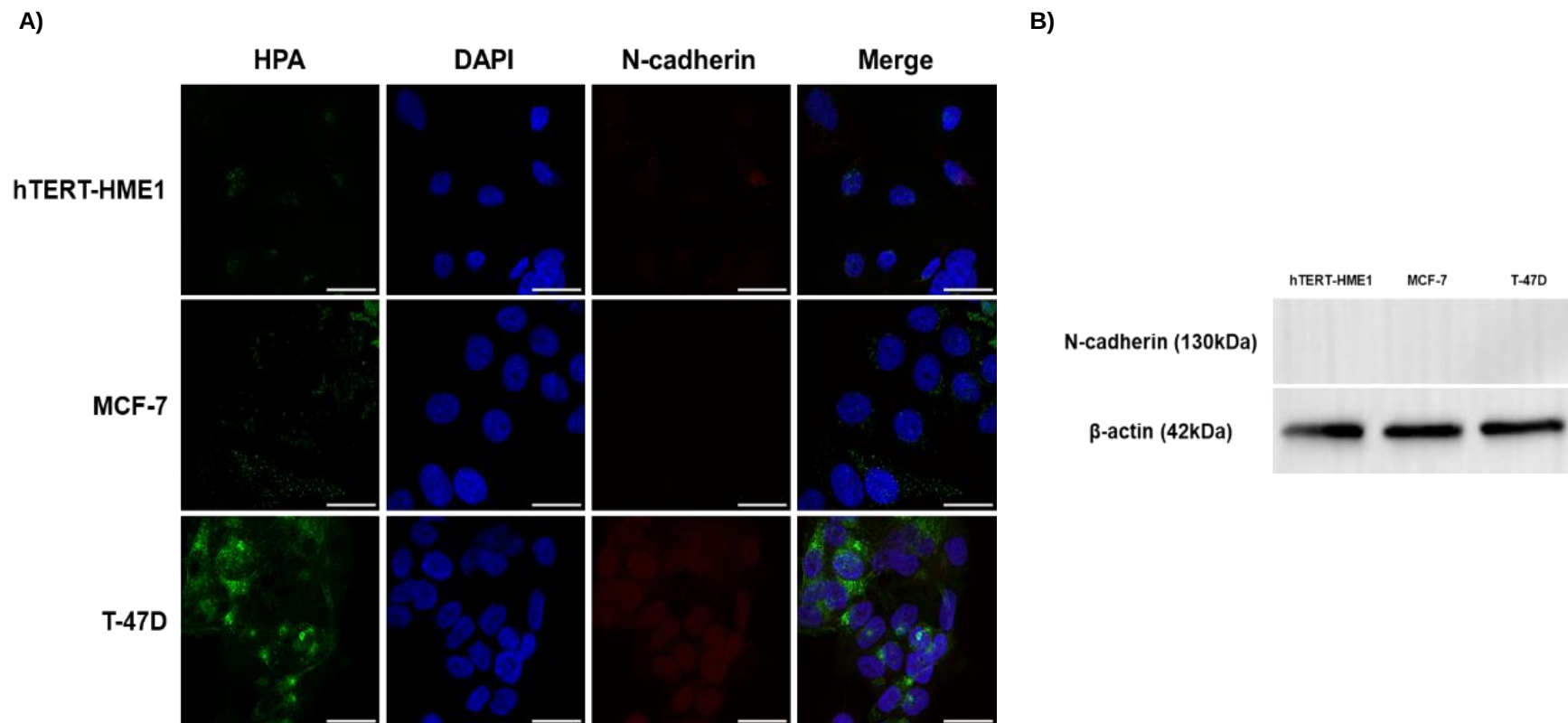


Figure 4.3 Levels of N-cadherin in hTERT-HME1, MCF-7 and T-47D cells. A) Representative Max Intensity Z-stack projections of N-cadherin immunofluorescence of hTERT-HME1, MCF-7 and T-47D cells. *Helix pomatia* agglutinin (HPA) (green), N-cadherin (red) and DAPI (blue). Scale bar = 25 μ m. B) Representative Western Blots for N-cadherin (130kDa) and loading control β -actin (42kDa) in hTERT-HME1, MCF-7 and T-47D cells.

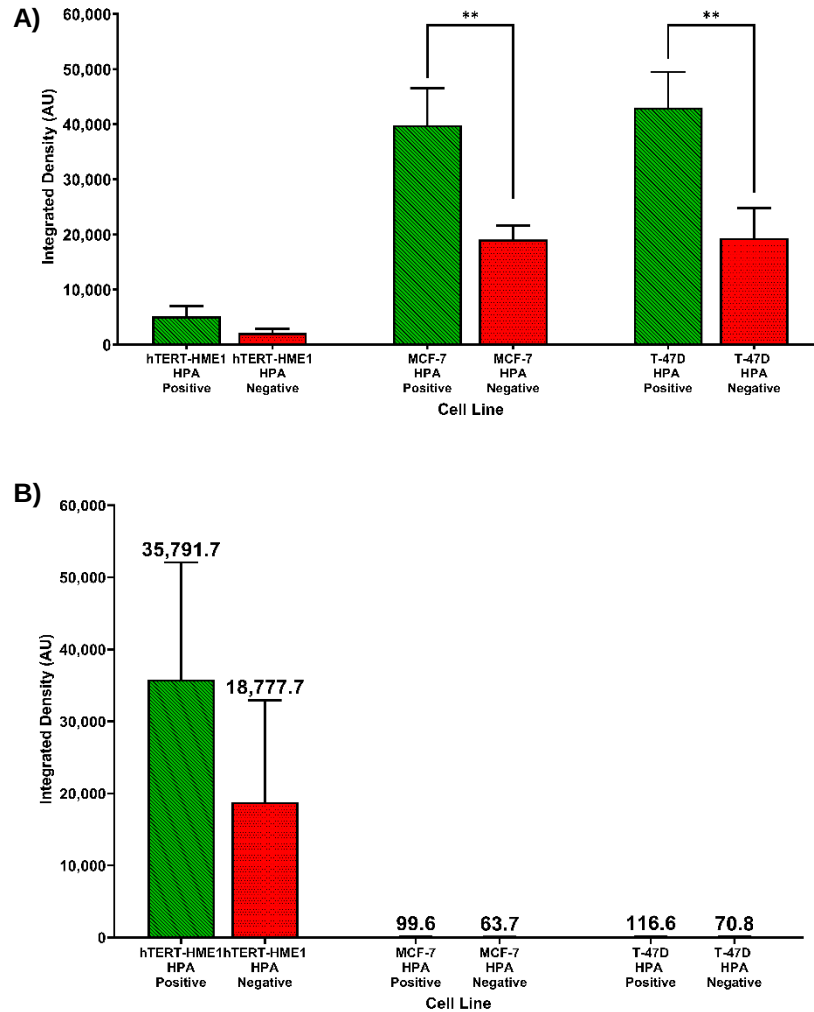


Figure 4.4 Levels of E-cadherin and vimentin protein marker in *Helix pomatia* agglutinin (HPA) -positive and -negative populations ($\pm$ SD). Each biological replicate value consists of an average of three coverslips, for which three images were acquired, and three biological replicates conducted. A) E-cadherin protein levels in HPA-positive and -negative cell populations. Unpaired Students t-test was used to test for significance, ** $p \leq 0.01$. B) Vimentin levels in HPA-positive and -negative cell populations. No significant difference was detected.

4.4.2 Wound healing assay

Both hTERT-HME1 and MCF-7 cells healed the 'wound' fully over 72 hours, Figure 4.5A. However, the T-47D cells only achieved a 27.8% (SD $\pm$ 10.1%) 'wound' closure over 72 hours, see Figure 4.5B. The MTT assay demonstrates that cell viability increased from 24 hours to 72 hours, a proxy to indicate cell division, see Figure 4.5C.

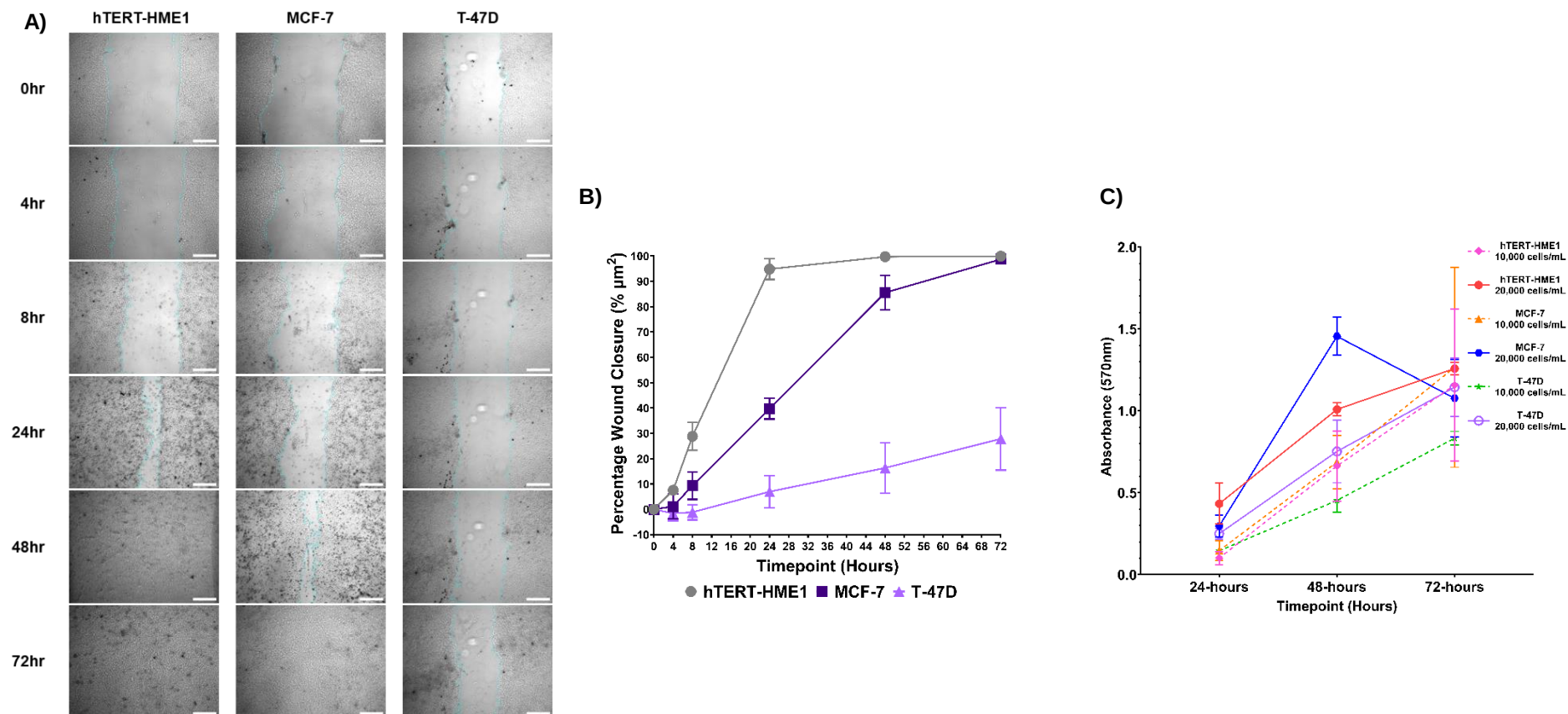


Figure 4.5 Wound healing and MTT assays. A) Representative phase contrast images of ‘wound’ closure imaged at 0, 4, 8, 24, 48 and 72-hour timepoints. The blue line indicates where the edge of the ‘wound’ was identified by the FIJI plugin “Wound_Healing_Size_Tool” [173]. Scale bar = 50 μm . B) Percentage of ‘wound’ closure (μm^2). Migration was calculated by subtracting from the ‘wound’ area timepoint at timepoint ‘0 hours’. An average of four technical replicates per biological replicate was calculated, and three biological replicates were completed ($\pm\text{SD}$). C) MTT assay used to measure the number of viable cells at 24, 48 and 72-hour timepoints. Each biological replicate consists of six technical replicates, and three biological replicates conducted per timepoint ($\pm\text{SD}$).

4.4.3 HPA labelling of the wound healing assay

The hTERT-HME1 cell line demonstrated negligible or no HPA labelling throughout all the timepoints and 'healed the wound' fully by the 72-hour timepoint, see Figure 4.6A. When quantified, the hTERT-HME1 cells did not show any significant difference in HPA labelling between the 'edge of the wound' and 'main body of cells' across all timepoints, see Figure 4.8A.

In the MCF-7 cell line, visually it is very clear that as the timepoints progress, the HPA-negative cells (cells exhibiting red membrane labelling only) were the predominant population at the 'edge of the healing wound'. When the 'wound' was fully healed there was a reversion back to a mixed population of HPA-positive and -negative cells, see Figure 4.6B and 4.7A. When quantified, there was a significant reduction in HPA labelling of MCF-7 cells at the 'edge of the wound' compared to the 'main body of cells' at the 8-hour and 24-hour timepoint and this difference remained visible, although the difference was non-significant, until the 48-hour timepoint. There was then a clear visible reversion back to a mixed population, with no significant difference between the 'edge of wound' and the 'main body of cells' as the 'wound' was fully healed at the 72-hour timepoint see Figure 4.8B.

At the 24, 48 and 72-hour timepoint, T-47D cells demonstrated a clear visual predominance of HPA-negative cells at the 'edge of the wound', see Figure 4.6C and 4.7B. This cell line, consistent with previous results, failed to heal the 'wound' fully in this period. When quantified, the T-47D cells demonstrated significantly less HPA labelling of the 'edge of the wound' at 24 hours and 72 hours, the effect was visible at 48 hours but the results were non-significant at this timepoint, see Figure 4.8C.

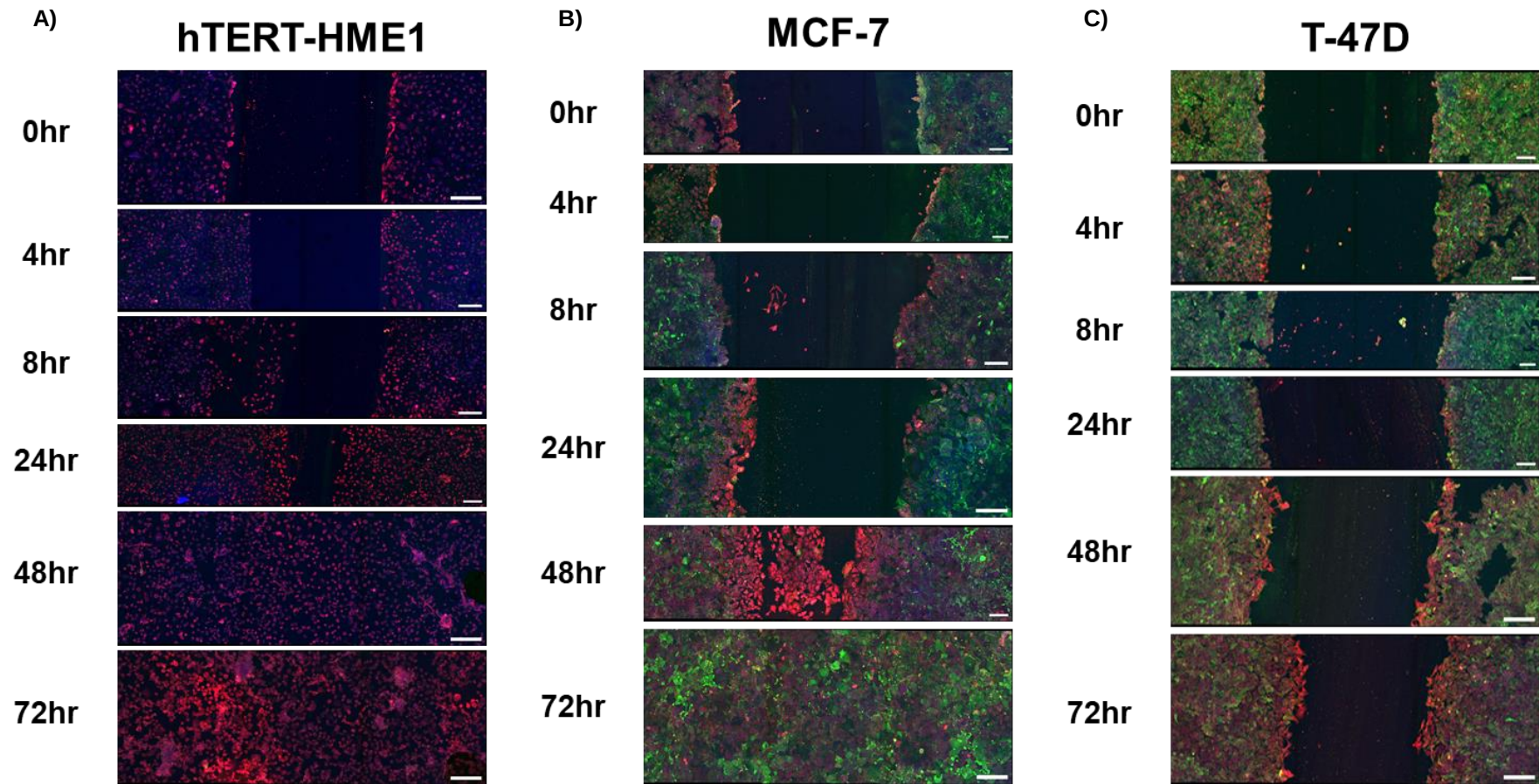


Figure 4.6 Fluorescent *Helix pomatia* agglutinin (HPA) labelling of the wound healing assay. Representative Max Intensity Z-stack projections of wound healing assays at 0, 4, 8, 24, 48 and 72 hours for A) hTERT-HME1, B) MCF-7 and C) T-47D cell lines. Labelled with HPA (green), CellMask™ cell membrane staining (red), and DAPI nucleus staining (blue). Scale bar = 150µm.

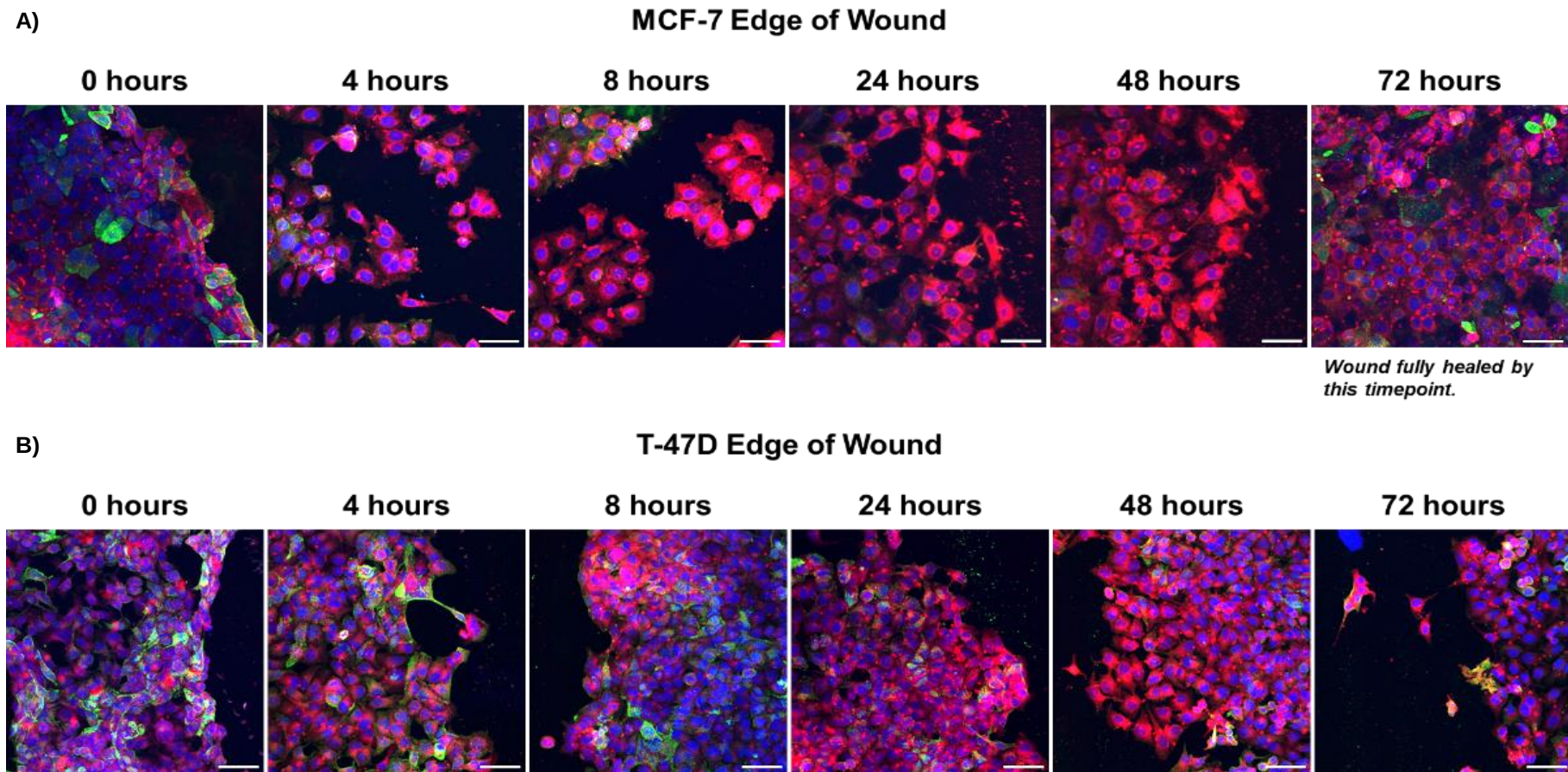


Figure 4.7 Regions of interest in fluorescent *Helix pomatia* agglutinin (HPA) labelling of the wound healing assay. Representative Max Intensity Z-stack projections of the 'leading edge' of the healing 'wound' at 0, 4, 8, 24, 48 and 72 hours for A) MCF-7 and B) T-47D cell lines. Labelled with HPA (green), CellMask™ cell membrane staining (red), and DAPI nucleus staining (blue). Scale bar = 50µm.

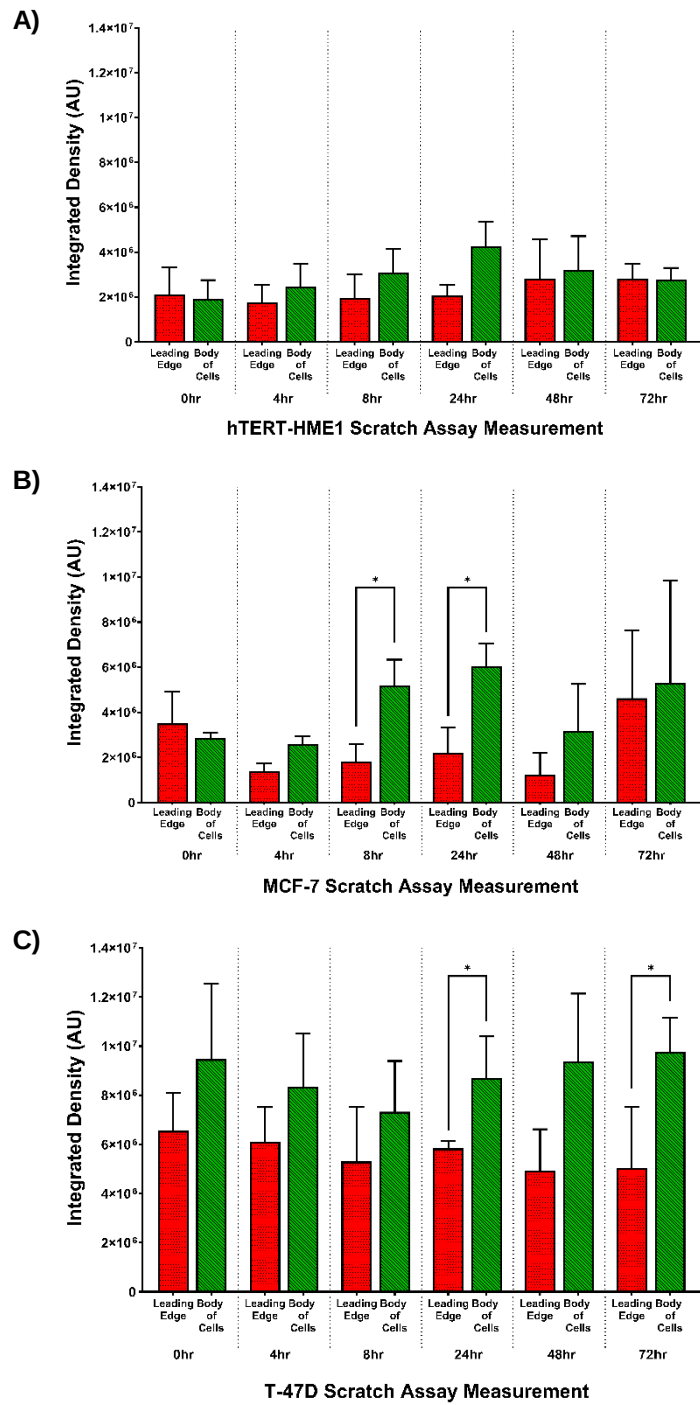


Figure 4.8 Mean fluorescence (AU) of *Helix pomatia* agglutinin (HPA) labelling at the 'leading edge' of the healing 'wound' and the 'main body' of cells in the wound healing assay. A) hTERT-HME1 cell line, B) MCF-7 cell line and C) T-47D cell line, three biological replicates ($\pm$ SD). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$.

4.4.4 E-cadherin labelling of the wound healing assay

Consistent with previous results that demonstrated little to no E-cadherin labelling in the hTERT-HME1 cell line, the wound healing assay demonstrated that there is little to no E-cadherin labelling in this cell line at all timepoints, see Figure 4.9A. When quantified, there was no significant difference in E-cadherin labelling between the 'edge of the wound' and the 'main body of cells' across all timepoints, see Figure 4.11A.

MCF-7 cells demonstrated a similar E-cadherin labelling pattern to that seen with HPA. Visually, the E-cadherin labelling was reduced at the 'edge of the wound' at the 8, 24 and 48-hour timepoint, see Figure 4.9B and 4.10A. MCF-7 cells appear visually to restore consistent E-cadherin labelling across the cell monolayer as the 'wound' is fully healed at the 72-hour timepoint. When quantified, there was a significant reduction in E-cadherin labelling at the 'edge of the wound' compared to the 'main body of cells' at the 4, 8 and 24-hour timepoints. There was then a reversion back to a mixed population, with no significant difference between the 'edge of wound' and the 'main body of cells' as the 'wound' fully healed at the 72-hour timepoint, see Figure 4.11B.

At the 48 and 72-hour timepoint, the T-47D cells demonstrated a visual reduction of E-cadherin labelling at the 'edge of the wound', see Figure 4.9C and 4.10B. Again, this cell line failed to fully 'heal the wound' in the 72 hours. When quantified, there was significantly less E-cadherin labelling of the 'edge of the wound' at 24 and 48 hours, with labelling still visibly reduced, but non-significantly so, at 72 hours, see Figure 4.11C.

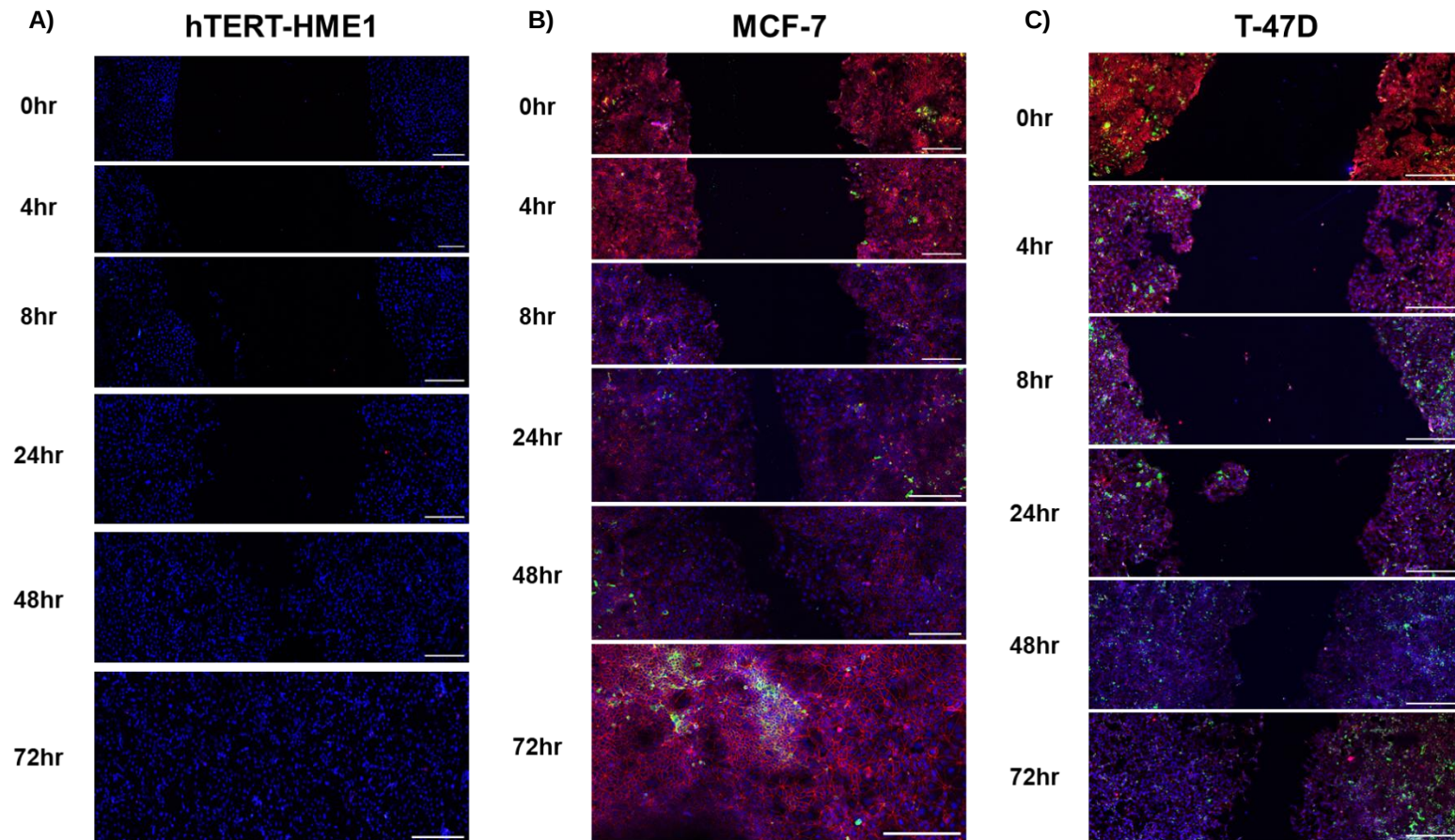


Figure 4.9 E-cadherin immunofluorescence and fluorescent *Helix pomatia* agglutinin (HPA) labelling of the wound healing assay. Representative Max Intensity Z-stack projections of wound healing assays at 0, 4, 8, 24, 48 and 72 hours for A) hTERT-HME1, B) MCF-7 and C) T-47D cell lines. Labelled with HPA (green), E-cadherin (red), and DAPI nucleus staining (blue). Scale bar = 150 μ m.

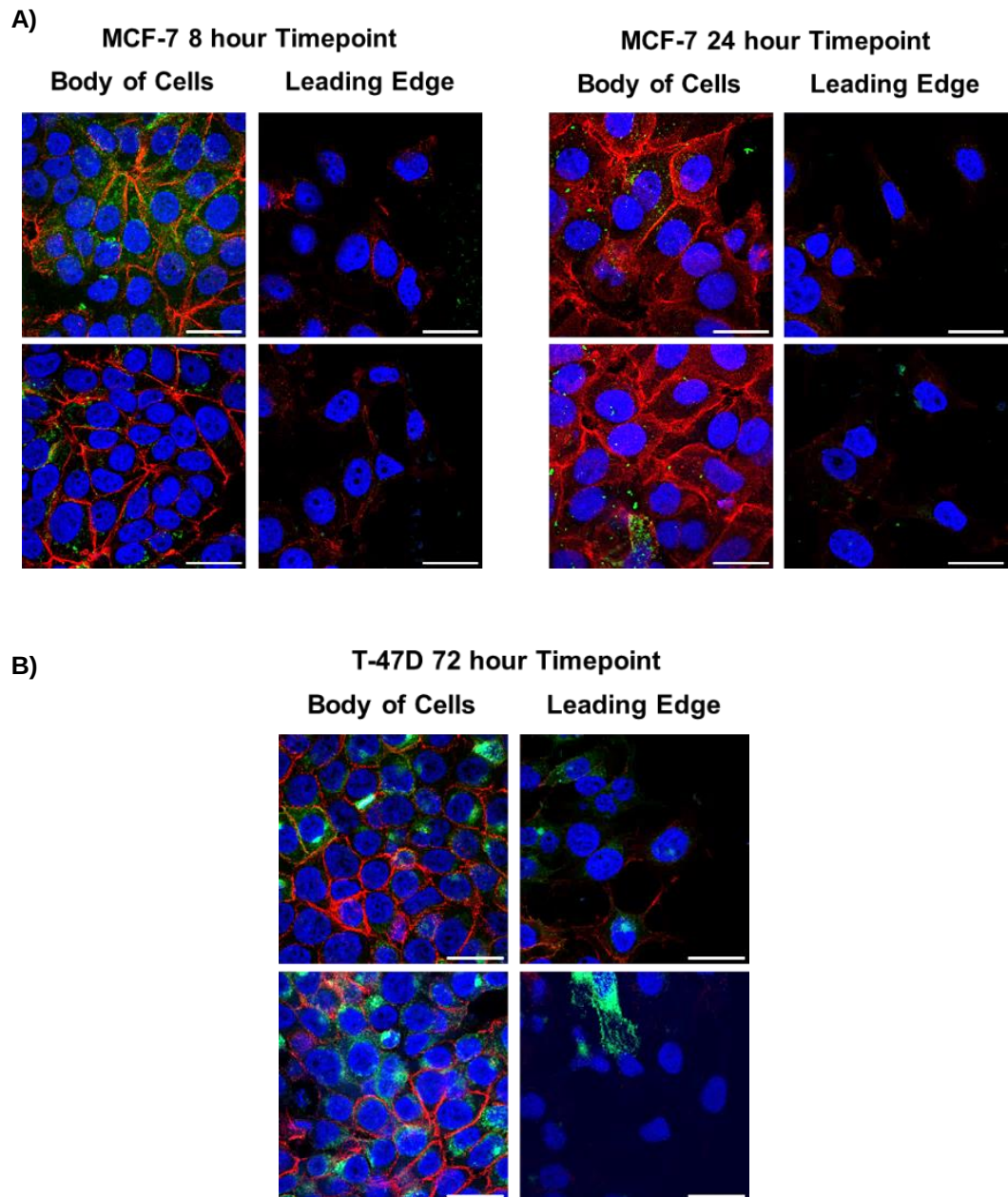


Figure 4.10 Regions of interest in E-cadherin immunofluorescence and fluorescent *Helix pomatia* agglutinin (HPA) labelling of the wound healing assay. Representative Max Intensity Z-stack projections of the 'leading edge' of the healing 'wound' at 8 and 24 hours for A) MCF-7 and at 72 hours for B) T-47D cell lines. Labelled with HPA (green), E-cadherin (red), and DAPI nucleus staining (blue). Scale bar = 50µm.

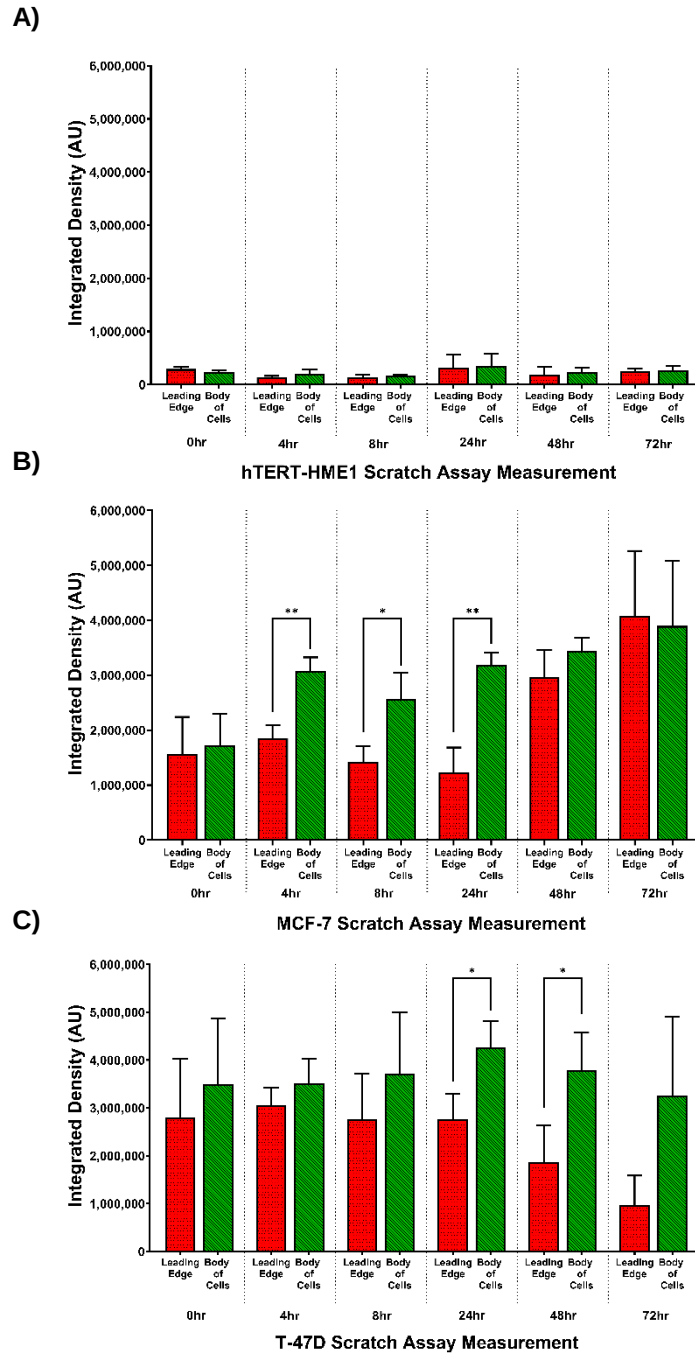


Figure 4.11 Mean fluorescence (AU) of E-cadherin labelling at the 'leading edge' of the healing 'wound' and the 'main body of cells' in the wound healing assay. A) hTERT-HME1 cell line, B) MCF-7 cell line and C) T-47D cell line, three biological replicates ($\pm$ SD). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$.

4.4.5 Static adhesion assay

Results demonstrated that masking HPA-binding truncated O-GalNAc glycans with HPA significantly reduced MCF-7 and T-47D cell adhesion to endothelial cells after 10 minutes compared to unmasked cells, $p < 0.05$, see Figure 4.12. This effect was not significant for the hTERT-HME1 cells, which have negligible levels of HPA-binding O-GalNAc glycans and showed very minimal adhesion compared to the breast cancer cell lines. Masking ConA-binding glycans led to a slight but non-significant reduction in breast cell adhesion to endothelial cells compared to unmasked cells in the hTERT-HME1 and MCF-7 cells. The reduction was significant ($p < 0.05$) in the T-47D cell line.

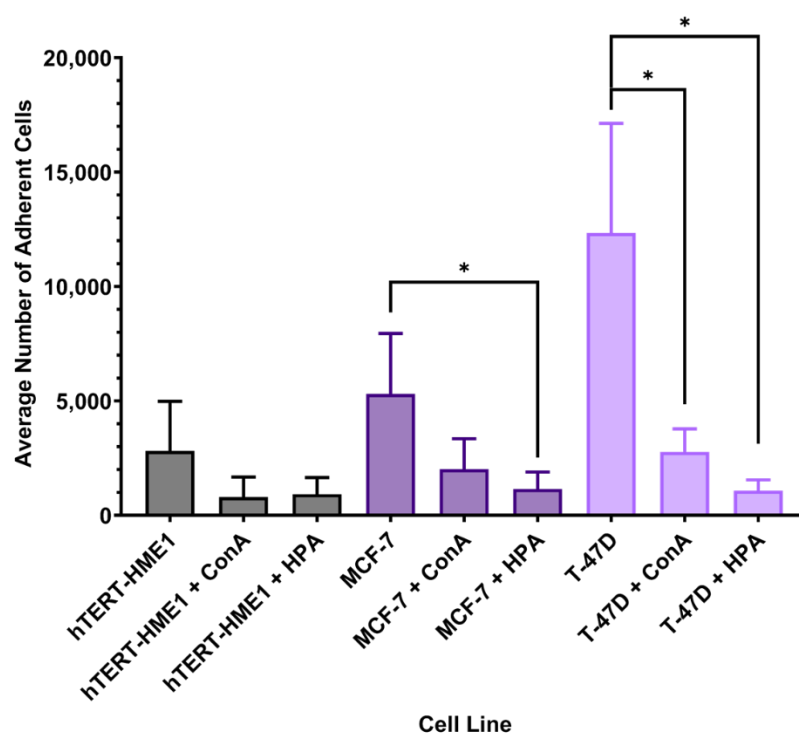


Figure 4.12 Static adhesion assay. The average number of adherent cells, hTERT-HME1, MCF-7 or T-47D, to a HUVEC endothelial cell monolayer after 10 minutes under static adhesion assay. Conditions include; non-masked cells, masking cell surface glycans with the Concanavalin A (ConA) lectin or *Helix pomatia* agglutinin (HPA) lectin. Each technical replicate consists of 6 replicates, three biological replicates per cell line were completed, data presented as the mean ($\pm$ SD). Unpaired Students t-test was used to test for significance, * $p < 0.05$.

4.5 Discussion

4.5.1 O-GalNAc glycans are associated with E-cadherin protein levels

To further investigate the relationship between O-GalNAc glycans and EMT, model breast epithelial cell lines hTERT-HME1, MCF-7 and T-47D were investigated for their levels of key protein markers of EMT, including the epithelial marker E-cadherin, and mesenchymal markers vimentin and N-cadherin. As EMT is often not a full binary switch into a 'full' mesenchymal state, with a spectrum of pEMT states present, as described in '1.4.2 Partial EMT', it is important to consider the possibility of a spectrum of EMT marker expression in these populations [128]. Loss of E-cadherin however, is still often considered the major indispensable hallmark of EMT [246, 247]. Often key EMT identifiers are also used alongside the functional markers of increased migration, elongated cell shape and increased invasion to help define the EMT phenotype of a cell [117].

Results presented in this chapter showed that original parental populations of MCF-7 and T-47D cell lines demonstrated significantly more E-cadherin immunopositivity than the hTERT-HME1 cell line. These results are supported by the literature which consistently indicates that both MCF-7 and T-47D express E-cadherin [248-251]. Previous studies have also not detected E-cadherin expression in hTERT-HME1 [249].

Conversely, the original parental population of hTERT-HME1 demonstrated significantly more vimentin immunopositivity than MCF-7 and T-47D cell lines. Previous literature supports this, and has reported that hTERT-HME1 cells express vimentin, an unexpected finding for what is characterised as an epithelial cell line with a 'normal phenotype' [251]. An explanation for this is that human telomerase reverse transcriptase (hTERT) immortalisation, which involves stable transfection of cells with hTERT, has been linked to promoting EMT, increasing vimentin protein expression and suppressing E-cadherin [252]. MCF-7 and T-47D have also consistently been documented in the literature as not expressing vimentin [248, 251, 253].

There was no significant difference in immunopositivity of vimentin between HPA-positive and -negative cells across all cell lines. Both MCF-7 and T-47D cells demonstrate very minimal vimentin immunolabelling. Raghu et al., [254] demonstrated that when mesenchymal trophoblast stem cells re-expressed GALNT3, resulting in the restoration of an epithelial cell state, it also caused a reduction in vimentin gene expression, as identified by qPCR. Indicating a potential link between O-GalNAc glycans and expression of the mesenchymal marker vimentin. A relationship between mesenchymal-like and epithelial-like phenotypes and vimentin was not observed in the study reported here.

In the study reported here, none of the model breast epithelial cell lines demonstrated immunopositivity for N-cadherin. The positive control, HUVECs, which are known to demonstrate immunopositivity for N-cadherin [245] demonstrated labelling. These results are also consistent with the literature where other studies have failed to detect N-cadherin expression in hTERT-HME1 [249], MCF-7 [249, 255] and T-47D cell lines [255, 256].

When cell lines were stratified into HPA-positive and -negative cell sub-populations, results showed that HPA-positive cells in MCF-7 and T-47D cell lines had significantly more E-cadherin immunopositivity than HPA-negative cells, $p < 0.01$. No significant difference was seen between HPA-binding sub-populations in the hTERT-HME1 cell line. E-cadherin is known to be modified by both O-linked glycosylation and N-linked glycosylation [257]. Importantly, Vester-Christensen et al., [258] demonstrated using mass spectrometry that E-cadherin possesses an O-GalNAc glycosite at an extracellular domain (EC1). Raghu et al., [254] demonstrated the importance of the O-GalNAc glycosite on E-cadherin, created by GALNT3, for the retention of an epithelial cell phenotype through correct localisation of E-cadherin and the formation of adherens junctions. In human mammary epithelial cells (HMECs), the loss of GALNT3, by short hairpin RNA (shRNA), resulted in intracellular retention of E-cadherin in the Golgi apparatus, as identified by co-localisation with giantin, induction of EMT and increased cell motility. In mesenchymal trophoblast stem cells, re-expression of GALNT3 resulted in the restoration of an epithelial cell state [254]. These results are consistent with results reported in this chapter that associate the O-GalNAc glycosylation of a cell with the epithelial phenotype and the presence of the epithelial protein E-cadherin.

4.5.2 O-GalNAc glycans are associated with functional aspects of EMT

Wound healing assays are the most common *in vitro* technique to measure cell migration and motility [173]. It is a relatively inexpensive and technically straightforward assay that requires no specialist equipment; however, reliability is variable and dependent on the accuracy of the operator when inflicting the 'wound' [171]. However, using a plugin for image analysis increases reproducibility [171, 173, 259]. Therefore, in this study, the wound healing assay, using a plugin for image analysis, was employed to investigate if cell motility was associated with O-GalNAc glycosylation status. It is important to consider that the effect of cell migration and cell division cannot be differentiated post-24 hours and therefore, the impact of cell division in healing the 'wound' must also be considered.

Interestingly, cell lines with the most E-cadherin immunopositivity, T-47D and then MCF-7, healed the 'wounds' inflicted in the assay the most slowly. Whereas, hTERT-HME1, which

demonstrates high vimentin immunopositivity was the quickest to heal the 'wound'. An explanation for this is that E-cadherin is a cell-cell adhesion molecule widely considered to suppress cell motility [260]. Liang et al., [261] demonstrated that increased E-cadherin expression as a result of metformin treatment in melanoma cell line B16F10 caused a marked inhibition of cell motility and ability to heal a 'wound' in a wound healing assay. Whereas, vimentin is widely associated with mesenchymal, single-cell, and motile-cell behaviours [262]. MCF-7 cells transfected with vimentin cDNA have been reported to move at ~66% faster speed and demonstrate more single-cell migration than seen in the untransfected, wild-type MCF-7 cells [263].

HPA-negative cell populations, demonstrated in previous experiments reported in Chapter 3 to exhibit a mesenchymal-like elongated morphology, were significantly more concentrated at the 'leading edge' of the healing 'wound' than in the undisturbed monolayer. In MCF-7 cells, this difference was significant at the 8 and 24-hour timepoints and remained visible, although non-significant, until 48 hours. In T-47D cells this concentration at the 'leading edge' was significant at 24 and 72 hours, and visible, but non-significant, at 48 hours. There was then a reversion back to a mixed population of HPA-positive and -negative cells at 72 hours in the MCF-7 cells as the 'wound' fully healed, with no significant difference in HPA labelling between the 'edge of the wound' and the 'main body of the wound'. The T-47D breast cancer epithelial cell line did not fully heal the 'wound' over 72 hours, and the literature also documents incomplete 'wound' closure over this time period using this cell line [244, 264]. hTERT-HME1 cells demonstrated little HPA labelling throughout the timepoints, and no significant difference between the 'edge of the wound' and the 'main body of cells' across all timepoints. This was to be expected since they generally exhibited negligible HPA labelling.

Altered glycosylation, such as increased sialylation of O-linked glycans, has been associated with increased cell motility in wound healing assays and adhesion to ECM components such as fibronectin in the literature [265, 266]. Overexpression of ST3GAL1 (makes sialylated core 1 glycans) is associated with cell growth, migration and invasion in ovarian cancer cell lines, SOV-3, OVCAR3 and A2780 [267]. Other studies have correlated Tn antigen levels with cell motility. For example, Gill et al., [72] demonstrated that ER localisation of GALNT2 in the MDA-MB-231 breast cancer cell line led to high levels of Tn antigen, up to 10-fold higher, detected by HPA and VVA in fluorescent microscopy compared to untransfected cells. In a wound healing assay, on fibronectin-coated plates, cells with ER-localised GALNT2 and high levels of Tn antigen, migrated around four times faster than untransfected cells and demonstrated larger lamellipodia. This faster migration was sustained throughout the 4 hours investigated, so the authors concluded that elevated O-linked glycosylation promotes faster and sustained cell migration. Chia et al., [73] demonstrated that extracellular signal-regulated kinase 8 (ERK8) depletion using RNAi increased Tn antigen levels, identified by HPA labelling, present in cervical cancer HeLa cells. ERK8-depleted cells, on

fibronectin-coated plates, healed the 'wound' two-fold faster than non-depleted cells. Knockout of both GALNT1 and -T2 resulted in reduced cell motility in a wound healing assay. Hofmann et al., [268] knocked down COSMC, an essential chaperone for C1GALT1 functionality, in Panc-1 pancreatic cancer cells and it resulted in a significant increase in cell migration in a wound healing assay. It was also associated with increased Tn and sialyl-Tn antigens identified by sialyl-Tn and Tn immunocytochemistry. None of the studies investigated the cell surface glycosylation profile of the cells during the wound healing, and therefore could not discern if Tn antigen-rich cells actually demonstrate enhanced migration.

In the study reported here, 'healing wounds' were also probed for immunopositivity for E-cadherin (epithelial marker) at each timepoint (0hr, 4hrs, 8hrs, 24hrs, 48hrs and 72hrs). Results showed that there was a significant reduction in E-cadherin labelling of MCF-7 cells at the 'edge of the wound' compared to the 'main body of cells' at the 4, 8 and 24-hour timepoints. There was then a reversion back to a mixed population, with no significant difference between the 'edge of wound' and the 'main body of cells' as the 'wound' was fully healed at the 72-hour timepoint. T-47D cells demonstrated significantly less E-cadherin labelling of the 'edge of the wound' compared to the 'main body of cells' at 24 and 48 hours, with a visible but non-significant reduction at 72 hours. This cellular heterogeneity is also widely reported *in vivo*, where there is the repression of E-cadherin at the edge of a solid tumour or the invasive front of multicellular clusters and it is highly expressed in the main body of the tumour [269]. Loss of E-cadherin is often regarded as the most characteristic protein marker of EMT in epithelial tumours. The exact mechanisms underlying why EMT is often localised at the edge of tumours are not fully understood but, it is thought to relate to access to ECM components including growth factors, hypoxic conditions and extracellular vesicles [225, 270].

Other studies have documented similar observations regarding E-cadherin localisation in wound healing assays to those reported in this thesis. Wager et al., [271] demonstrated that the 'leading edge' of a 'wound' inflicted in a monolayer of normal epidermal keratinocyte cells, NHEK, showed slightly reduced transcript levels of E-cadherin after 24 hours, alongside increased expression of vimentin. These results suggested a potential link to cells undergoing pEMT, neither full loss of epithelial characteristic nor full gain of mesenchymal, at the 'leading edge' of a healing 'wound' [271]. Ritch et al., [272] utilised ovarian cancer cell line PEO1 in a wound healing assay and immunolabelled for both E-cadherin and N-cadherin. At the 'edge of the wound,' PEO1 demonstrated minimal clusters of E-cadherin labelled cells with a majority of cells individually arranged and labelling for N-cadherin. Cells away from the 'wound' displayed a heterogeneous profile of E-cadherin cells surrounded by N-cadherin labelling cells.

In the experiments reported here, healing 'wounds' were also probed for levels of immunopositivity of vimentin (mesenchymal marker) at each timepoint (0hr, 4hrs, 8hrs,

24hrs, 48hrs and 72hrs), see Supplementary Figure 17. Consistent with previous results, only the hTERT-HME1 cells labelled for vimentin and there was no visible vimentin immunopositivity in cells at the 'edge of the healing wound' or the 'main body of cells' at any of the timepoints in the MCF-7 and T-47D cell line assays. Previous literature has reported detectable vimentin in wound healing assays, and in a cell line that did not originally express the protein. The normal human breast cell line MCF710A were labelled for vimentin in a wound healing assay and the cell's migratory speed and trajectories were measured using nuclei tracking. Vimentin mRNA and protein expression was induced solely in cells at the 'edge of the wound' and these cells actively migrated towards the centre of the 'wound'. Cells expressing the most vimentin were found to migrate fastest. Vimentin expression was not present when the 'wound' was initially created, and expression disappeared following the 'closure of the wound' [273]. In human skin, migrating epithelial cells in wound margins have been shown to express vimentin and gain mesenchymal features, potentially linked to a pEMT state, following exposure to TNF- α (an EMT-inducer) [274]. Differences between the observations of studies in the literature and those reported here could be explained by the use of different cell lines, as well as the differences in the use of EMT-inducing agents such as TNF- α .

A static adhesion assay was employed to investigate the functional involvement of terminal O-GalNAc glycans on the adhesive capabilities of model epithelial cell lines. This type of assay is commonly used to model the adhesion of cells to vascular endothelium under conditions of stasis, such as when cells are slowed or trapped in small capillaries, as during metastatic spread. The assay does not require specialised equipment and is quick to perform [275]. A cytoplasmic stain, HPTS, was chosen to label the breast cells since it is less likely than compounds that label cell surface molecules to affect adhesion [276]. Masking O-GalNAc glycans with HPA significantly reduced MCF-7 and T-47D cell adhesion to endothelial cells after 10 minutes compared to unmasked cells, $p < 0.05$. This suggests a functional role in cell adhesion for truncated O-GalNAc glycans, seen in O-GalNAc-rich MCF-7 and T-47D cells, under static conditions. Cell adhesion was also significantly reduced in T47D cells, but not for MCF-7 cells, when glycans were masked with ConA, suggesting that α -D-mannosyl and α -D-glucosyl residues may also be involved in adhesive processes for T-47D cells. hTERT-HME1 cells, which demonstrate negligible levels of terminal O-GalNAc glycans, showed minimal adhesion compared to the other breast cancer epithelial cell lines, and masking with either ConA or HPA had no effect.

Truncated O-GalNAc glycans have previously been implicated in cell adhesion and have been shown to interact with galectin-3, facilitating initiation adhesion events with endothelial cells [277]. This is critical during metastasis, as it allows the cancer cells to adhere to the vascular endothelium to extravasate and form a metastatic site. Bapu et al., [172] demonstrated that masking truncated O-GalNAc glycans, by pre-incubating MCF-7, BT-474 and ZR-75-1 cells with HPA, resulted in a significant reduction in cell adhesion to a

monolayer in MCF-7 and ZR-75-1 cells under conditions of sweeping flow. These results are consistent with those demonstrated in this thesis. This effect demonstrated by [172] was also evident when GalNAc monosaccharide was added, which should bind to putative GalNAc-receptors on the endothelial cells, thus blocking interactions between epithelial cell surface GalNAc-glycans and endothelial cell receptors thereby reducing adhesion. However, free GalNAc had damaging effects on the endothelial monolayer. Anti-Tn monoclonal antibodies have also been employed to investigate the role of O-GalNAc glycans in cell adhesion and demonstrated inhibition of MCF-7 cell adhesion to endothelial cells [278] and inhibition of ZR-75-30 breast cancer cell line adhesion to HBL-100, a normal breast epithelial cell line, implicating an additional role in facilitating tumour and adjacent normal cell adhesions [279].

Taken together, the observations reported in this chapter are consistent with the reversible, dynamic process of EMT whereby adhesive, epithelial-like HPA-positive cells can trans-differentiate into HPA-negative, migratory mesenchymal-like cells, and then also revert, by MET. Results reported here are consistent with cells undergoing a pEMT state rather than full differentiation, as they do not fully lose their epithelial markers nor do they fully gain the mesenchymal ones. Results suggest that loss of E-cadherin is associated with the loss of O-GalNAc glycans, putatively the Tn antigen, and this process is enough for a cell to become motile in a wound healing assay. The reverse process, MET, results in O-GalNAc glycan-rich cells which are more adhesive to endothelial cells, a key step in dissemination via the bloodstream and ultimately metastatic progression. Thus, plasticity in glycosylation provides cancer cells with access to dynamic phenotypes that they can exploit to become successful at each point in the metastatic cascade.

4.6 Key findings

- MCF-7 and T-47D cells have significantly more E-cadherin (epithelial marker) than hTERT-HME1 cells. hTERT-HME1 cells feature significantly more vimentin (mesenchymal marker) than MCF-7 and T-47D cells. None of these cell lines feature detectable levels of N-cadherin (mesenchymal marker).
- HPA-positive, O-GalNAc glycan-rich, MCF-7 and T-47D cells demonstrated significantly more E-cadherin than their HPA-negative, O-GalNAc glycan-poor, counterparts. There was no significant relationship between HPA-binding and vimentin immunolabelling.

- HPA-negative, O-GalNAc glycan poor, MCF-7 and T-47D cells were concentrated at the 'leading edge of the healing wound'. There were significantly lower levels of O-GalNAc glycans on cells at the 'leading edge of the healing wound' than the 'main body of cells' at 8 and 24 hours in MCF-7 cells and 24 and 72 hours in T-47D cells. When the MCF-7 'wound' was fully healed there was a reversion to a mixed population of HPA-positive and -negative cells, with no significant difference in levels of O-GalNAc glycans.
- MCF-7 and T-47D cells at the leading 'edge of the healing wound' also demonstrated significantly reduced E-cadherin compared to those in the 'main body of cells'. This was significant at 4, 8 and 24 hours in MCF-7 cells and at 24 and 48 hours in T-47D cells.
- MCF-7 and T-47D cell line adhesion to an endothelial monolayer in a static adhesion assay could be significantly reduced by masking the O-GalNAc glycans with the HPA lectin, indicating a potential functional role in cancer cell adhesion to endothelium during metastasis.

Chapter 5

Investigating the effect of transforming growth factor- β on O-GalNAc glycans

5. Investigating the effect of transforming growth factor- β on O-GalNAc glycans

5.1 Background

5.1.1 Cytokine-induced EMT *in vitro*

A variety of cytokines have been implicated in the induction of EMT, including EGF, fibroblast growth factor (FGF), hepatocyte growth factor (HGF), insulin-like growth factor (IGF) 1 and 2 and TGF- β (reviewed by [280]). Other cytokines, secreted by the tumour stroma can act to promote EMT, including interleukin-6 (IL-6) [281], interleukin-8 (IL-8) [282] and TNF- α [283]. Among these cytokines, the pleiotropic cytokine TGF- β has received the most attention due to its efficiency in inducing EMT and its strong link to EMT induction in both normal development and disease states (reviewed by [284] and [285]). The role of TGF- β in inducing EMT was first recognised by Miettinen et al., [286] in cell culture of mouse mammary epithelial cells, NMuMG, through observation of changes in morphology (to elongated, spindle-shaped morphology), decreased expression of E-cadherin and increased fibronectin, detected by Western Blotting and immunofluorescence. Cytokines act by binding to the respective cytokine receptors and inducing respective signalling pathways. The pathways of EMT-inducing cytokines converge and result in the regulation of a variety of EMT-inducing transcription factors, such as Snail, TWIST and ZEB1/2 which act to suppress the transcription of E-cadherin and induce transcription of N-cadherin, vimentin and fibronectin promoting cell polarity loss and destroying tight junctions [287, 288], see Figure 1.16 for an overview. Incomplete or partial activation of signalling pathways and transcription factors can result in the pEMT state (reviewed by [289]).

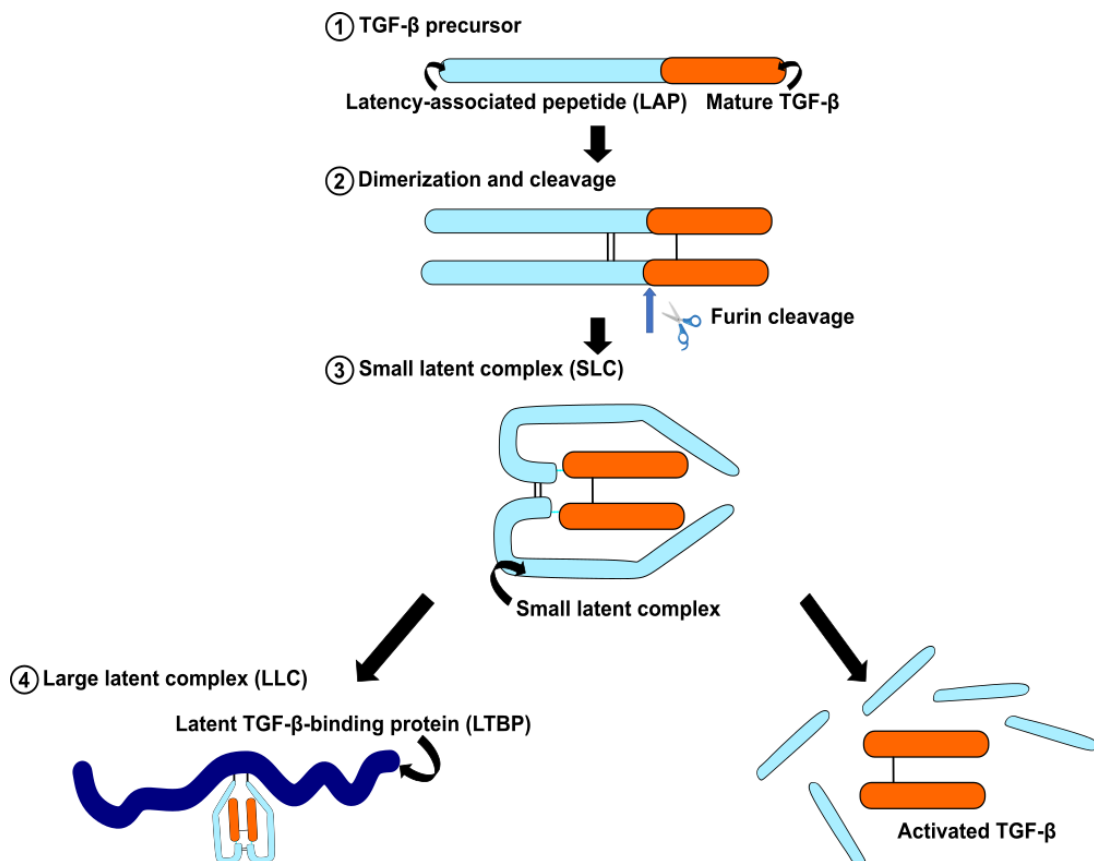
5.1.2 TGF- β pathway

In humans, the TGF- β superfamily of cytokines consists of 33 members, which includes three TGF- β isoforms in the TGF- β subfamily (TGF- β 1, TGF- β 2 and TGF- β 3), bone morphogenetic proteins (BMPs) and growth differentiation factors (GDFs) (reviewed by [290]). An interesting duality exists where TGF- β acts as both a tumour suppressor and tumour promoter during various stages of cancer development and metastasis. This is often referred to as the “TGF- β paradox” (reviewed by [291] and [292]). TGF- β inhibits proliferation and can induce apoptosis of premalignant cells under normal physiological conditions.

However, in the later stages of malignancy, TGF- β can induce EMT and facilitate cell migration and invasion (reviewed by [285]). This switch in late-stage cancer is thought to be due to a mixture of genetic, epigenetic and tumour microenvironment changes [291]. TGF- β 1 is the most abundant, prototypical isoform of the TGF- β family. It is expressed ubiquitously in human tissues and is typically upregulated to a greater degree than the other isoforms in cancer [293, 294].

TGF- β isoforms are secreted as inactive, small latent complexes (SLC), comprising a TGF- β dimer of inactive precursors bound to the latency-associated peptide (LAP). TGF- β dimers are cleaved by furin from their pro-peptides in the Golgi apparatus however, they remain associated with the pro-peptide via noncovalent interactions creating the SLC. The SLC can also be found bound to the latent TGF- β -binding protein (LTBP) to form the so-called large latent complex (LLC), see Figure 5.1 (reviewed by [294]). TGF- β can be activated from these latent complexes by a variety of factors including MMP-2 and MMP-9 [295], thrombospondin-1 [296], integrins α V β 6 and α V β 8 [297, 298], reactive oxygen species [299] or an acidic pH [300]. Active TGF- β can bind as a ligand with high affinity to one of the cell surface TGF- β receptors (T β R) [301].

Downstream signalling occurs when T β RII binds the TGF- β ligand, which catalyses recruitment and phosphorylation of T β RI. This results in the activation of the canonical Smad pathway, which in turn, results in the activation of receptor-regulated SMADs (R-SMADs), SMAD2 and SMAD3, which become phosphorylated and form a trimer with SMAD4, a common SMAD (Co-SMAD). This trimer is then translocated into the nucleus where it associates and cooperates with transcription factors to activate or repress gene transcription, see Figure 5.2. SMAD7, an inhibitory SMAD (I-SMAD), acts to inhibit the activation of the receptor-regulated SMADs in this pathway [285]. TGF- β can also exert effects through other noncanonical pathways that are typically activated by other tyrosine kinase receptors in response to their respective ligands (reviewed by [302]).



- ①** TGF-β precursor synthesised in the endoplasmic reticulum.
- ②** TGF-β precursor dimerizes and is then cleaved by furin, a convertase.
- ③** This forms the small latent complex (SLC) in the Golgi apparatus.
- ④** The SLC can bind to the latent TGF-β-binding protein (LTBP) to form a large latent complex (LCC). TGF-β can be activated from the SLC and LLC by a variety of factors including MMP-2, MMP-9, thrombospondin-1, integrins, ROS or an acidic pH.

Figure 5.1 Transforming growth factor-β (TGF-β) latent complex formation and activation. TGF-β precursor is synthesised in the endoplasmic reticulum, dimerizes and is then cleaved by furin, a convertase, to form the small latent complex (SLC) in the Golgi apparatus. The SLC can bind to the latent TGF-β-binding protein (LTBP) to form a large latent complex (LCC). TGF-β can be activated from the SLC and LCC by a variety of factors including matrix metalloproteinase-2 (MMP-2), matrix metalloproteinase-9 (MMP-9), thrombospondin-1, integrins, reactive oxygen species (ROS) or an acidic pH. Adapted from [303].

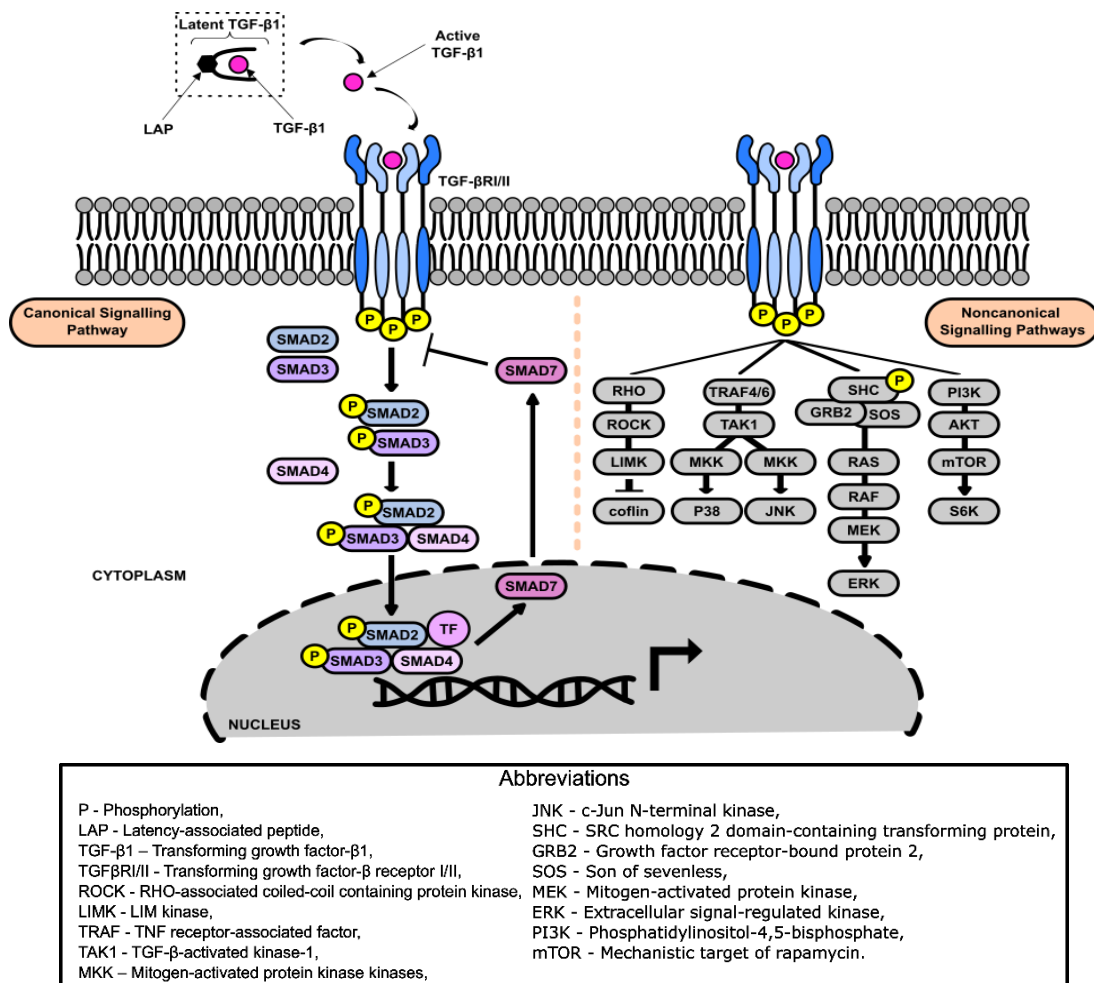


Figure 5.2 Overview of transforming growth factor-β1 (TGF-β1) canonical and noncanonical signalling pathways. Latent TGF-β1 can exist as a small latent complex, consisting of latency-associated peptide (LAP) and TGF-β1 which are connected by non-covalent bonds. Active TGF-β1 can be released from the small latent complex by proteases, for example, matrix metalloproteinases (MMPs) or by thrombospondin-1. TGF-β1 typically exerts its effect through the canonical SMAD pathway. It can also exert effects through other noncanonical pathways, such as RHO GTPases, P38, JNK, ERK and PI3K-AKT. Adapted from [304].

It is well-accepted that breast epithelial cell cultures can be treated with TGF-β1 to reliably induce EMT and EMT functionality [305]. This includes the cell lines MCF-7 [306-308] and T-47D [307]. Interestingly, it has been reported that in HMLE breast epithelial cells, prolonged exposure to TGF-β1 (treatment for 24 days) results in cells stably residing in the mesenchymal phenotype even after the TGF-β1 has been removed [309]. This evidence suggests that TGF-β1 treatment of MCF-7 and T-47D cell lines allows for the induction of EMT that is detectable by changes in morphology, protein markers and functionality.

5.1.3 Glycosylation changes in cytokine-induced EMT

Literature investigating changes in glycosylation as a result of TGF- β treatment in the context of cancer metastasis is limited. However, Kurcon et al., [310] demonstrated that TGF- β 1 treatment (5ng/mL for 48 hours) of A549 lung cancer cells resulted in significant increases in B3GLCT (transfers Glc to Fuc) (3.5-fold) and ST3GAL5 (transfers NeuNAc to Gal) (1.5-fold) enzyme mRNA, detected by qRT-PCR, compared to untreated cells. These results were also confirmed at the protein level using Western Blotting. Induction of EMT was confirmed by morphology assessments, and EMT marker Western Blotting and qRT-PCR. However, potentially altered glycans produced as a result of the enzyme increases were not investigated.

Du et al., [311] visualized the dynamics of sialylated glycans over 84 hours after TGF- β 1 treatment in human keratinocyte, HaCaT cells. EMT was confirmed using morphology assessments and increased N-cadherin protein levels (assessed by Western Blot) and fibronectin and MMP14 mRNA expression (assessed by qRT-PCR). Sialylated glycans were metabolically labelled with an azide functional analogue of sialic acid precursor N-acetylmannosamine (ManNAc), allowing the identification of newly synthesised glycans that are sialylated as a result of EMT. Cells were preincubated with the metabolically labelled ManNAc for 48 hours and then TGF- β 1 treatment was conducted for up to 84 hours. Results showed a bi-phasic change in cell surface sialylation. At 24 hours, when EMT was being induced, there was a 30% reduction in sialic acid, then in the timepoints up to 84 hours when cells were in a mesenchymal state, there was an upregulation of sialylation of cell surface proteins, assessed by flow cytometry and confocal microscopy. Proteomic analysis of TGF- β 1 treated metabolically labelled cells identified sialylglycoproteins that were more abundant compared to the control. Gene ontology using the Database for Annotation, Visualisation and Integrated Discovery (DAVID) revealed that the biological process categories of cell adhesion and cell motion were enriched, which indicates sialylation after TGF- β 1 induced EMT could have an important functional role in EMT. Cumulatively, evidence suggests that TGF- β treatment can have an impact on the cell surface glycosylation profile of a cell.

5.2 Aims and Objectives

To further elucidate if glycosylation changes in EMT, TGF- β 1 cytokine treatment of MCF-7 and T-47D cell lines was utilised to force a cytokine-induced EMT state. The experiments described in this chapter aimed to investigate if TGF- β 1-induced EMT states in the MCF-7

and T-47D cell lines led to any alteration in the abundance of EMT protein markers and O-GalNAc glycans.

The objectives of this chapter are to:

- 1) Confirm that TGF- β 1 can induce EMT in MCF-7 and T-47D cell lines.
- 2) Determine if TGF- β 1 induced EMT results in morphology changes associated with a mesenchymal-like phenotype in model breast epithelial cell lines, MCF-7 and T-47D.
- 3) Characterise model breast epithelial cell lines, MCF-7 and T-47D, for their levels of protein EMT markers, E-cadherin (epithelial marker), N-cadherin and vimentin (mesenchymal markers), after TGF- β 1 treatment.
- 4) Characterise the model breast epithelial cell lines, MCF-7 and T-47D, for their levels of O-GalNAc glycans after TGF- β 1 treatment.

5.3 Methods

5.3.1 TGF- β 1 treatment

MCF-7 and T-47D cell lines were induced into EMT using a TGF- β 1 cytokine treatment as outlined in '2.8.1 TGF- β 1 treatment' and '2.8.2 TGF- β 1 treatment and Western Blotting'.

5.3.2 Cell morphology analysis

Morphological analysis of cells was conducted using the phase contrast images of TGF- β 1 treated and control conditions as outlined in '2.8.3 Identifying effects of TGF- β 1 treatment on cell morphology, HPA labelling and EMT markers'. This was performed at the 24-, 48-, 72-hour and 7-day timepoints for MCF-7 cells and the 72-hour and 7-day timepoints for T-47D cells, since the T-47D cell line did not demonstrate visible morphological changes at the 8-, 24- or 48-hour timepoint, see Supplementary Figure 18.

5.3.3 EMT marker analysis

TGF- β 1 treated and control cells were assessed for their protein levels of key EMT markers, E-cadherin (epithelial marker), N-cadherin and vimentin (mesenchymal markers). This was achieved using immunofluorescence and Western Blotting as outlined in '2.8.1 TGF- β 1 treatment', '2.8.2 TGF- β 1 treatment and Western Blotting' and '2.8.3 Identifying effects of TGF- β 1 treatment on cell morphology, HPA labelling and EMT markers'. Western Blot analysis was conducted at the 72-hour timepoint for both MCF-7 and T-47D cells. Similarly, to the morphology assessments T-47D cells did not demonstrate visible changes in EMT markers E-cadherin or vimentin immunofluorescence during optimisation at the 8- and 24-hour timepoints, see Supplementary Figure 19A and B. Therefore, T-47D cells were investigated at 72 hours and 7 days to mirror morphology assessments. The MCF-7 cell line was investigated at 24-, 48-, 72-hour and 7-day timepoints.

5.3.4 Assessment of cellular O-GalNAc glycans following TGF- β 1 treatment

Cellular O-GalNAc glycans were assessed through fluorescent HPA lectin labelling using methodology described in '2.2.4 Fluorescent lectin labelling' on TGF- β 1 treated and control cells. The T-47D cell line did not show any visible changes in O-GalNAc glycan labelling in optimisation experiments at the 8 and 24-hour timepoint, see Supplementary Figure 19C. This outcome, coupled with the previous reasoning from cell morphology analysis and EMT markers is why they were only investigated at the 72-hour and 7-day timepoint.

5.4 Results

5.4.1 TGF- β 1 treatment and MCF-7 and T-47D cell morphology

MCF-7 cells acquired an elongated, mesenchymal-like morphology after 24 hours of TGF- β 1 treatment, see Figure 5.3. Control cells retained their cobblestone-like, rounded epithelial morphology throughout all timepoints investigated. TGF- β 1 treated MCF-7 cells were significantly more elongated than control cells across all timepoints, assessed by measurement of aspect ratio. Across all timepoints, TGF- β 1 treated MCF-7 cells demonstrated an aspect ratio score that breached the >2:1 aspect ratio, indicating an elongated morphology, see Figure 5.4A. When circularity was measured, control cells demonstrated a significantly more rounded morphology compared to TGF- β 1 treated cells, see Figure 5.4B.

The changes in the morphology of T-47D cells were less visibly clear-cut compared to those of the MCF-7 cells. However, TGF- β 1 treated T-47D cells still became visibly more elongated than control at the 72-hour and 7-day timepoints, see Figure 5.5. The aspect ratio of TGF- β 1 treated T-47D cells were significantly more elongated than control cells, >2:1, at the 72-hour and 7-day timepoints investigated, see Figure 5.6A, indicating an elongated phenotype. When circularity was measured, control cells were also significantly more rounded than TGF- β 1 treated T-47D cells, see Figure 5.6B.

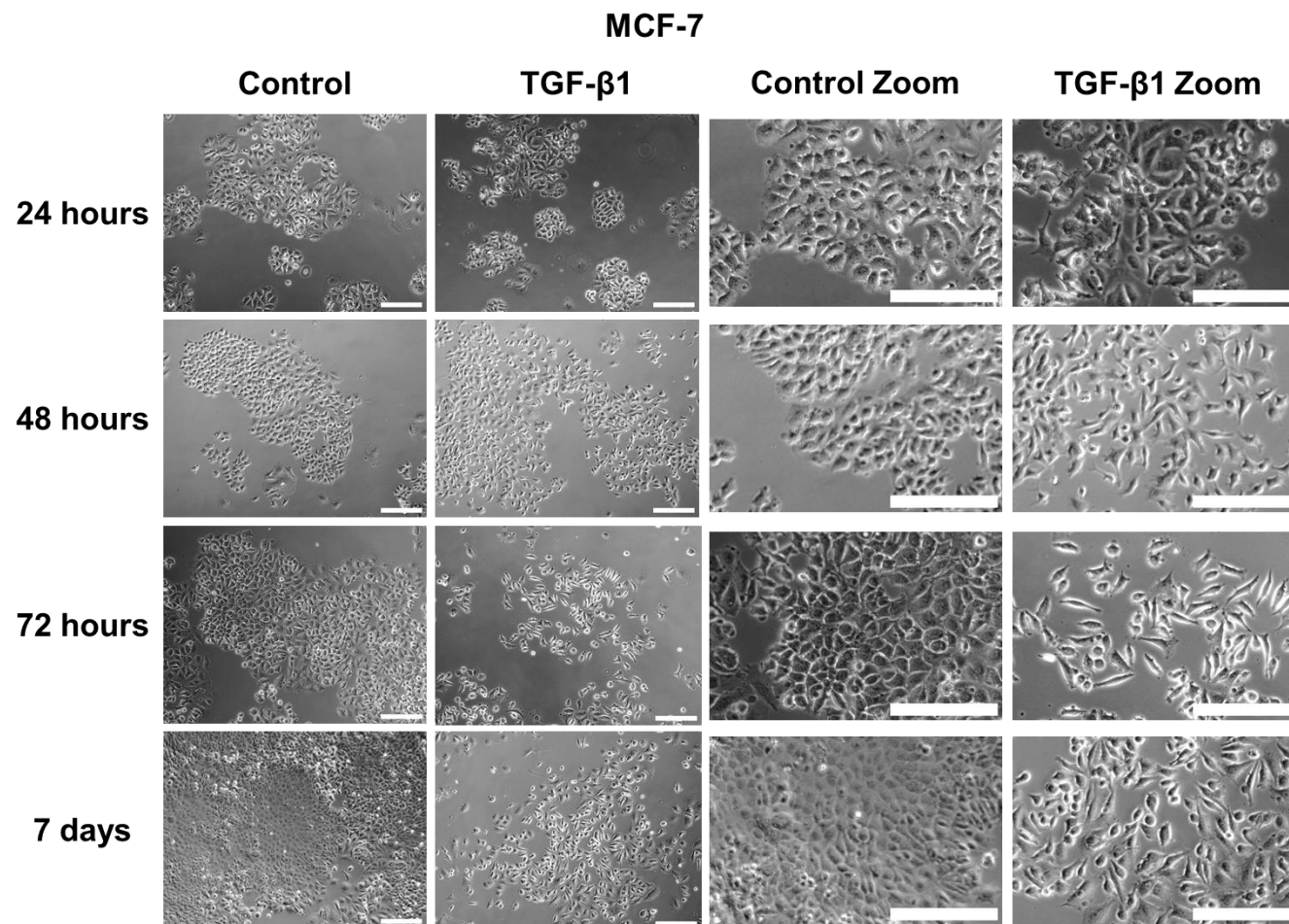


Figure 5.3 Transforming growth factor- β 1 (TGF- β 1) treatment effect on MCF-7 morphology. Representative phase contrast images of control and TGF- β 1 treated MCF-7 cells at 24, 48, 72 hours and 7 days. Scale bar = 50 μ m.

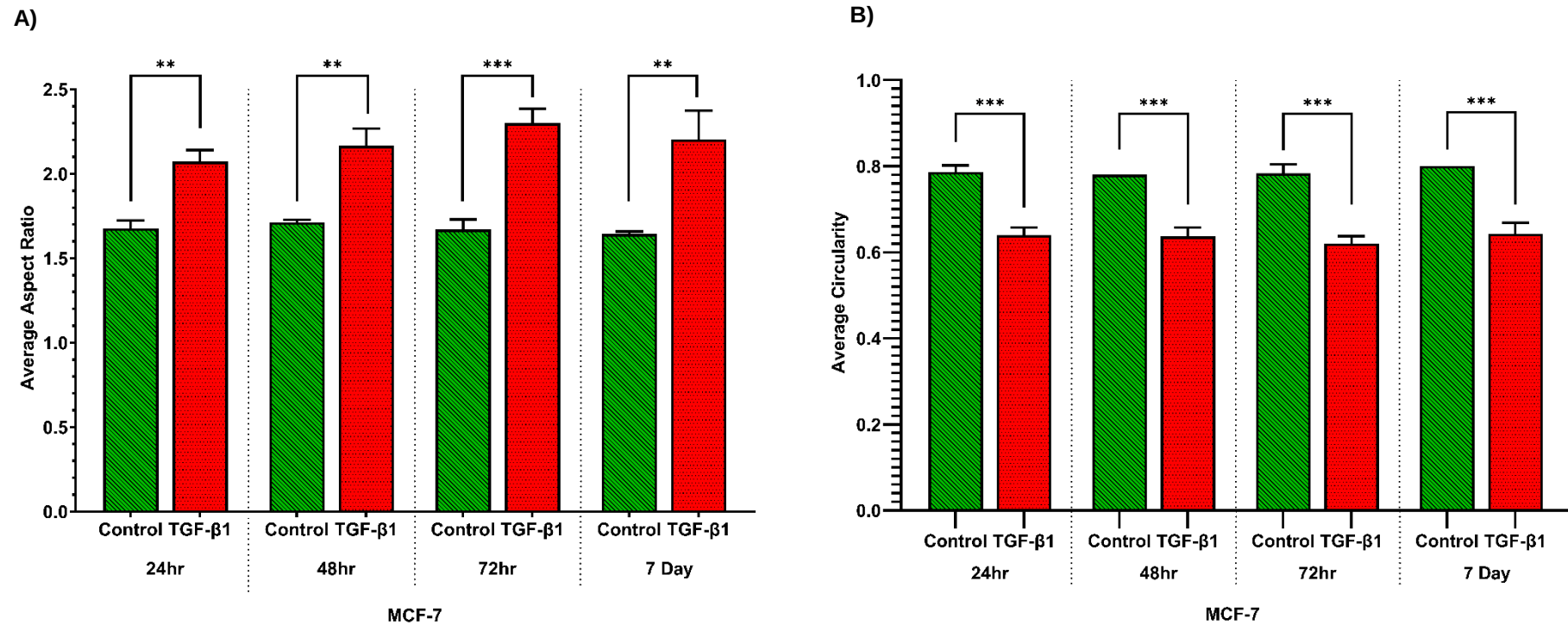


Figure 5.4 Transforming growth factor- β 1 (TGF- β 1) treatment morphology assessment in MCF-7 cells. Each biological replicate value consists of an average of three images, and up to 300 cells per image were quantified, three biological replicates were performed. Unpaired Students t-test was used to test for significance between cell line conditions, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. A) Average aspect ratio ($\pm$ SD). B) Average circularity ($\pm$ SD).

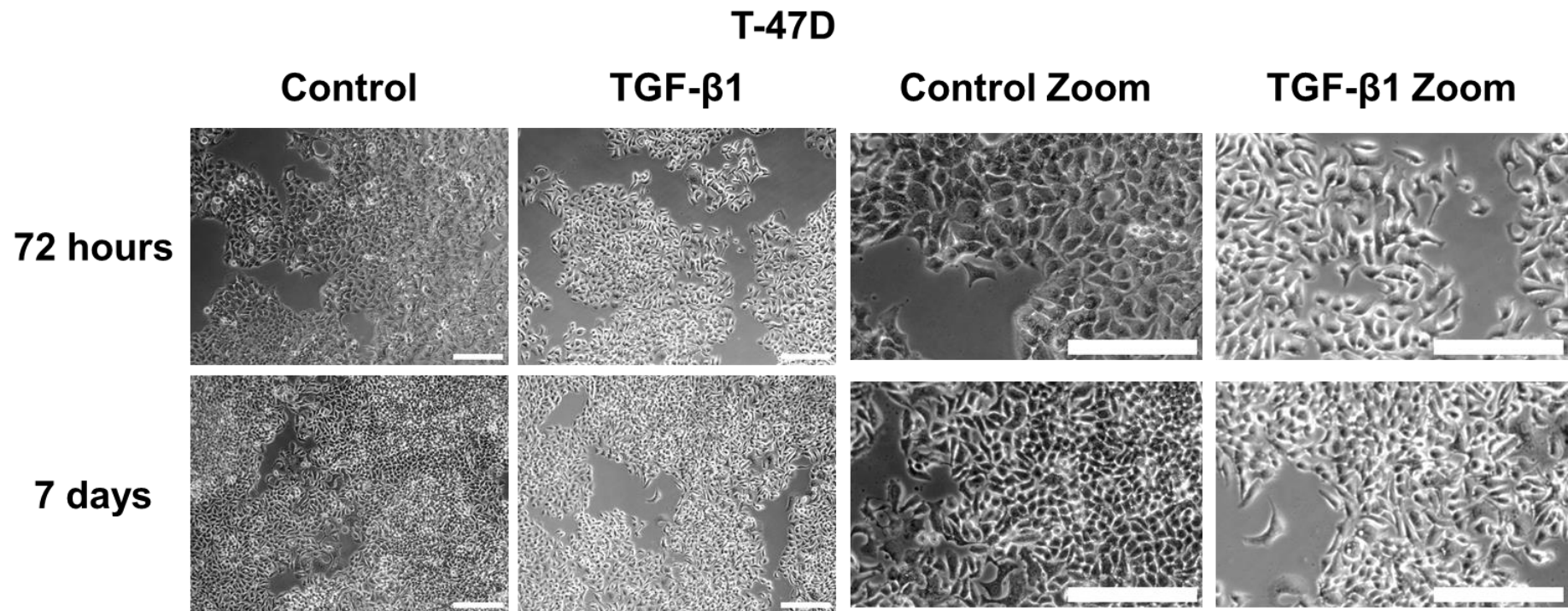


Figure 5.5 Transforming growth factor- β 1 (TGF- β 1) treatment effect on T-47D morphology. Representative phase contrast images of control and TGF- β 1 treated T-47D cells at 72 hours and 7 days. Scale bar = 50 μ m.

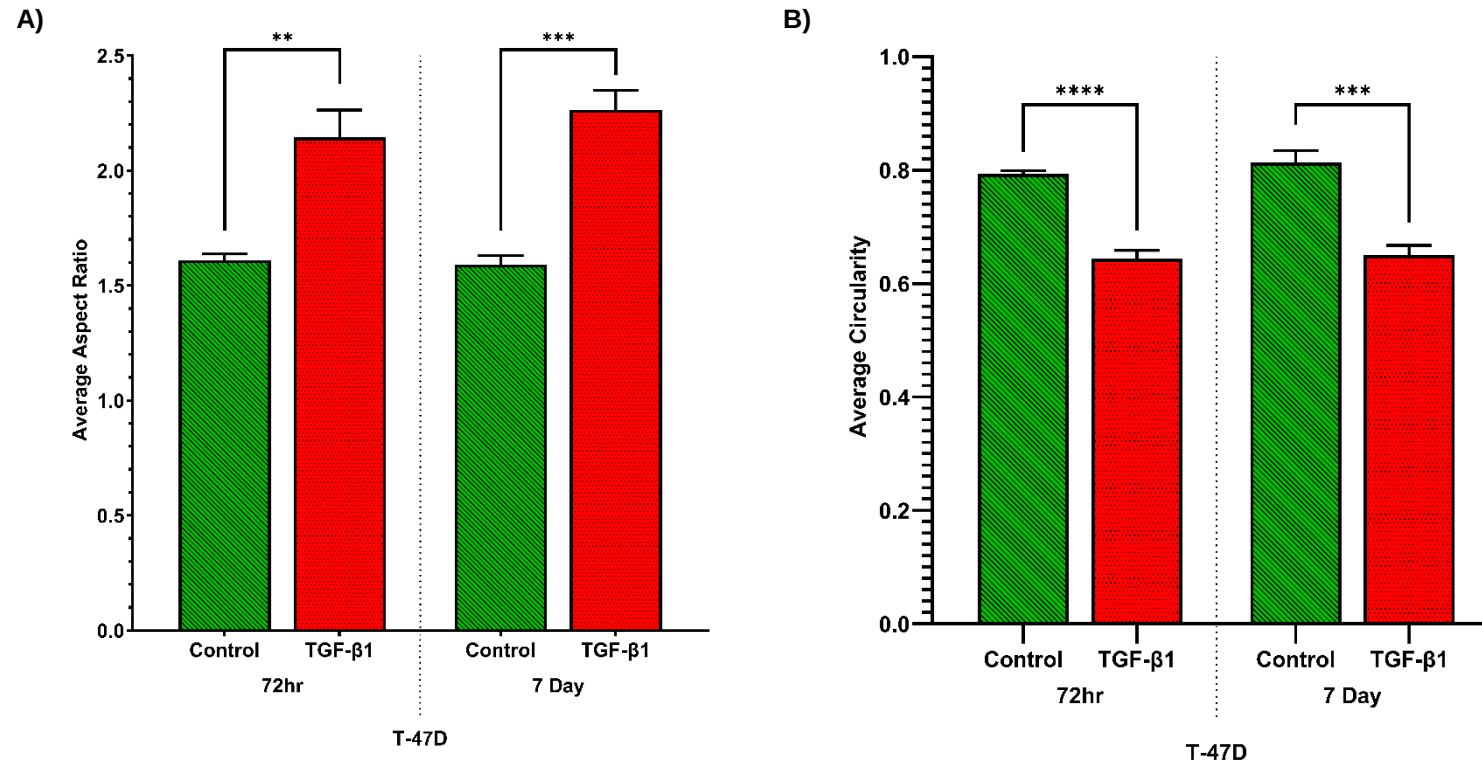


Figure 5.6 Transforming growth factor-β1 (TGF-β1) treatment morphology assessment in T-47D cells. Each biological replicate value consists of an average of three images, and up to 300 cells per image were quantified, three biological replicates were performed. Unpaired Students t-test was used to test for significance between cell line conditions, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. A) Average aspect ratio ($\pm$ SD). B) Average circularity ($\pm$ SD).

5.4.2 TGF- β 1 treatment and EMT marker protein expression

Both MCF-7 and T-47D cell line controls demonstrated strong labelling for E-cadherin across all timepoints investigated, see Figure 5.7A and B respectively. TGF- β 1 treated MCF-7 cells demonstrated visible, and quantifiably significantly lower E-cadherin levels than control, across all timepoints, see Figure 5.8A. TGF- β 1 treated T-47D cells demonstrated a visible reduction in E-cadherin protein labelling at both 72 hours and 7 days. However, this difference was not significant ($p=0.056$ at 7 days), see Figure 5.8B. Interestingly, neither N-cadherin nor vimentin protein levels were increased by TGF- β 1 treatment in either the MCF-7 or T-47D cell lines, see Figure 5.9 and 5.10.

EMT marker immunofluorescence results were confirmed by Western Blot analysis. Results showed a clear visible reduction in E-cadherin protein levels for both MCF-7 and T-47D cell lines treated with TGF- β 1 compared to control, see Figure 5.11. Western Blots also confirm that there was no N-cadherin or vimentin detectable in either the control or TGF- β 1 treated cells. Relative protein expression of EMT marker E-cadherin, normalised to β -actin, in MCF-7 TGF- β 1 treated cells, was significantly lower compared to the control ($p<0.05$). There was a visible, but non-significant reduction, of E-cadherin in TGF- β 1 treated T-47D cells compared to control, mirroring the results of the immunofluorescence.

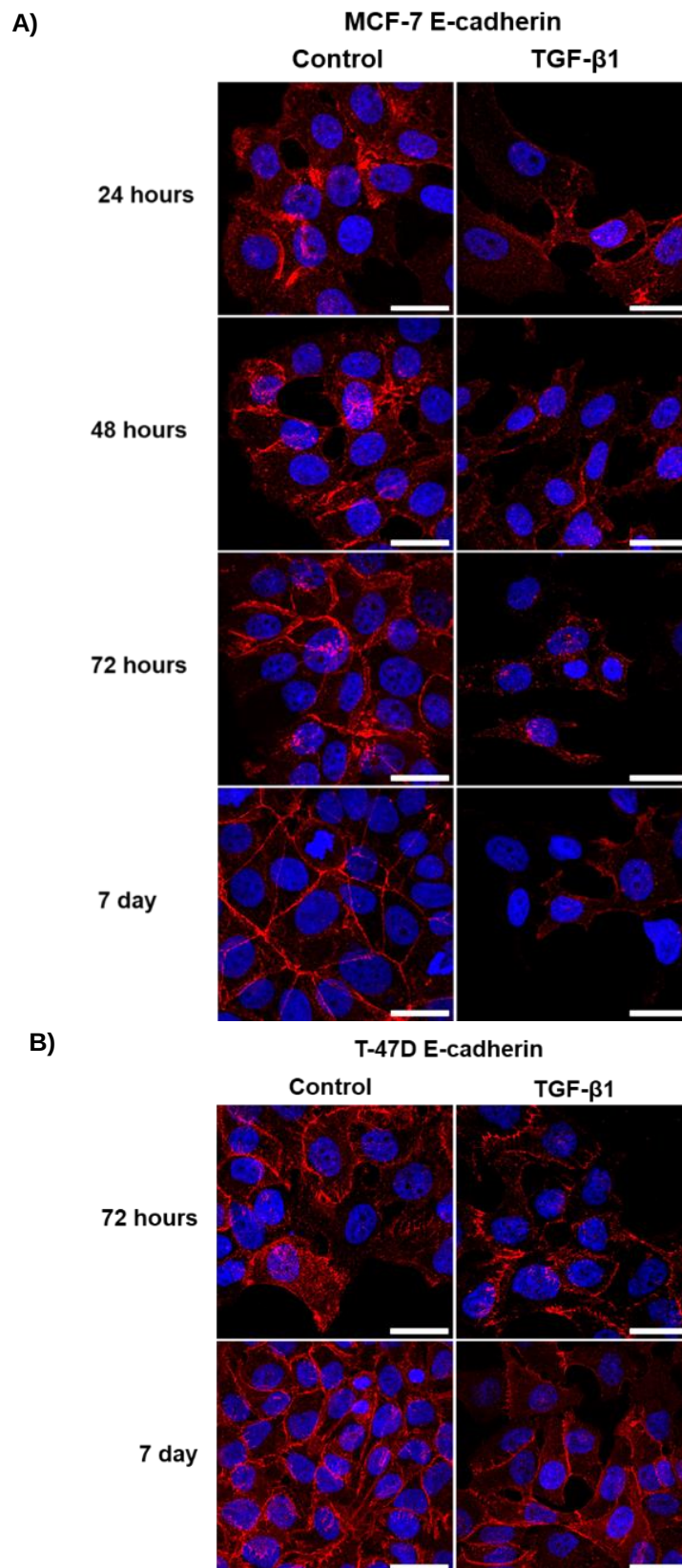
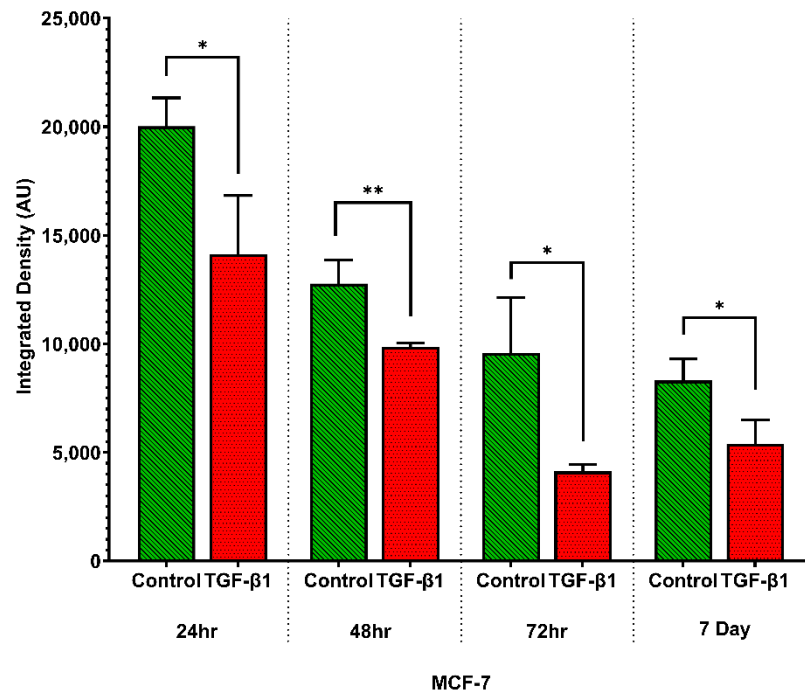


Figure 5.7 Transforming growth factor- β 1 (TGF- β 1) treatment and E-cadherin immunofluorescence. Representative Max Intensity Z-stack projections of E-cadherin immunofluorescence in A) MCF-7 cell line, B) T-47D cell line. E-cadherin (red) and DAPI (blue). Scale bar = 25 μ m.

A)



B)

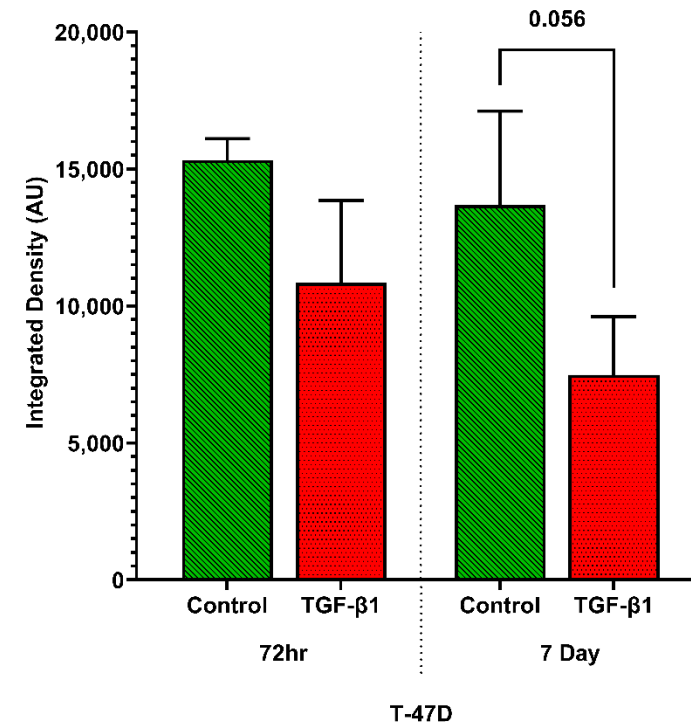


Figure 5.8 E-cadherin immunofluorescence after transforming growth factor- β 1 (TGF- β 1) treatment. Average integrated density (AU) of E-cadherin labelling per cell of A) MCF-7 cells, B) T-47D cells. Three biological replicates ($\pm$ SD). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$.

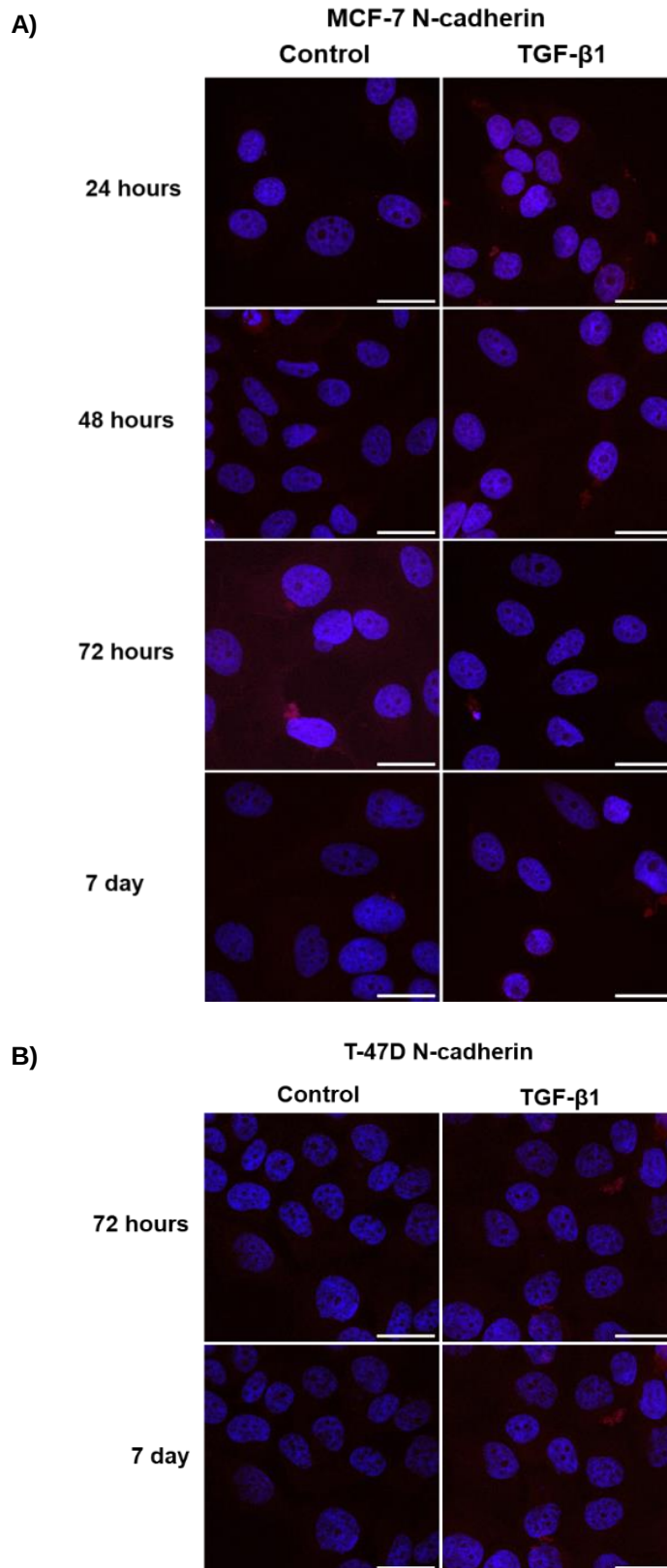


Figure 5.9 Transforming growth factor- β 1 (TGF- β 1) treatment and N-cadherin immunofluorescence. Representative Max Intensity Z-stack projections of N-cadherin immunofluorescence in A) MCF-7 cell line, B) T-47D cell line. N-cadherin (red) and DAPI (blue). Scale bar = 25 μ m.

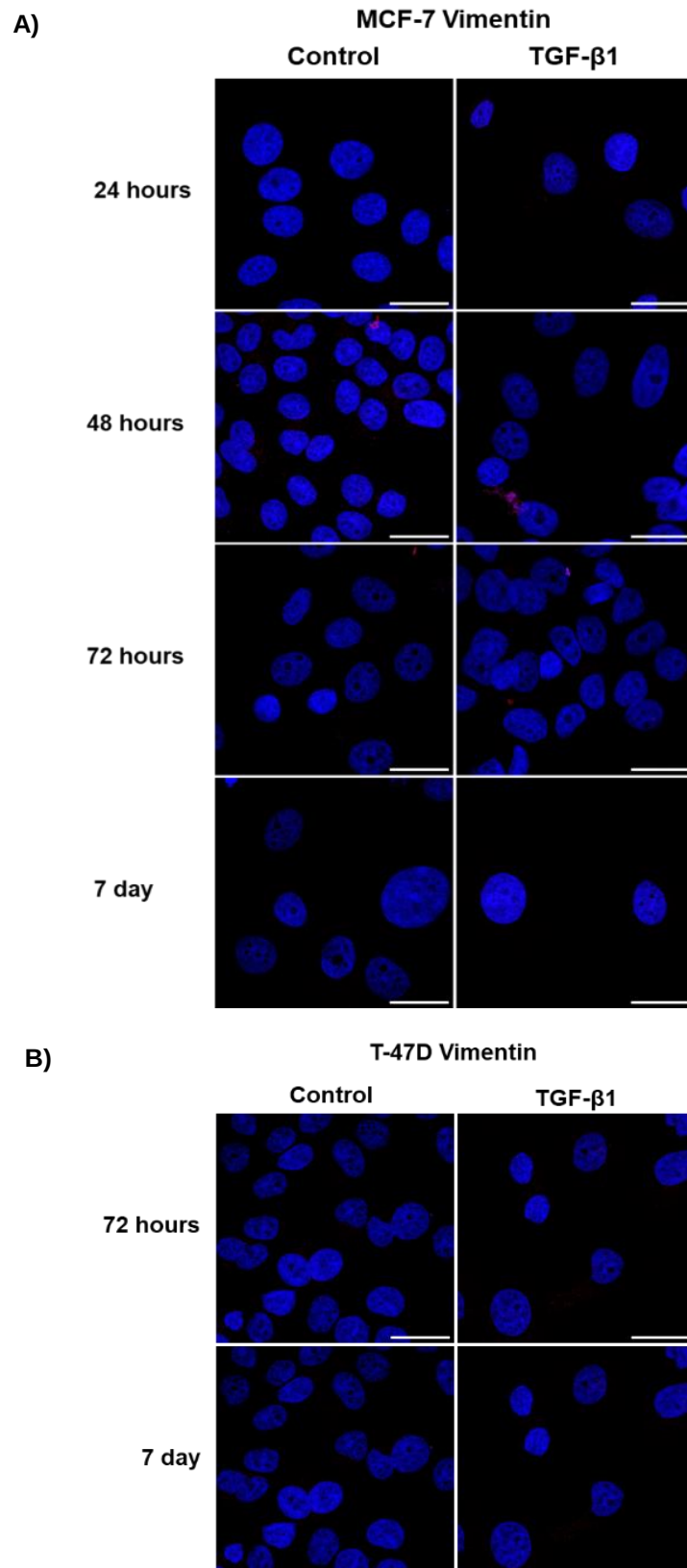


Figure 5.10 Transforming growth factor- β 1 (TGF- β 1) treatment and vimentin immunofluorescence. Representative Max Intensity Z-stack projections of vimentin immunofluorescence in A) MCF-7 cell line, B) T-47D cell line. Vimentin (red) and DAPI (blue). Scale bar = 25 μ m.

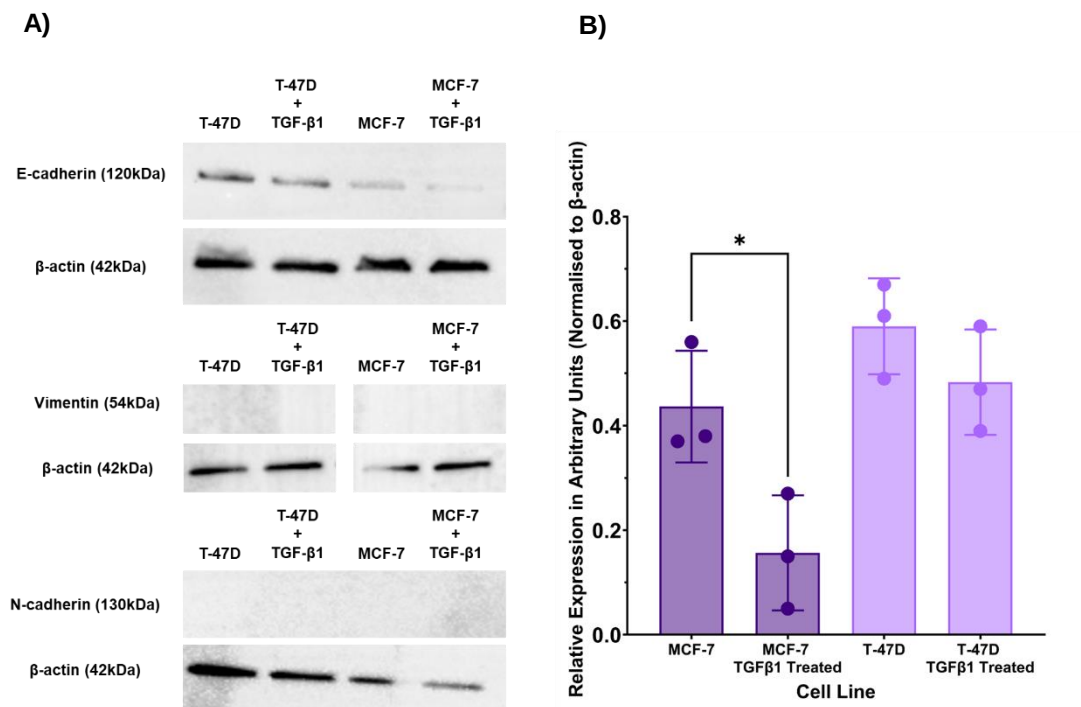


Figure 5.11 Transforming growth factor-β1 (TGF-β1) treatment and Western Blotting for epithelial-mesenchymal transition (EMT) markers. A) Representative Western Blots for E-cadherin (120kDa), vimentin (54kDa) and N-cadherin (130kDa) and loading control β-actin (42kDa) for control (untreated) and TGF-β1 treated T-47D and MCF-7 cells at 72 hours. B) Relative protein levels of EMT marker E-cadherin normalised to β-actin in control (untreated) and TGF-β1 treated MCF-7 and T-47D cells, displayed as average of three biological replicates (±SD). Unpaired Students t-test was used to test for significance, * p<0.05.

5.4.3 TGF-β1 treatment and O-GalNAc glycans

TGF-β1 treatment of both MCF-7 and T-47D cells had a visible impact on the levels of cell surface O-GalNAc glycans, detected by HPA labelling, see Figure 5.12A and B respectively. At all timepoints, 24, 48, 72 hours and 7 days, MCF-7 cells showed visibly lower O-GalNAc glycan labelling, as detected by HPA binding, compared to controls. The corresponding changes in cell morphology, to an elongated mesenchymal-like morphology were also visible in HPA-negative cells. TGF-β1 treated MCF-7 cells showed a visible but, not significant reduction in O-GalNAc glycans at 24, 48 and 72 hours compared to the control. There was a significant difference in O-GalNAc glycans at 7 days of TGF-β1 treatment compared to control in MCF-7 cells, see Figure 5.13A. It is interesting to note that the entire cell population did not become HPA-negative upon TGF-β1 cytokine treatment, a small population of cells or individual cells retained cell surface O-GalNAc glycans.

At both timepoints, 72 hours and 7 days, TGF- β 1 treated T-47D cells also showed visibly less O-GalNAc glycans compared to the control. These changes were associated with a visible mesenchymal-like phenotype. TGF- β 1 treated T-47D cells showed a visible but not significant reduction in O-GalNAc glycans at 72 hours. There was a significant difference in O-GalNAc glycans levels at 7 days of TGF- β 1 treatment compared to control, see Figure 5.13B.

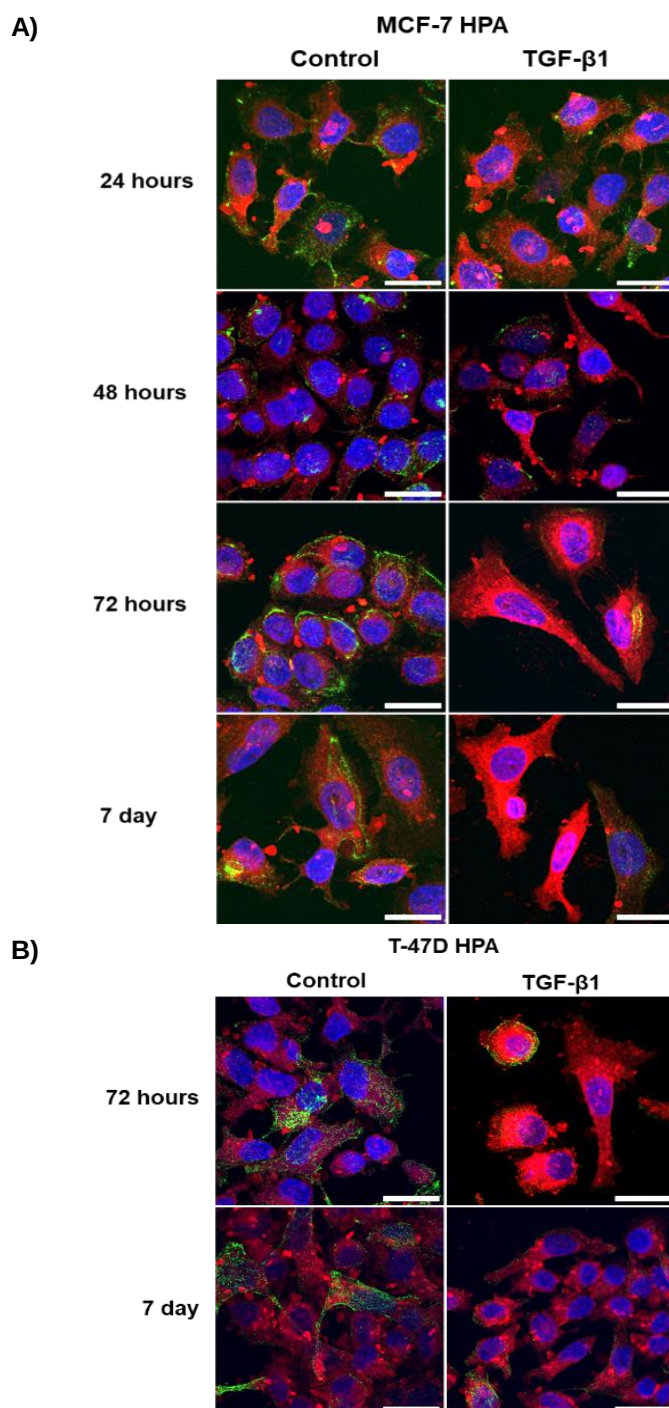
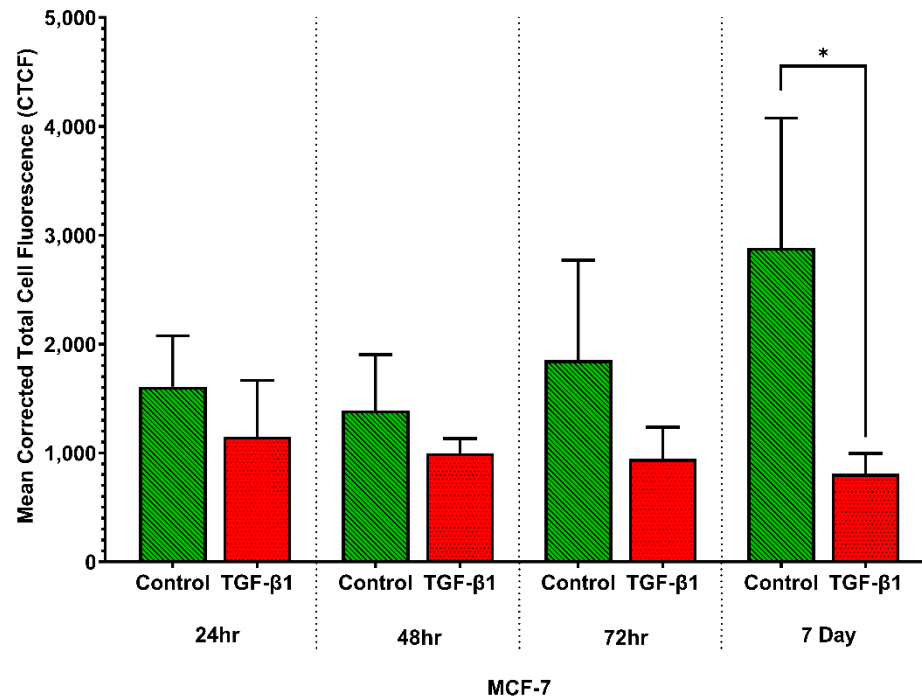


Figure 5.12 Transforming growth factor- β 1 (TGF- β 1) treatment and fluorescent *Helix pomatia* agglutinin (HPA) lectin labelling. Representative Max Intensity Z-stack projections of fluorescent HPA lectin labelling in A) MCF-7 cell line, B) T-47D cell line. HPA (green), CellMask™ (red) and DAPI (blue). Scale bar = 25 μ m.

A)



B)

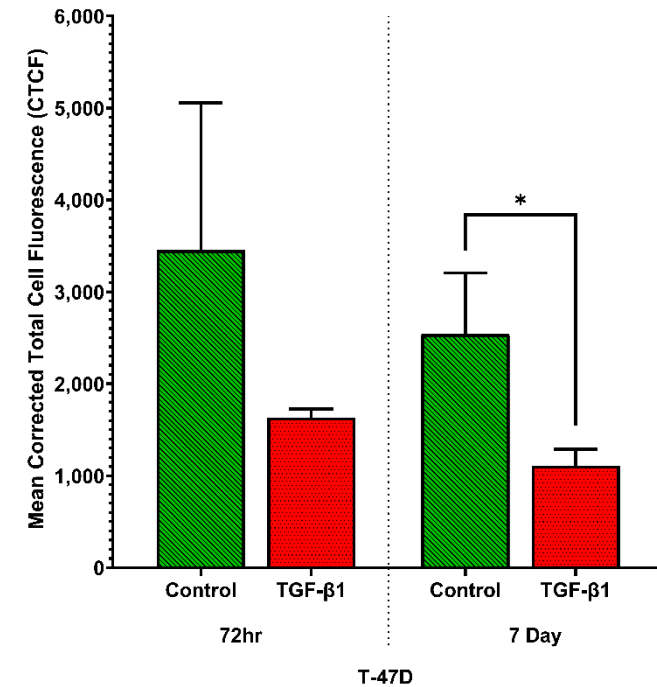


Figure 5.13 Mean corrected total cell fluorescence (CTCF) of *Helix pomatia* agglutinin (HPA) labelling after transforming growth factor-β1 (TGF-β1) treatment. CTCF of HPA labelling in A) MCF-7 control and TGF-β1 treated cells, B) T-47D control and TGF-β1 treated cells. Three biological replicates (±SD). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$.

5.5 Discussion

5.5.1 TGF- β 1 treatment induces an elongated mesenchymal-like cell morphology

To confirm that TGF- β 1 treatment is capable of inducing EMT in MCF-7 and T-47D cells it was necessary to identify a change in cell morphology, from a rounded-epithelial phenotype to an elongated-mesenchymal phenotype [223]. This was assessed using the methods previously described in Chapter 3, taking measurements of aspect ratio (length/width), which indicates cell elongation, and circularity ($4\pi \times (\text{area}) / (\text{perimeter})^2$) using ImageJ [171]. Both cell lines gained a mesenchymal-like morphology following TGF- β 1 treatment, across each timepoint investigated. In both cell lines, TGF- β 1 treated cells were significantly more elongated, breaching the $>2:1$ aspect ratio, which indicates an elongated phenotype, and were less circular than the control, untreated, cells. These results are indicative of the acquisition of an EMT phenotype. These morphological changes are consistent with other studies that have also demonstrated that TGF- β 1 treatment, at varying concentrations and exposure timepoints, induces cell morphology changes in the cell lines MCF-7 [307, 312, 313] and T-47D [307].

5.5.2 TGF- β 1 treatment affects EMT marker, E-cadherin, protein levels

In addition to morphological changes, it was also necessary to assess protein levels of key EMT markers, such as the epithelial marker, E-cadherin, which is often referred to as the “gatekeeper of the epithelial state” [121] and mesenchymal markers such as N-cadherin and vimentin. TGF- β 1 treatment significantly reduced E-cadherin protein levels, assessed by immunolabelling, across all timepoints (24, 48, 72 hours and 7 days) in MCF-7 cells. E-cadherin protein immunolabelling was decreased visibly at 72 hours and this reached near significance at 7 days ($p=0.056$) in T-47D cells following TGF- β 1 treatment. There was no visible induction of N-cadherin or vimentin protein expression. These results are consistent with those reported in Chapters 3 and 4 that demonstrated that O-GalNAc glycan-negative cells were elongated, more motile and displayed significantly less E-cadherin protein but, did not feature detectable levels of N-cadherin or vimentin protein. E-cadherin is an important tumour suppressor gene, and reduced expression or loss of E-cadherin is regarded as the major mechanism of triggering the loss of cell-cell adhesion and therefore initiating and promoting cell migration and metastasis [314].

Loss of E-cadherin via post-translational downregulation has also been described as a marker of 'early EMT' [315, 316] whereas, N-cadherin expression has been described as a marker of a 'late EMT' state [317]. These results could therefore be indicative of an 'early EMT' state, or a pEMT state because cells have a significant reduction in epithelial marker E-cadherin but, have not gained mesenchymal markers. pEMT states are common as cancer cells often neither fully lose their epithelial markers nor fully gain mesenchymal markers, existing in a spectrum of hybrid protein expression states (reviewed by [123]).

It is interesting to note that TGF- β 1 treated MCF-7 and T-47D cells do not fully lose E-cadherin protein, despite the large reduction compared to control conditions. This could also add to evidence supporting a pEMT state. Aban et al., [316] demonstrated in human embryonic stem cells that partial downregulation of E-cadherin, achieved by CRISPR, resulted in a pEMT state whereby cells demonstrated a mesenchymal-like phenotype. It also stimulated collective cell migration and increased 'wound' closure in a wound healing assay. The partial loss of E-cadherin could act as the triggering signal at the start of EMT.

It is also rare *in vivo* that tumour cells commit to a full-EMT state and tumour cells that are in a pEMT state often do not have high expression levels of EMT transcription factors (reviewed by [318]). Puram et al., [319] demonstrated using *in situ* spatial localisation of cells in head and neck squamous cell carcinoma tissue samples, that cells demonstrating pEMT were localised to the 'leading edge' of primary tumours. Identified by labelling for a variety of pEMT genes like TGFBI. Cells were then sorted using fluorescence-activated cell sorting (FACS) into pEMT^{high} and pEMT^{low} based on TGFBI immunolabelling. Interestingly, some gene expression of epithelial markers was generally maintained despite an overall reduction of epithelial gene expression in pEMT^{high} cells, indicating that cells at the invasive, migratory edge of a tumour do not necessarily completely lose epithelial markers. Cheung et al., [320] demonstrated collective migration in 3D primary breast carcinoma organoids, and cells at the invasive edges still retained a level of E-cadherin expression which was postulated to maintain the cell-cell adhesion with follower cells. Cells at the 'leading edge' were also molecularly distinct and did not express common EMT markers vimentin, TWIST or Slug.

Results presented in this chapter are not fully consistent with all the reported literature, however. Mahdi et al., [307], who, as previously described, reported morphology changes after TGF- β 1 treatment similar to those reported here, also found that not only was E-cadherin significantly reduced following TGF- β 1 treatment but, vimentin was also significantly increased in both MCF-7 and T-47D cells at both the protein and mRNA level, as detected by Western Blotting and qRT-PCR. High levels of protein and mRNA expression were seen even in controls. Results presented in this thesis show that neither N-cadherin nor vimentin protein was detectable, even in the controls. Other literature, however, suggests that MCF-7 cells do not express N-cadherin proteins [321], nor do T-47D cells

[255, 256]. MCF-7 and T-47D cells have also consistently been documented in the literature as not expressing vimentin [248, 251, 253].

5.5.3 TGF- β 1 treatment reduces cell surface O-GalNAc glycans

MCF-7 and T-47D cells treated with TGF- β 1 had visibly reduced levels of cell surface O-GalNAc glycans, detected by HPA labelling, which also corresponded to changes in cell morphology, to an elongated mesenchymal-like morphology. Following TGF- β 1 treatment, both MCF-7 and T-47D cells demonstrated significantly less O-GalNAc glycans at the 7-day timepoint. Interestingly, in both cell lines, the whole cell population did not become HPA-negative upon TGF- β 1 treatment; instead, a small population of cells or individual cells retained cell surface O-GalNAc glycans. This is the first time that alterations in O-GalNAc glycans have been documented as changing in response to TGF- β 1 treatment in breast cancer cell lines.

Most literature focuses on the N-glycome and frequently describes changes in N-linked glycosylation and correlating glycosylation enzymes following TGF- β 1 treatment [311, 322]. N-linked glycosylation is often the focus of studies due to ease of sample preparation for analysis by mass spectrometry and because of the challenges associated with overlapping enzyme specificities and roles in other forms of glycosylation [310]. However, Yang et al., [323] utilised lectins, in an enzyme-linked lectin assay, to investigate the effect of TGF- β 1 (5ng/mL for 48 hours) and TNF- α (40ng/mL for 48 hours) treatment on bovine and human chondrocytes. TGF- β 1 treatment resulted in significant changes in bovine chondrocyte cell surface glycosylation but, not human chondrocytes. TGF- β 1 treatment in bovine chondrocytes resulted in a significant increase in binding of the Gal-binding PNA, and a significant decrease in the binding of sialyl α 2-3-binding *Maackia amurensis* agglutinin (MAA) and GlcNAc-binding wheat germ agglutinin (WGA) compared to untreated cells. Suggesting significantly more core 1 O-glycans, and reduced GlcNAc and sialyl α 2-3 glycans. TNF- α treatment of bovine chondrocytes resulted in a significant increase in GlcNAc-binding WGA, sialyl α 2-6-binding *Sambucus nigra* agglutinin (SNA), core 1-binding PNA, sialyl α 2-3-binding MAA, and mannose-binding ConA lectin binding compared to untreated cells, suggesting significantly more GlcNAc, core 1 glycans, sialyl α 2-3 and sialyl α 2-6 Gal glycans and mannose N-linked glycans. Human chondrocytes treated with TNF- α showed a significant increase in SNA lectin binding and a significant decrease in HPA lectin binding, which suggests significantly fewer O-GalNAc glycans and more sialyl α 2-6 glycans.

Lucena et al., [324] demonstrated that A459 lung cancer cells treated with TGF- β 1 (5ng/mL) for up to 48 hours, could be induced into EMT as detected by changes in morphology and altered EMT protein marker expression, reduced E-cadherin and increased vimentin. Using

a variety of lectins in flow cytometry they demonstrated a significant increase in labelling for SNA (indicating more sialyl α 2-6 glycans) and a visible, but non-significant, decrease in PNA (indicating decreased core 1 glycans) and VVA labelling (indicating decreased O-GalNAc glycans). Delmotte et al., [325] studied primary human bronchial mucosa cells, and post-exposure to TNF- α (20ng/mL for 4 hours) they saw an increase in α 2-3-sialyltransferase activity and mRNA expression of ST3GAL3 and ST3GAL4, indicating a potential increase in sialylated glycans. In cultured primary human endothelial cells, TNF- α treatment (200 units/mL for 72 hours) also resulted in increased sialyltransferase activity (3.5-fold), and increased binding of SNA (recognising α 2-6 linked sialyl glycans) to cell lysate proteins [326]. Cumulatively, this evidence suggests that EMT, induced by cytokines such as TNF- α and TGF- β , is linked to changes in cell surface glycosylation which is important in the functionality of the cell. Changes to sialylation of glycans were not investigated in the study reported here, and could provide a potential explanation for the reduction of O-GalNAc glycans seen following TGF- β 1 treatment in both MCF-7 and T-47D cells.

Results presented here suggest that TGF- β 1 treatment of both MCF-7 and T-47D cells induced morphological changes resulting in an elongated, mesenchymal-like phenotype. It also resulted in the reduction of epithelial marker E-cadherin protein but, did not induce mesenchymal markers N-cadherin or vimentin proteins. This could be indicative of an 'early EMT' state or a pEMT state, whereby the loss of E-cadherin is an initiating event or is sufficient for pEMT in these cells. TGF- β 1 treatment was also found to result in a significant reduction in the level of cell surface O-GalNAc glycans, detected by HPA binding, and so could allude to a potential role of altered O-glycosylation in early or pEMT states.

5.6 Key findings

- TGF- β 1 treatment of MCF-7 and T-47D cells results in altered cell morphology. Cells became significantly less rounded and significantly more elongated compared to control across all timepoints investigated per cell line.
- TGF- β 1 treatment of MCF-7 cells resulted in significantly less E-cadherin (epithelial marker) across all treatment timepoints, 24, 48, 72 hours and 7 days. This was confirmed in Western Blot analysis for the 72-hour timepoint investigated.
- TGF- β 1 treatment of T-47D cells resulted in a visible, but not statistically significant, reduction of E-cadherin (epithelial marker) across all treatment timepoints, 72 hours and 7 days ($p=0.056$ at 7 days). This was confirmed in Western Blot analysis for the 72-hour timepoint.

- TGF- β 1 treatment of MCF-7 and T-47D cells did not induce an increase in N-cadherin or vimentin (mesenchymal markers). This was confirmed in Western Blot analysis for the 72-hour timepoint.
- TGF- β 1 treatment of MCF-7 and T-47D cells resulted in a significant reduction of cell surface O-GalNAc glycans at the 7-day timepoint. This effect was also visible, but not statistically significant, at all preceding timepoints investigated for each cell line.

Chapter 6

RNA sequencing

6. RNA sequencing

6.1 Background

6.1.1 RNA sequencing

RNA sequencing utilises next-generation sequencing (NGS) and is a high-throughput methodology to quantify RNA molecules. It allows for higher reproducibility, coverage, and resolution of the transcriptome than other methods such as microarray-based approaches, enabling single-nucleotide resolution [327]. However, one caveat is that unless single-cell RNA sequencing is performed, conventional RNA sequencing will not capture transcriptomic differences of individual cells, which can be important when analysing a heterogeneous tumour cell population. It is also important to consider that RNA sequencing only identifies a 'snapshot' of time and space in a highly dynamic transcriptome of a cell (reviewed by [328]). As the work documented in this thesis is concerned with cell line populations, rather than individual cell differences from a tumour sample, RNA sequencing was used to investigate the transcriptomes of various cell populations of interest.

6.1.2 Glycosylation and RNA sequencing

Although the genome does not directly encode the glycome - the complete repertoire of cell glycans and glycoconjugates - the transcription of the ~700 glycosylation-related genes ('glycogenes'), which accounts for approximately 3-4% of the mammalian genome, can be interrogated to provide a useful insight into what is happening to the cellular glycosylation machinery (reviewed by [329]). Additional complexity exists when investigating glycome-related gene expression due to the relatively low abundance of transcripts, overlapping functionality and enzymatic redundancy of many glycosyltransferases. For example, it can be hard to discern individual contributions of the multiple different GALNTs [330].

Some studies have demonstrated that increased levels of glycosyltransferase transcripts generally correlate with increased levels of corresponding glycans [330, 331]. For example, Nairn et al., [331] investigated mouse embryonic stem cells as they differentiated into embryoid bodies and extraembryonic endodermal cells. They measured transcript levels of glycosylation-related genes by qRT-PCR and also interrogated the glycan profiles by mass spectrometry. The abundance of mono- and disialylated-core 1 structures was found to be

increased by 20% in extraembryonic endodermal cells, which correlated with a 6-10-fold increase in ST3GAL1 and ST6GALNAC5 glycosyltransferases compared to embryonic stem cells. However, there were also instances where there was no corresponding increase in glycan following increased transcription. For example, there was increased GCNT1 and GCNT3 (required for making core 2 glycans) despite a decrease in core 2 glycans on the cell. Interestingly, it is often noted that the correlation between mRNA abundance and protein abundance is also poor. This is attributed to varying regulatory mechanisms that act as translational modulators and post-translational modifications which affect individual protein half-life, such as ubiquitination and phosphorylation (reviewed by [332]). More recently, others [333] showed that RNA sequencing data taken from the Human Protein Atlas and glycans present on cells, detected by lectins and mass spectrometry, could generally be predicted from glycogene transcript levels. This included a correlation between the expression level of GCNT3 and levels of core 2 glycans present.

6.2 Aims and Objectives

The experiments described in this chapter aimed to investigate what transcriptional differences exist between HPA-positive and HPA-negative populations in MCF-7 and T-47D cell lines by performing RNA sequencing. The experiments described in this chapter also aimed to elucidate the transcriptional differences between TGF- β 1 treated MCF-7 cells and MCF-7 HPA-positive cells by performing RNA sequencing.

The objectives of this chapter are to:

- 1) Perform RNA sequencing of MCF-7 HPA-positive and HPA-negative subpopulations isolated from a mixed HPA-positive and HPA-negative population and identify the most significant DEGs and KEGGs in MCF-7 HPA-negative cells compared to MCF-7 HPA-positive cells.
- 2) Perform RNA sequencing of T-47D HPA-positive and HPA-negative subpopulations isolated from a mixed HPA-positive and HPA-negative population and identify the most significant DEGs and KEGGs in T-47D HPA-negative compared to T-47D HPA-positive cells.
- 3) Perform RNA sequencing of TGF- β 1 treated MCF-7 cells and analyse them in comparison to MCF-7 HPA-positive cells. To identify the most significant DEGs and KEGGs in TGF- β 1 treated MCF-7 cells compared to MCF-7 HPA-positive cells.

- 4) Identify any DEGs in the mucin-type O-glycan biosynthesis KEGG pathway between MCF-7 and T-47D HPA-positive and HPA-negative subpopulations, and between TGF- β 1 treated MCF-7 cells and MCF-7 HPA-positive cells.

6.3 Methods

6.3.1 Lectin-coated magnetic bead isolation of glycosylation-specific cell populations

To investigate transcriptomes of HPA-binding and non-binding cells of MCF-7 and T-47D cell lines, cells were separated into HPA-binding and non-binding populations using HPA-coated magnetic beads from total mixed cell populations as outlined in '2.9 *Lectin-coated magnetic bead isolation of glycosylation-specific cell populations*'.

6.3.2 Total RNA isolation, RNA sequencing and analysis

To investigate differences in transcriptomes of samples, total RNA was isolated from the following samples; MCF-7 HPA-positive cells, MCF-7 HPA-negative cells, T-47D HPA-positive cells, T-47D HPA-negative cells and MCF-7 TGF- β 1 treated cells (72-hour timepoint, 10ng/mL TGF- β 1 treatment, replenished every other day, as described in '2.8.1 *TGF- β 1 treatment*'). RNA isolation of samples and RNA sequencing were conducted as outlined in '2.10 *Ribonucleic acid (RNA) sequencing*'. Four biological replicates of each sample type were conducted.

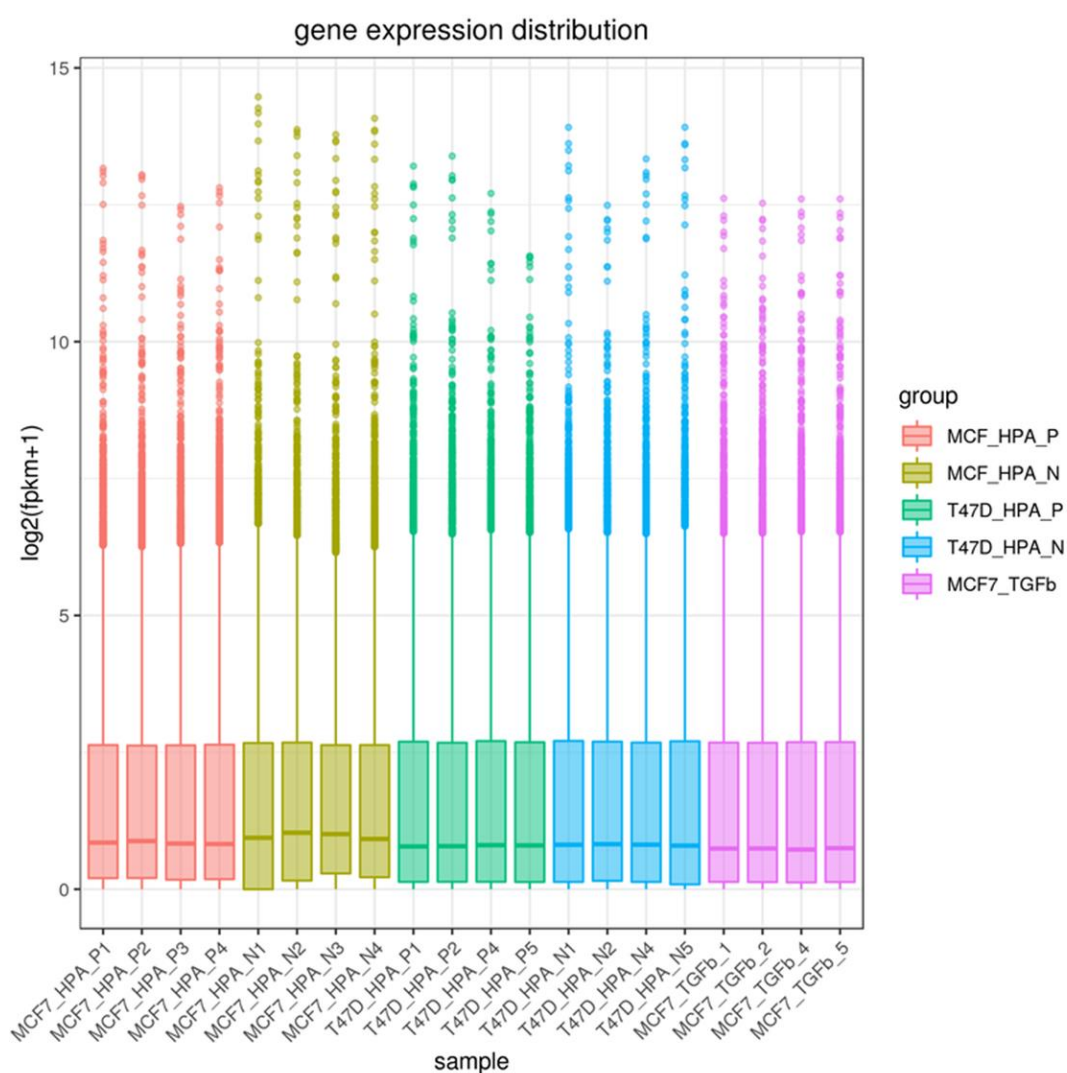
6.4 Results

6.4.1 RNA sequencing sample quality overview

Sample quality control performed by the external company Novogene confirmed that all samples had high score RNA integrity numbers, were of high quality (high Q30 scores), had equal representation of guanine and cytosine bases, the majority were clean reads and were mainly from exon regions, see Supplementary Figure 20-24 and Supplementary Table 2-3.

6.4.2 Gene expression overview

Sample gene expression number and density, presented as fragments per kilobase of transcript per million mapped reads (FPKM), which normalises for sequencing depth and gene length, was relatively constant across all samples, see Figure 6.1.

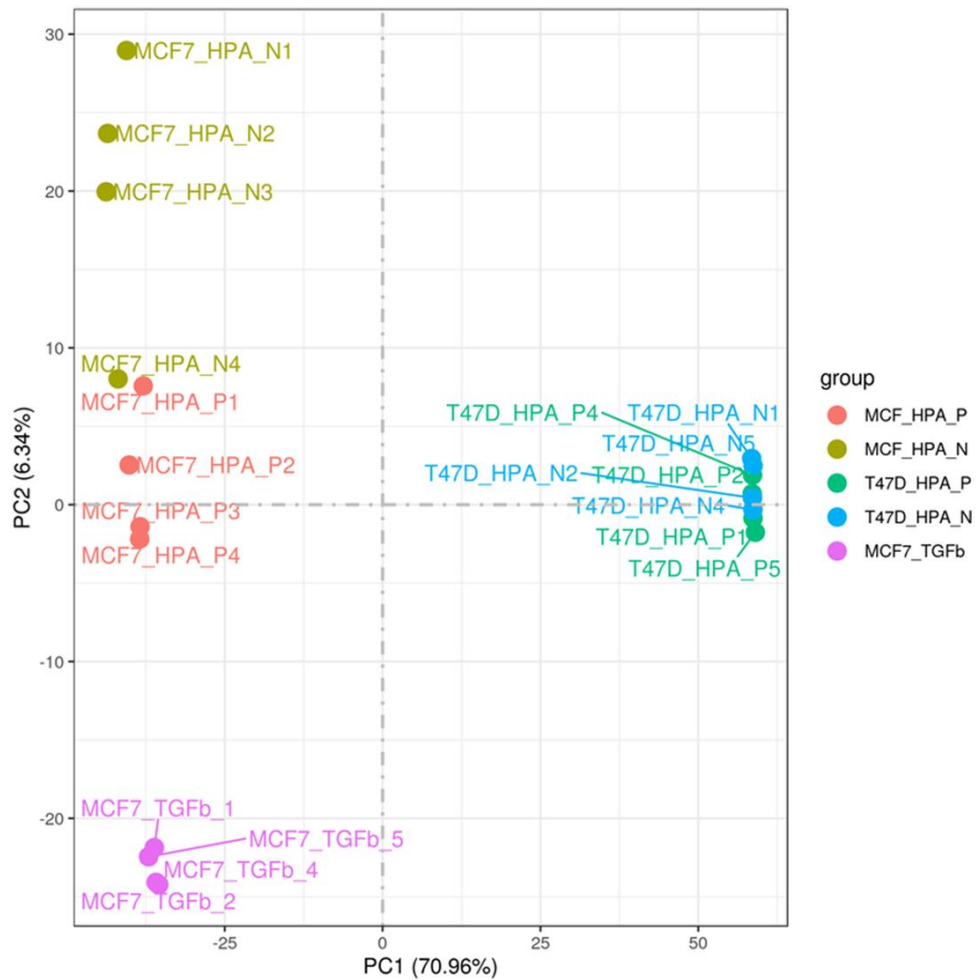


Group Key	
MCF_HPA_P = MCF-7 HPA-positive	T47D_HPA_N = T-47D HPA-negative
MCF_HPA_N = MCF-7 HPA-negative	MCF7_TGFB = MCF-7 TGF-β1 treated
T47D_HPA_P = T-47D HPA-positive	

Figure 6.1 Sample gene expression distribution box plot of MCF-7 *Helix pomatia* agglutinin (HPA)-positive, MCF-7 HPA-negative, T-47D HPA-positive, T-47D HPA-negative and MCF-7 transforming growth factor-β1 (TGF-β1) treated cells.

6.4.3 Principal component analysis (PCA)

PCA was used to assess inter-group differences and intra-group repeatability. The first principal component (PC1) accounted for 70.96% of the variation in the data and the second principal component (PC2) accounted for 6.34% of the variation in the data, see Figure 6.2.



Group Key	
MCF_HPA_P = MCF-7 HPA-positive	T47D_HPA_N = T-47D HPA-negative
MCF_HPA_N = MCF-7 HPA-negative	MCF7_TGFb = MCF-7 TGF-β1 treated
T47D_HPA_P = T-47D HPA-positive	

Figure 6.2 Principal component analysis (PCA) of MCF-7 *Helix pomatia* agglutinin (HPA)-positive, MCF-7 HPA-negative, T-47D HPA-positive, T-47D HPA-negative and MCF-7 transforming growth factor-β1 (TGF-β1) treated cells.

6.4.4 Gene co-expression Venn diagrams

Gene co-expression Venn diagrams demonstrate the number of uniquely expressed gene transcripts in a group. MCF-7 HPA-negative cells demonstrated 953 uniquely expressed genes compared to the MCF-7 HPA-positive cells, see Figure 6.3A. T-47D HPA-negative cells demonstrated 392 uniquely expressed genes compared to T-47D HPA-positive cells, see Figure 6.3B. MCF-7 TGF- β 1 treated cells demonstrated 428 uniquely expressed genes compared to MCF-7 HPA-positive cells, see Figure 6.3C. When MCF-7 HPA-positive, HPA-negative and TGF- β 1 treated cells were all compared there were 189, 805 and 208 uniquely expressed genes respectively, see Figure 6.3D.

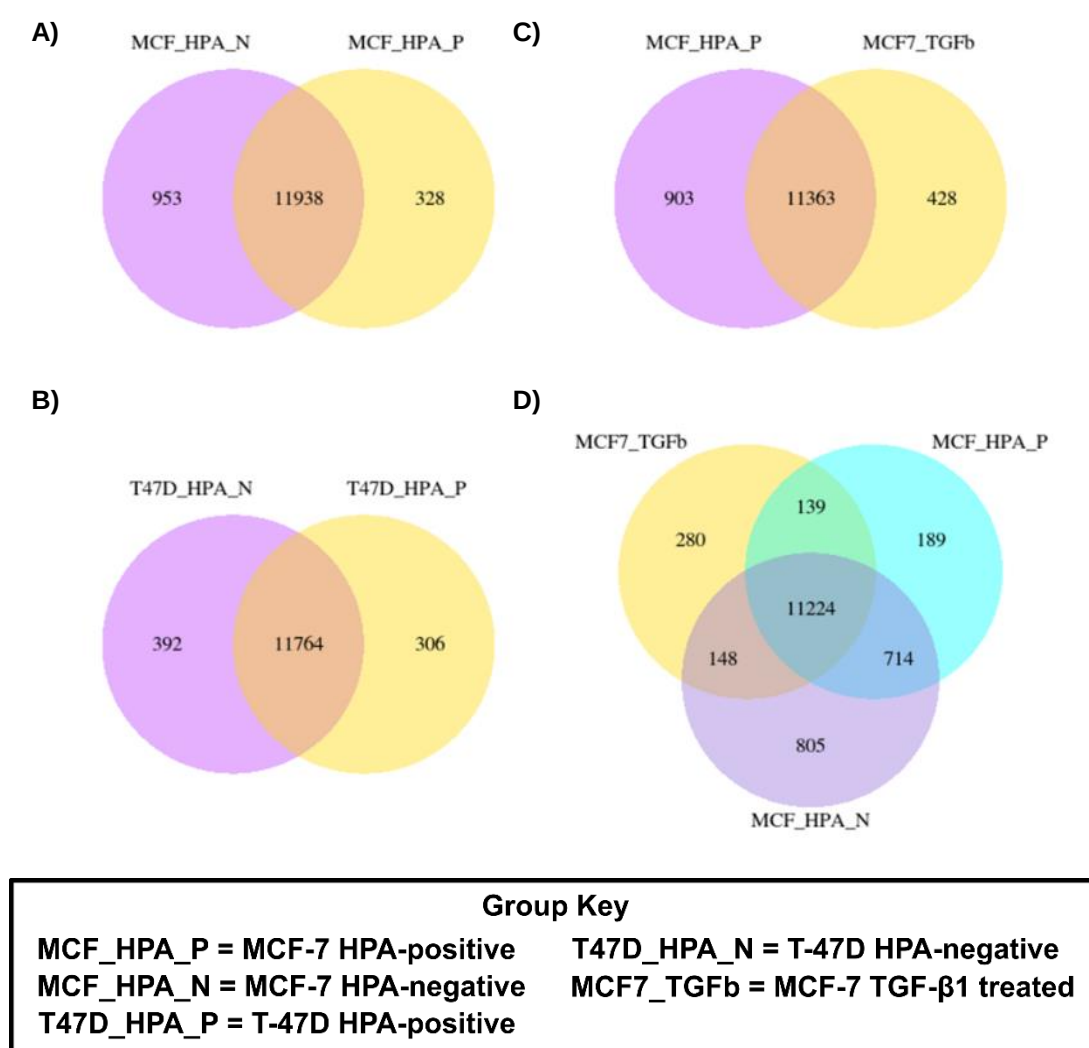


Figure 6.3 Gene co-expression Venn diagrams, an average of four biological replicates. Comparison groups include A) MCF-7 *Helix pomatia* agglutinin (HPA)-negative and MCF-7 HPA-positive, B) T-47D HPA-negative and T-47D HPA-positive, C) MCF-7 HPA-positive and MCF-7 transforming growth factor- β 1(TGF- β 1) treated and D) MCF-7 TGF- β 1 treated, MCF-7 HPA-positive and MCF-7 HPA-negative.

6.4.5 Total number of DEGs across samples

MCF-7 HPA-negative cells demonstrated a total of 4,601 DEGs, of which 2,386 were down-regulated and 2,215 were up-regulated compared to MCF-7 HPA-positive cells, see Figure 6.4. MCF-7 TGF- β 1 treated cells demonstrated 9,481 DEGs, of which 5,085 were down-regulated and 4,396 were up-regulated compared to MCF-7 HPA-negative cells. MCF-7 TGF- β 1 treated cells demonstrated 9,006 DEGs, of which 4,515 were down-regulated and 4,491 were up-regulated compared to MCF-7 HPA-positive cells. Finally, T-47D HPA-negative cells demonstrated 1,371 DEGs, of which 560 were down-regulated and 811 were up-regulated compared to T-47D HPA-positive cells.

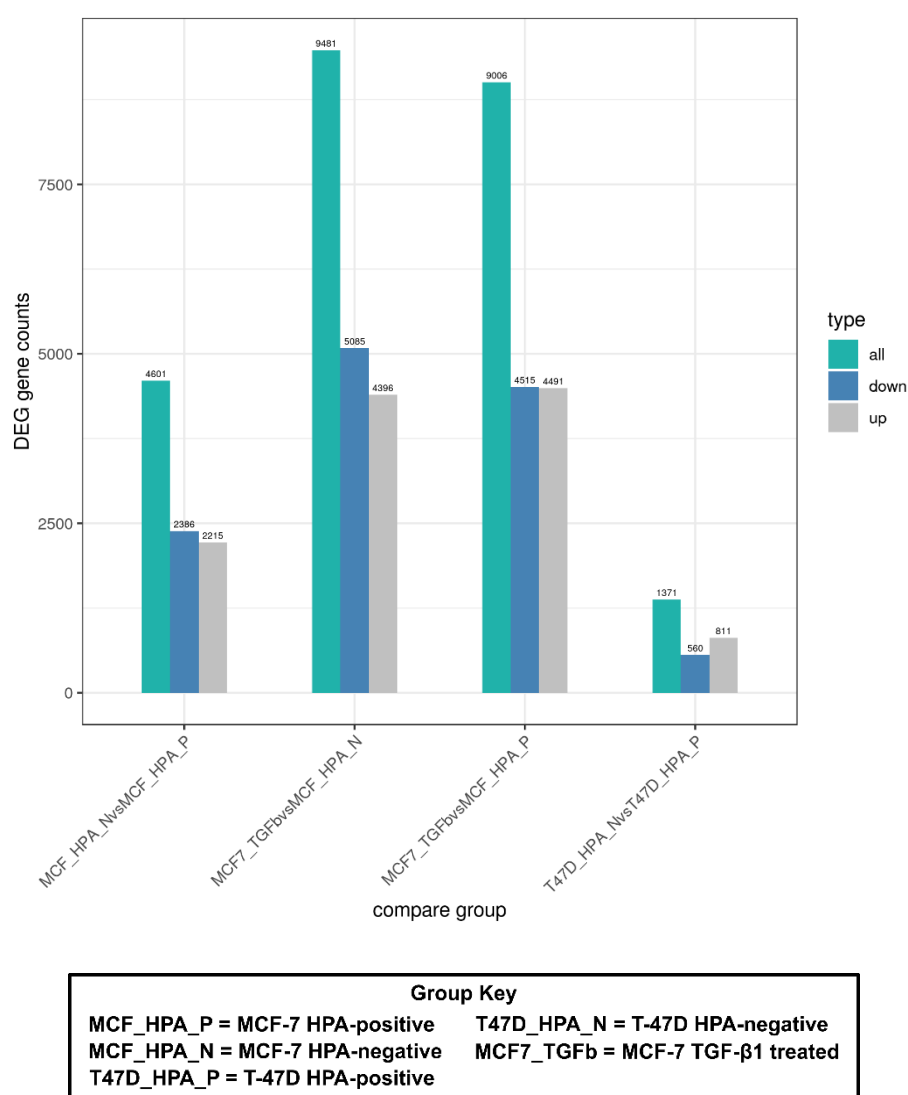


Figure 6.4 Total differentially expressed genes (DEGs), down-regulated DEG count and up-regulated DEG counts, an average of four biological replicates. Comparison groups include; MCF-7 *Helix pomatia* agglutinin (HPA)-negative versus MCF-7 HPA-positive, MCF-7 transforming growth factor- β 1 (TGF- β 1) treated versus MCF-7 HPA-negative, MCF-7 TGF- β 1 treated versus MCF-7 HPA-positive and T-47D HPA-negative versus T-47D HPA-positive.

6.4.6 MCF-7 HPA-negative and HPA-positive cell differential gene expression analysis

The volcano plot illustrated in Figure 6.5 demonstrates the fold-change ($\log_2\text{FoldChange}$) of all significant DEGs ($p < 0.05$) in MCF-7 HPA-negative samples compared to MCF-7 HPA-positive samples. Top ten most significant DEGs have been hand-annotated and described in Table 6.1; significantly upregulated are displayed in pink, significantly downregulated are displayed in green, and non-significant genes are displayed in blue.

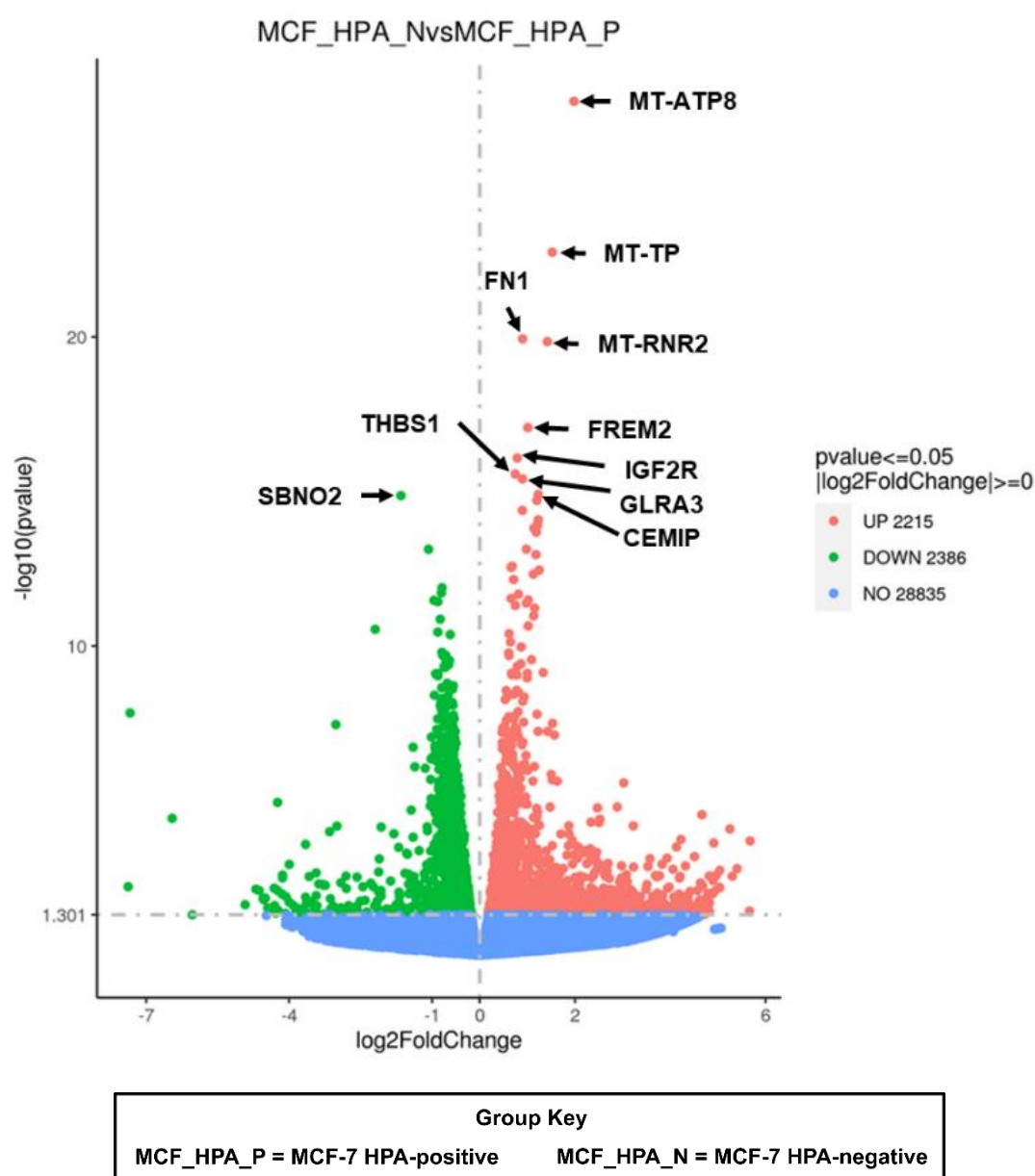


Figure 6.5 Volcano plot of significant differentially expressed genes (DEGs) in MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells, an average of four biological replicates. Top 10 most significant DEGs hand annotated. Significance defined at $p < 0.05$.

Table 6.1 Top 10 significant differentially expressed genes (DEGs) and fold-changes between MCF-7 *Helix pomatia* agglutinin (HPA)-negative compared to MCF-7 HPA-positive cells. Regulation direction is HPA-negative compared to HPA-positive, with values an average of four biological replicates.

Top 10 most significant DEGs in MCF-7 HPA-negative cells						
	Fold-change	Regulation direction	pvalue	Gene_name	Codes for:	Reference to involvement in EMT
1	3.94	Up	<0.025	MT-ATP8	Mitochondrially encoded ATP synthase membrane subunit 8	-
2	2.87	Up	<0.025	MT-TP	Mitochondrially Encoded TRNA-Pro (CCN)	-
3	1.87	Up	<0.025	FN1	Fibronectin 1	[334]
4	2.67	Up	<0.025	MT-RNR2	Mitochondrially Encoded 16S RNA	-
5	2.01	Up	<0.025	FREM2	FRAS1-Related Extracellular Matrix 2	[335]
6	1.73	Up	<0.025	IGF2R	Insulin-like growth factor 2	[336]
7	1.68	Up	<0.025	THBS1	Thrombospondin 1	[337]
8	1.86	Up	<0.025	GLRA3	Glycine receptor alpha-3	-
9	2.34	Up	<0.025	CEMIP	Cell migration inducing hyaluronidase 1	[338]
10	3.14	Down	<0.025	SBNO2	Strawberry Notch Homolog 2	-

6.4.7 T-47D HPA-negative and HPA-positive cell differential gene expression analysis

The volcano plot illustrated in Figure 6.6 demonstrates the fold-change ($\log_2\text{FoldChange}$) of all significant DEGs ($p < 0.05$) in T-47D HPA-negative samples compared to T-47D HPA-positive. Top ten most significant DEGs have been hand-annotated and described in Table 6.2; significantly upregulated are displayed in pink, significantly downregulated are displayed in green, and non-significant genes are displayed in blue.

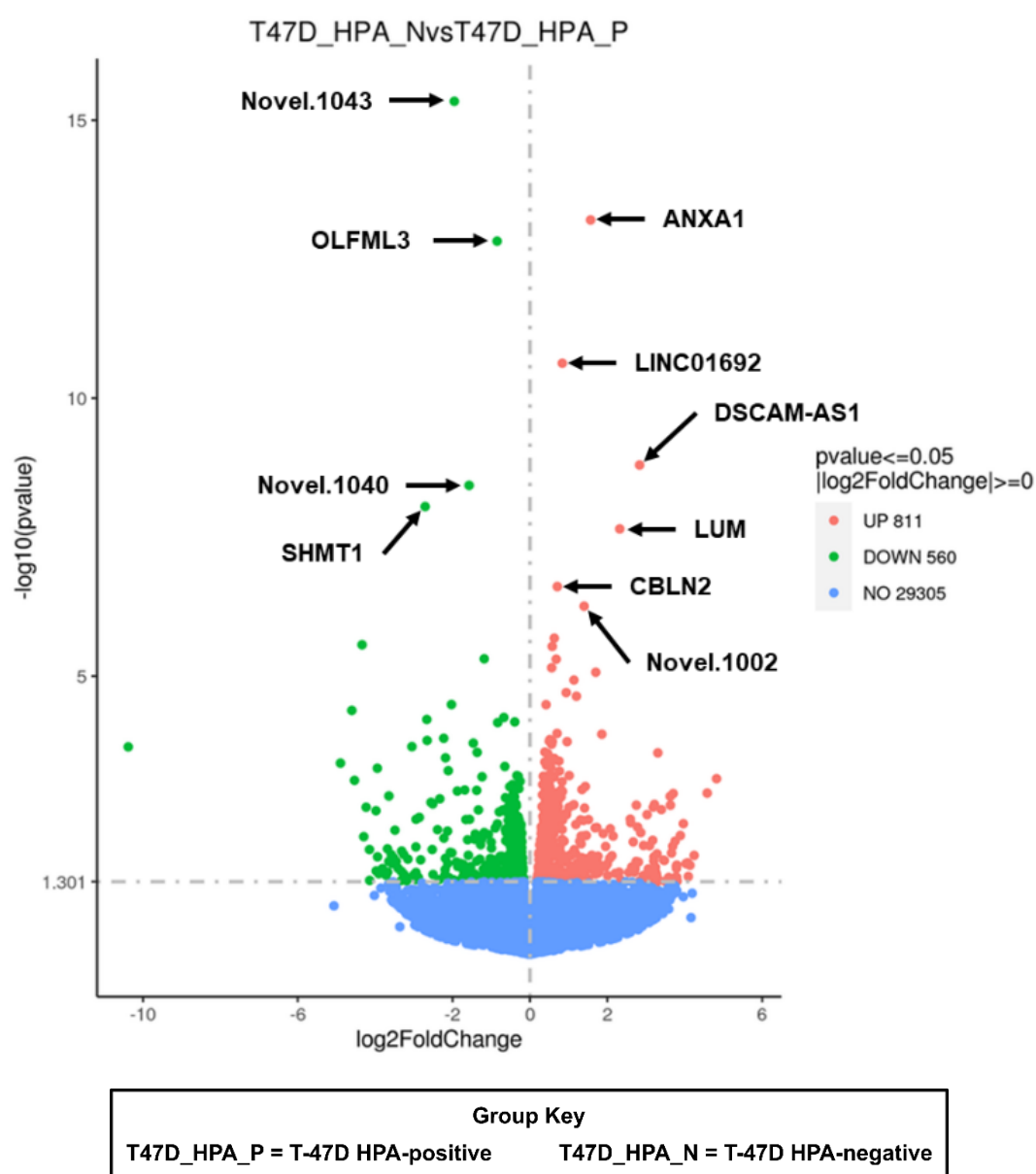


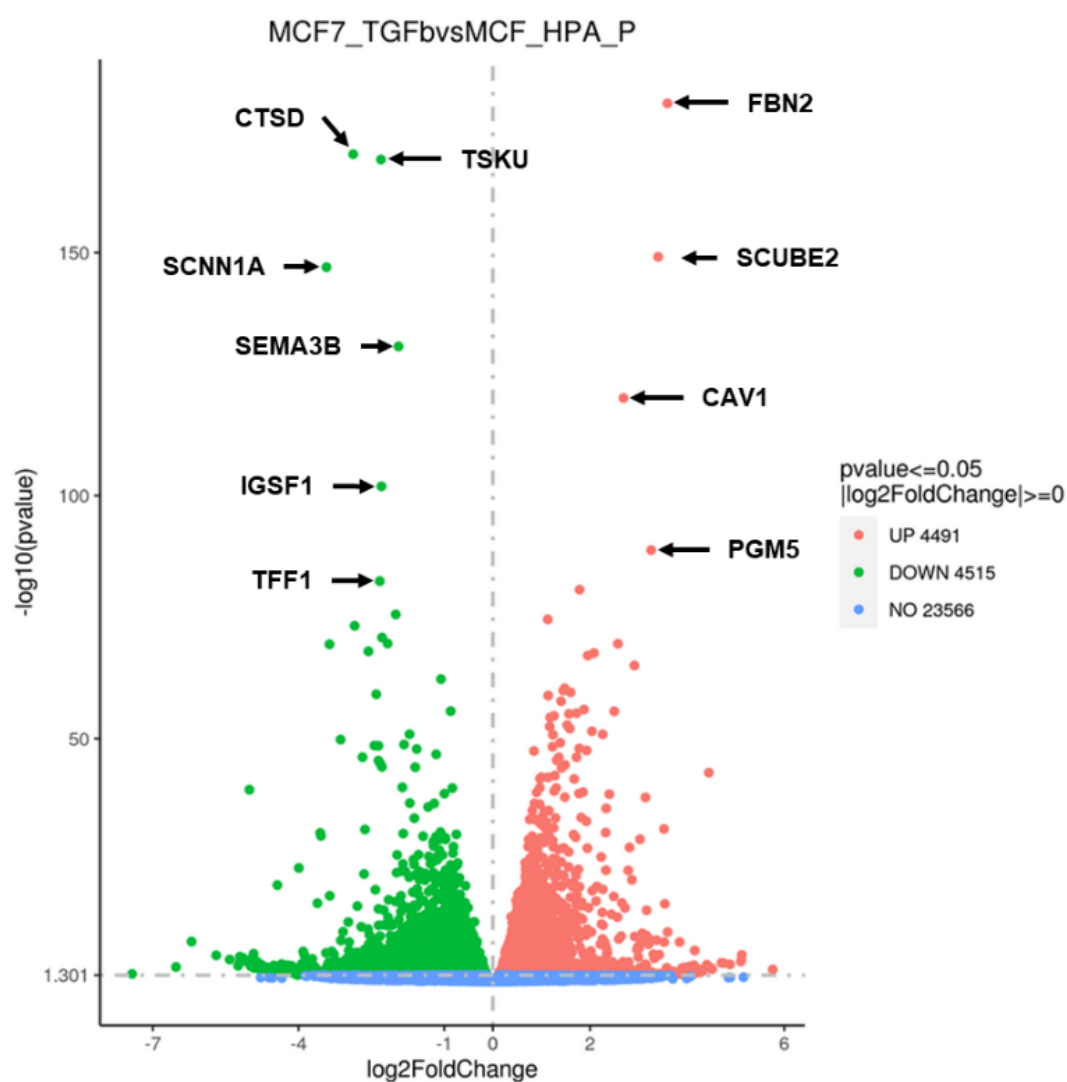
Figure 6.6 Volcano plot of significant differentially expressed genes (DEGs) in T-47D *Helix pomatia* agglutinin (HPA)-negative cells compared to T-47D HPA-positive cells, an average of four biological replicates. Top 10 most significant DEGs hand annotated. Significance defined at $p < 0.05$.

Table 6.2 Top 10 significant differentially expressed genes (DEGs) and fold-changes between T-47D *Helix pomatia* agglutinin (HPA)-negative compared to T-47D HPA-positive cells. Regulation direction is HPA-negative compared to HPA-positive, with values an average of four biological replicates.

Top 10 most significant DEGS in T-47D HPA-negative cells						
	Fold-change	Regulation direction	pvalue	Gene_name	Codes for:	Reference to involvement in EMT
1	3.87	Down	<0.025	Novel.1043	Novel (Sphingomyelin phosphodiesterase 4)	-
2	2.97	Up	<0.025	ANXA1	Annexin A1	[339]
3	1.80	Down	<0.025	OLFML3	Olfactomedin Like 3	[340]
4	1.79	Up	<0.025	LINC01692	Long Intergenic Non-Protein Coding RNA 1692	-
5	7.13	Up	<0.025	DSCAM-AS1	DSCAM Antisense RNA 1	[341]
6	2.97	Down	<0.025	Novel.1040	Novel (Electrogenic aspartate/glutamate antiporter SLC25A12)	-
7	6.53	Down	<0.025	SHMT1	Serine Hydroxymethyltransferase 1	[342]
8	5.00	Up	<0.025	LUM	Lumican	[343]
9	1.63	Up	<0.025	CBLN2	Cerebellin 2 Precursor	-
10	2.64	Up	<0.025	Novel.1002	Novel (Beta-1,3-N-Acetylgalactosaminyltransferase 2)	-

6.4.8 MCF-7 TGF- β 1 treated cells and MCF-7 HPA-positive cells differential gene expression analysis

The volcano plot illustrated in Figure 6.7 demonstrates the fold-change ($\log_2\text{FoldChange}$) of all significant DEGs ($p < 0.05$) in MCF-7 TGF- β 1 treated samples compared to MCF-7 HPA-positive. Top ten most significant DEGs have been hand-annotated and described in Table 6.3; significantly upregulated are displayed in pink, significantly downregulated are displayed in green, and non-significant genes are displayed in blue.



Group Key	
MCF_HPA_P = MCF-7 HPA-positive	MCF7_TGFb = MCF-7 TGF- β 1 treated

Figure 6.7 Volcano plot of significant differentially expressed genes (DEGs) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 *Helix pomatia* agglutinin (HPA)-positive cells, an average of four biological replicates. Top 10 most significant DEGs hand annotated. Significance defined at $p < 0.05$.

Table 6.3 Top 10 significant differentially expressed genes (DEGs) and fold-changes between MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells and MCF-7 *Helix pomatia* agglutinin (HPA)-positive cells. Regulation direction is TGF- β 1 treated compared to HPA-positive, with values an average of four biological replicates.

Top 10 most significant DEGs in MCF-7 TGF- β 1 treated cells						
	Fold-change	Regulation direction	pvalue	Gene_name	Codes for:	Reference to involvement in EMT
1	12.08	Up	<0.025	FBN2	Fibrillin 2	[344]
2	7.34	Down	<0.025	CTSD	Cathepsin D	[345]
3	4.94	Down	<0.025	TSKU	Tsukushi, Small Leucine Rich Proteoglycan	[346]
4	10.57	Up	<0.025	SCUBE2	Signal Peptide, CUB Domain and EGF Like Domain Containing 2	[347]
5	10.75	Down	<0.025	SCNN1A	Sodium Channel Epithelial 1 Subunit Alpha	[348]
6	3.84	Down	<0.025	SEMA3B	Semaphorin 3B	[349]
7	6.45	Up	<0.025	CAV1	Caveolin 1	[350]
8	4.91	Down	<0.025	IGSF1	Immunoglobulin Superfamily Member 1	[351]
9	9.58	Up	<0.025	PGM5	Phosphoglucomutase 5	[352]
10	5.03	Down	<0.025	TFF1	Trefoil Factor 1	[353]

6.4.9 KEGG enrichment analysis

KEGG enrichment analysis was also conducted to identify high-order functionality of the significant DEGs for each sample type. Of note, compared to MCF-7 HPA-positive cells, MCF-7 HPA-negative cells significantly upregulated DEGs were significantly enriched for KEGGs; including ECM-receptor interactions (25/822 DEGs, $p < 0.025$), proteoglycans in cancer (42/882 DEGs, $p < 0.025$), microRNAs in cancer (33/882 DEGs, $p < 0.025$), signalling pathways regulating pluripotency of stem cells (28/882 DEGs, $p < 0.025$) and focal adhesions (41/882 DEGs, $p < 0.025$), see Figure 6.8A. Significantly enriched KEGGs of MCF-7 HPA-negative significantly downregulated DEGs included; carbon metabolism (40/1,161 DEGs, $p < 0.025$), proteasome (19/1,161 DEGs, $p < 0.025$) and oxidative phosphorylation (50/1,161 DEGs, $p < 0.025$), see Figure 6.9A.

Significantly upregulated DEGs of MCF-7 TGF- β 1 treated cells, in comparison to MCF-7 HPA-positive cells, were significantly enriched for KEGGs; including proteoglycans in cancer (86/1,915 DEGs, $p < 0.025$), Wnt signalling pathway (60/1,915 DEGs, $p < 0.025$), TGF- β

signalling pathway (45/1,915 DEGs, $p < 0.025$), signalling pathways regulating pluripotency of stem cells (49/1,195 DEGs, $p < 0.025$), focal adhesion (79/1,915 DEGs, $p < 0.025$) and microRNAs in cancer (56/1,915 DEGs, $p < 0.025$), see Figure 6.8B. Whereas, significantly downregulated DEGs were significantly enriched for KEGGs; including chemical carcinogenesis – reactive oxygen species (94/1,674 DEGs, $p < 0.025$), DNA replication (16/1,674 DEGs, $p < 0.025$), oxidative phosphorylation (62/1,674 DEGs, $p < 0.025$) and the proteasome (19/1,674 DEGs, $p < 0.025$), see Figure 6.9B.

KEGG enrichment analysis of T-47D HPA-positive significantly upregulated DEGs revealed significant enrichment for; signalling pathways regulating pluripotency of stem cells (12/295 DEGs, $p < 0.025$), polycomb repressive complex (10/295 DEGs, $p < 0.025$), microRNAs in cancer (13/295 DEGs, $p < 0.025$), and mucin-type O-glycan biosynthesis (4/295 DEGs, $p < 0.05$), see Figure 6.8C. Whereas, significantly downregulated DEGs were associated with KEGGs; including the pentose phosphate pathway (5/282 DEGs, $p < 0.05$), spliceosome (14/282 DEGs, $p < 0.05$) and DNA replication (5/282 DEGs, $p < 0.05$), see Figure 6.9C.

Interestingly, for both up- and downregulated DEGs, proteoglycans in cancer KEGG (hsa05205) was significantly enriched for MCF-7 HPA-negative cells (64/2,043 DEGs, $p < 0.025$) compared to MCF-7 HPA-positive cells, see Supplementary Figure 25 and in MCF-7 TGF- β 1 treated cells (119/3,589 DEGs, $p < 0.025$) compared to MCF-7 HPA-positive cells, see Supplementary Figure 26. It was not significant for T-47D HPA-negative cells (19/577 DEGs, $p = 0.07$) compared to T-47D HPA-positive cells, see Supplementary Figure 27. All DEGs that demonstrated a fold-change of ≥ 1.5 and $p < 0.05$, indicating those with the most biological significance, were collated into Supplementary Table 4-6.

Similarly, signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) was significantly enriched for both up- and downregulated DEGs across all comparison groups, MCF-7 HPA-negative cells (39/2,043 DEGs, $p < 0.05$) in comparison to MCF-7 HPA-positive cells, see Supplementary Figure 28, MCF-7 TGF- β 1 treated cells (62/3,589 DEGs, $p < 0.05$) compared to MCF-7 HPA-positive cells, see Supplementary Figure 29 and T-47D HPA-negative cells (17/577 DEGs, $p < 0.025$) in comparison to T-47D HPA-positive cells, see Supplementary Figure 30. All DEGs that demonstrated a fold-change of ≥ 1.5 and $p < 0.05$, indicating those with the most biological significance, were collated into Supplementary Table 7-9.

A) MCF-7 HPA-negative vs MCF-7 HPA-positive KEGG - upregulated

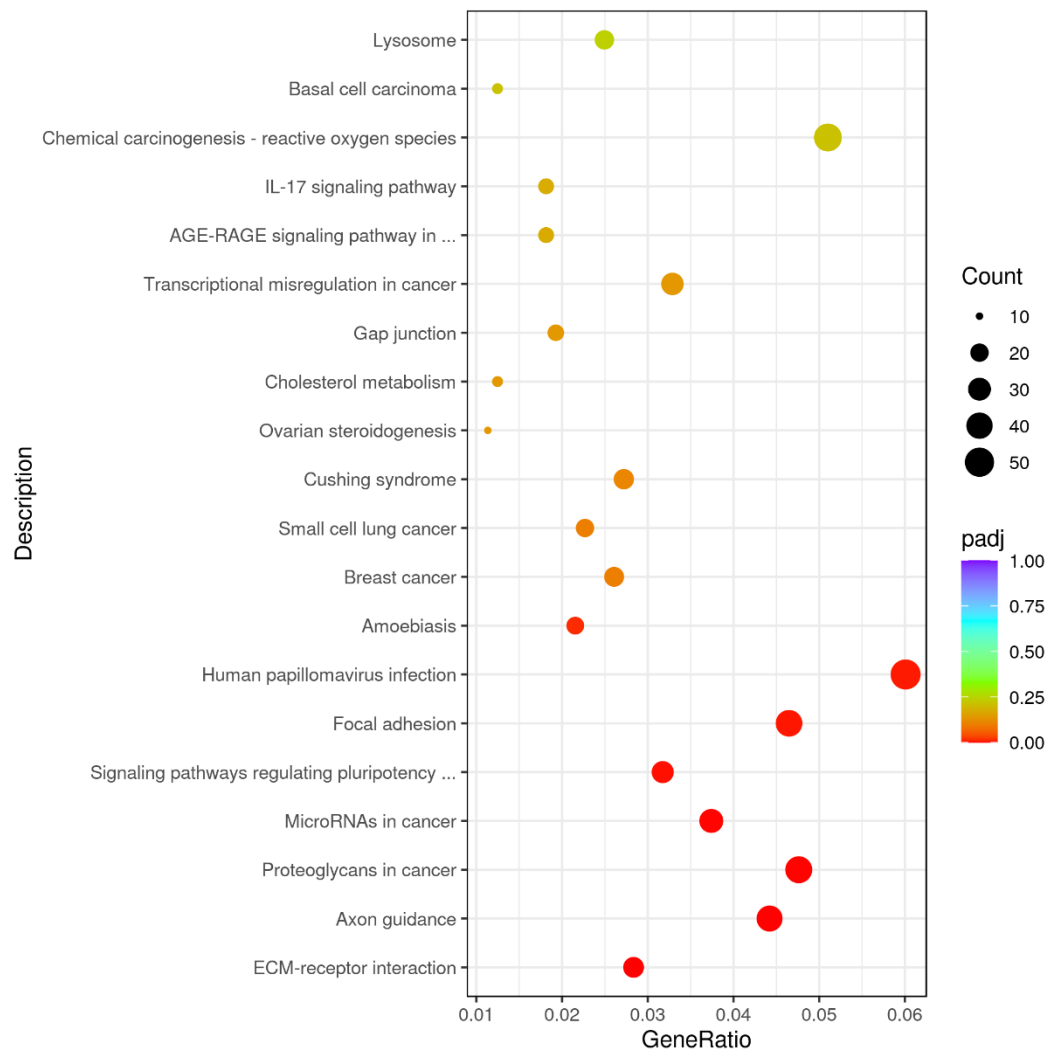


Figure 6.8 KEGG enrichment scatter plots of the 20 most significant KEGG pathways for upregulated differentially expressed genes (DEGs), each sample is an average of four biological replicates. DEGs KEGG pathway scatter plot for A) MCF-7 *Helix pomatia* agglutinin (HPA)-negative cell upregulated, B) MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cell upregulated and C) T-47D HPA-negative cell upregulated DEGs (1/3).

B) MCF-7 TGF- β 1 vs MCF-7 HPA-positive KEGG - upregulated

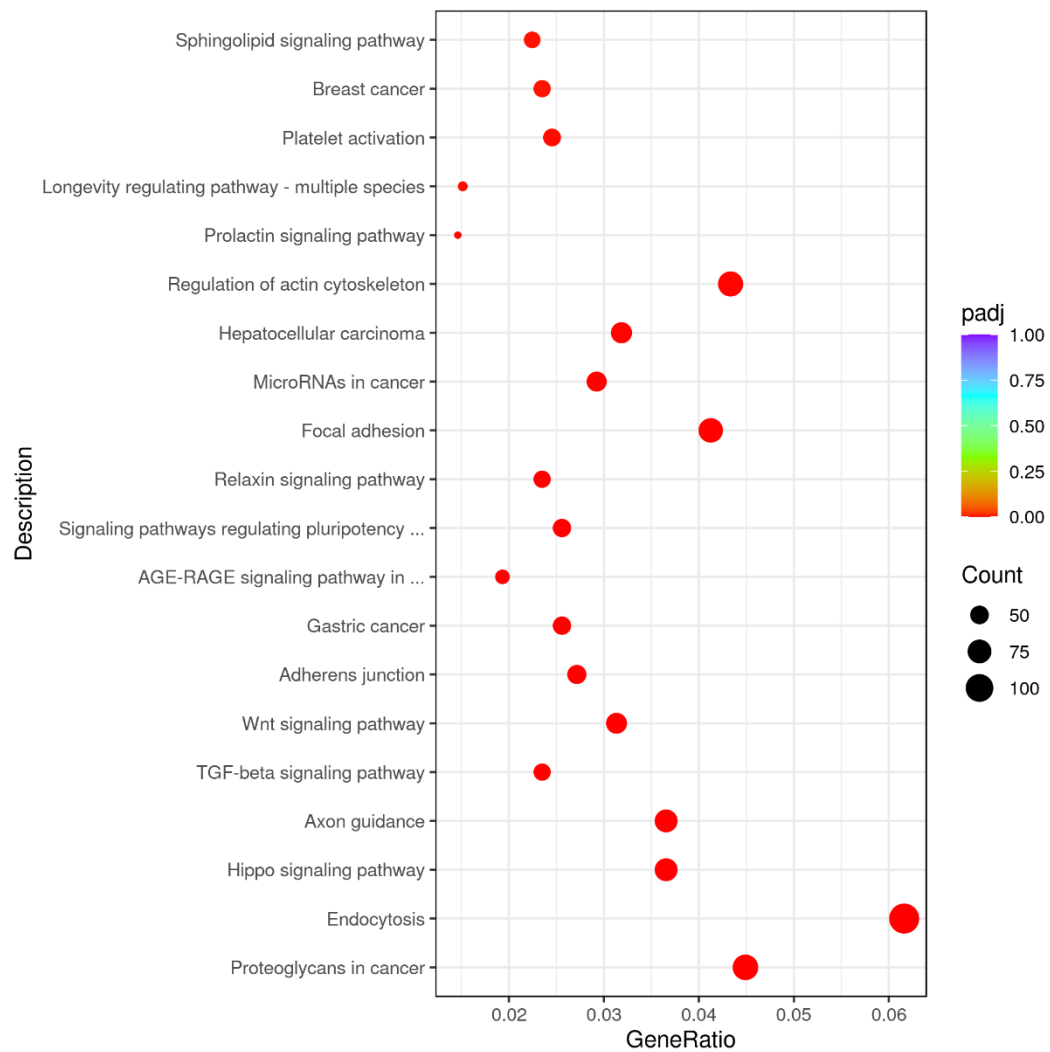


Figure 6.8 (Continued 2/3).

C) T-47D HPA-negative vs T-47D HPA-positive KEGG - upregulated

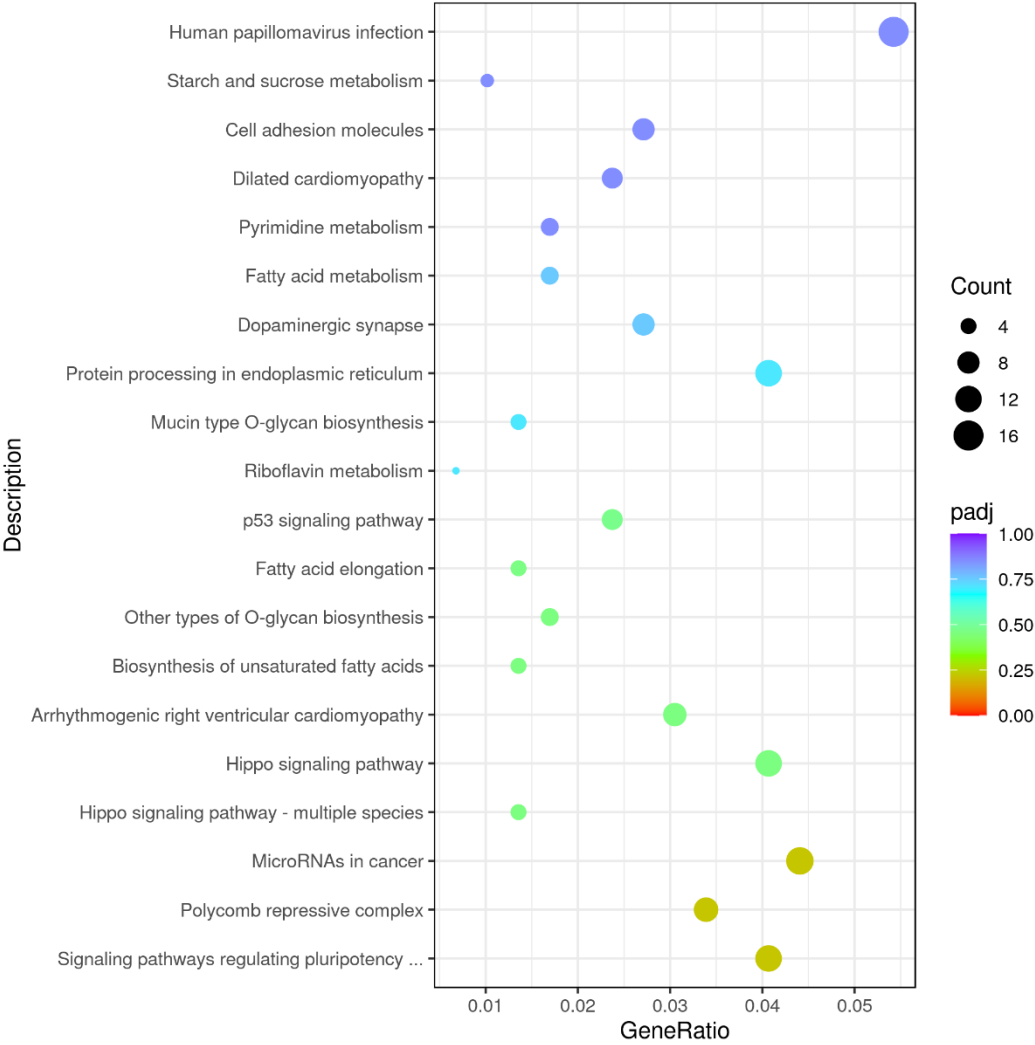


Figure 6.8 (Continued 3/3).

A) MCF-7 HPA-negative vs MCF-7 HPA-positive KEGG - downregulated

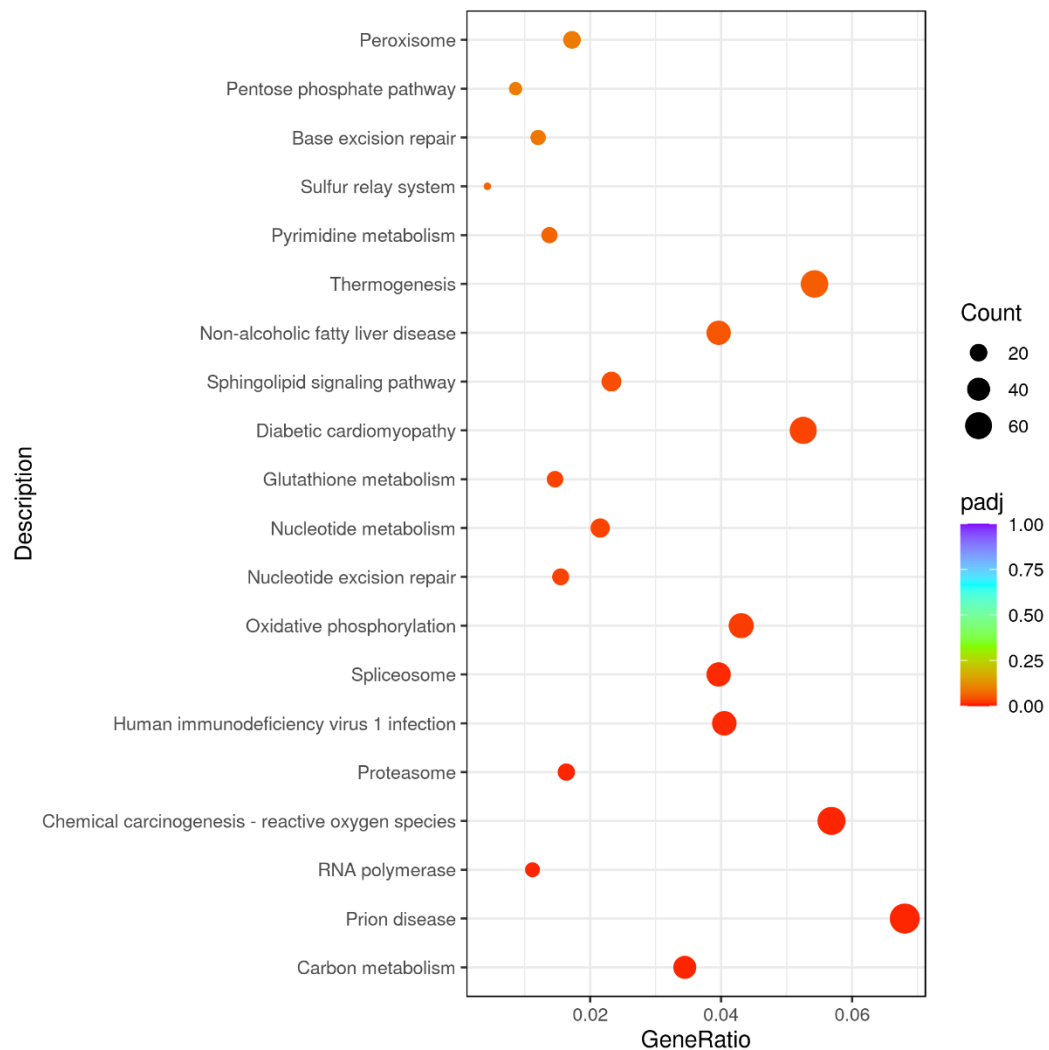


Figure 6.9 KEGG enrichment scatter plots of the 20 most significant KEGG pathways for downregulated differentially expressed genes (DEGs), each sample is an average of four biological replicates. DEGs KEGG pathway scatter plot for A) MCF-7 *Helix pomatia* agglutinin (HPA)-negative cell downregulated, B) MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cell downregulated and C) T-47D HPA-negative cell downregulated DEGs (1/3).

B) MCF-7 TGF- β 1 vs MCF-7 HPA-positive KEGG - downregulated

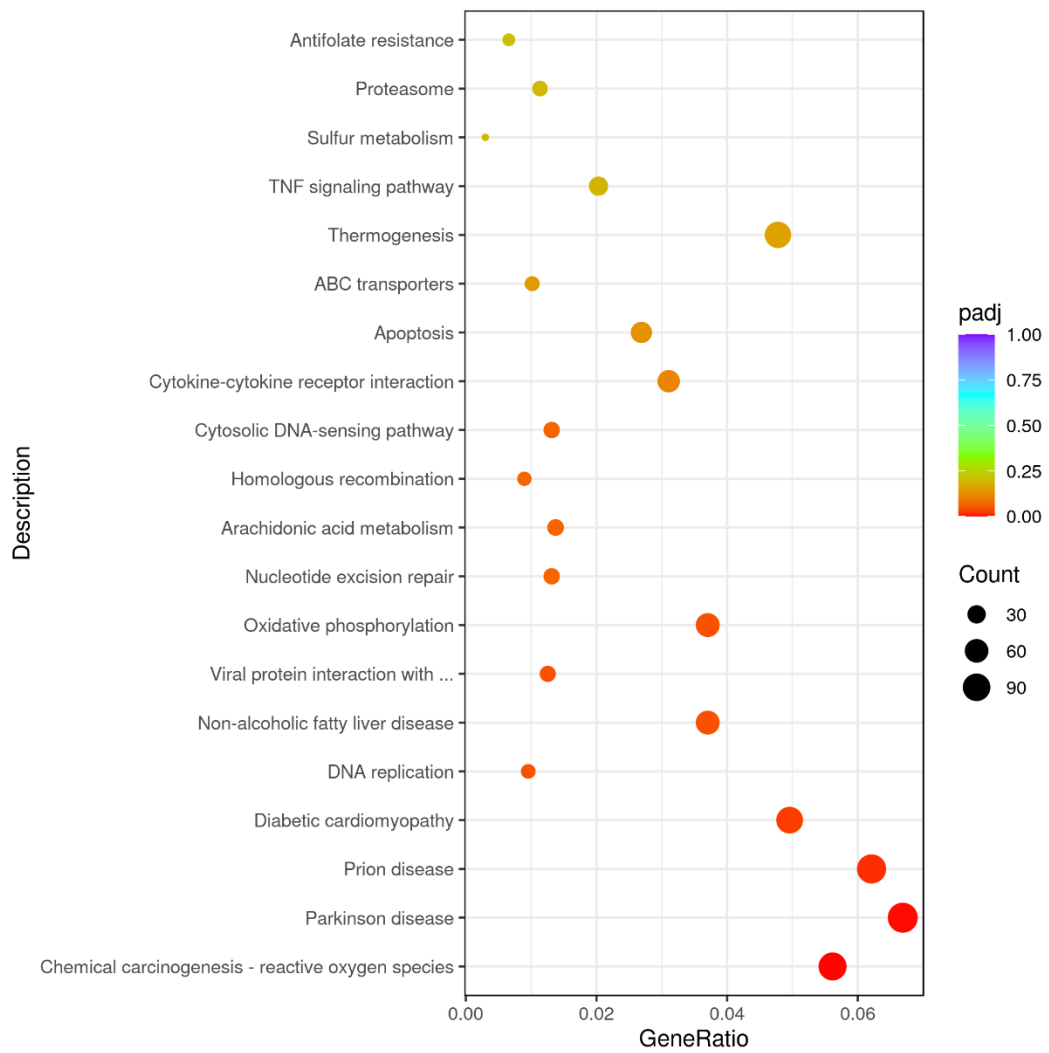


Figure 6.9 (Continued 2/3).

C) T-47D HPA-negative vs T-47D HPA-positive KEGG - downregulated

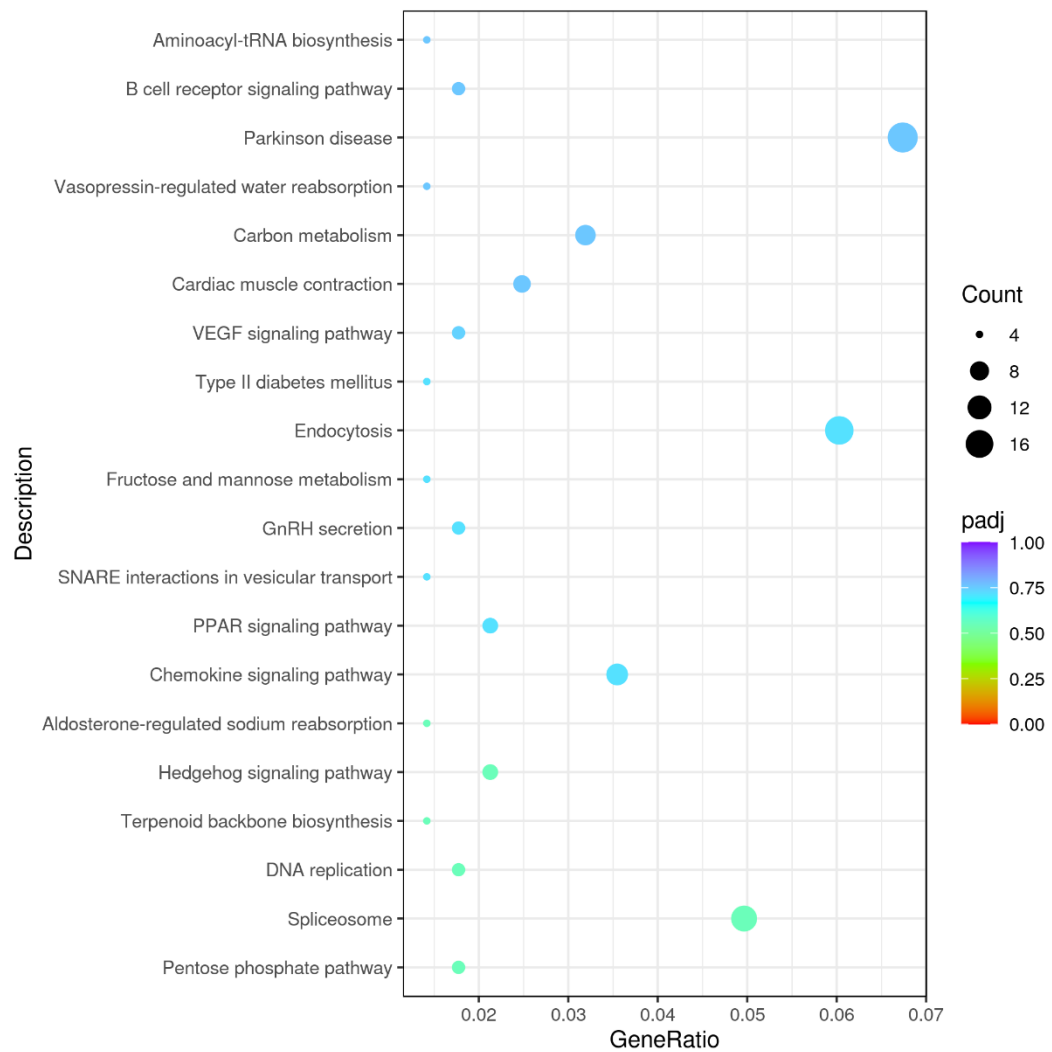


Figure 6.9 (Continued 3/3).

6.4.10 Mucin-type O-glycan biosynthesis KEGG (hsa00512)

MCF-7 HPA-negative cells were not significantly enriched for mucin-type O-glycan biosynthesis KEGG (hsa00512) (6/2,043 DEGs, $p=0.81$) however, six genes denoted as assigned to the KEGG were all individually significantly up- or down-regulated in comparison to MCF-7 HPA-positive cells. Although the overall KEGG was not significant these results were still considered due to the potential functional impact of individual DEGs identified which follow a clear pathway to elongation of the Tn antigen to a single sialylated-core 1 structure, and not a disialylated-core 1 structure, see Figure 6.10. DEGs include GALNT7 (+1.31 fold-change, $p<0.025$), SAP30L-AS1 (+1.59 fold-change, $p<0.025$), C1GALT1 (+1.29 fold-change, $p<0.05$), ST3GAL1 (+1.24 fold-change, $p<0.025$), ST3GAL2 (+1.40 fold-change, $p<0.025$) and ST6GALNAC4 (-1.32 fold-change, $p<0.025$), see Table 6.4.

Whereas, T-47D HPA-negative cells were significantly enriched for mucin-type O-glycan biosynthesis KEGG (hsa00512) (6/577 DEGs, $p<0.05$). Interestingly, many DEGs denoted as assigned to the KEGG are the same as, and also highlight a similar pathway to, the MCF-7 HPA-negative cells, see Figure 6.11. DEGs include GALNT3 (+1.21 fold-change, $p<0.05$), GALNT7 (+1.25 fold-change, $p<0.025$), SAP30L-AS1 (+1.37 fold-change, $p<0.05$), C1GALT1 (+1.35 fold-change, $p<0.025$), GALNT18 (-1.27 fold-change, $p<0.05$) and ST6GALNAC4 (-1.36 fold-change, $p<0.025$), see Table 6.5.

TGF- β 1 treated MCF-7 cells were not significantly enriched for mucin-type O-glycan biosynthesis KEGG (hsa00512) (16/3,589 DEGs, $p=0.29$) compared to MCF-7 HPA-positive cells. These cells demonstrated a similar, but also distinct, DEGs in the KEGG pathway in comparison to HPA-negative MCF-7 and T-47D cells, with promotion of core 2 O-glycan biosynthesis, see Figure 6.12. DEGs include GALNT1 (+1.47 fold-change, $p<0.025$), GALNT7 (+ 1.26 fold-change, $p<0.025$), GALNT10 (+1.87 fold-change, $p<0.025$), GALNT11 (+1.43 fold-change, $p<0.025$), GALNT12 (+ 2.01 fold-change, $p<0.025$), GALNT16 (+2.76 fold-change, $p<0.025$), GCNT4 (+1.83 fold-change, $p<0.025$), B4GALT5 (+1.31 fold-change, $p<0.025$), GALNT6 (-1.13 fold-change, $p<0.025$), GALNT14 (-1.83 fold-change, $p<0.025$), GALNT18 (-1.21 fold-change, $p<0.025$), B3GNT3 (-1.42 fold-change, $p<0.025$), B3GNT6 (-10.97 fold-change, $p<0.025$), ST3GAL1 (-1.74 fold-change, $p<0.025$), ST3GAL2 (-1.38 fold-change, $p<0.025$) and ST6GALNAC4 (-1.55 fold-change, $p<0.025$), see Table 6.6.

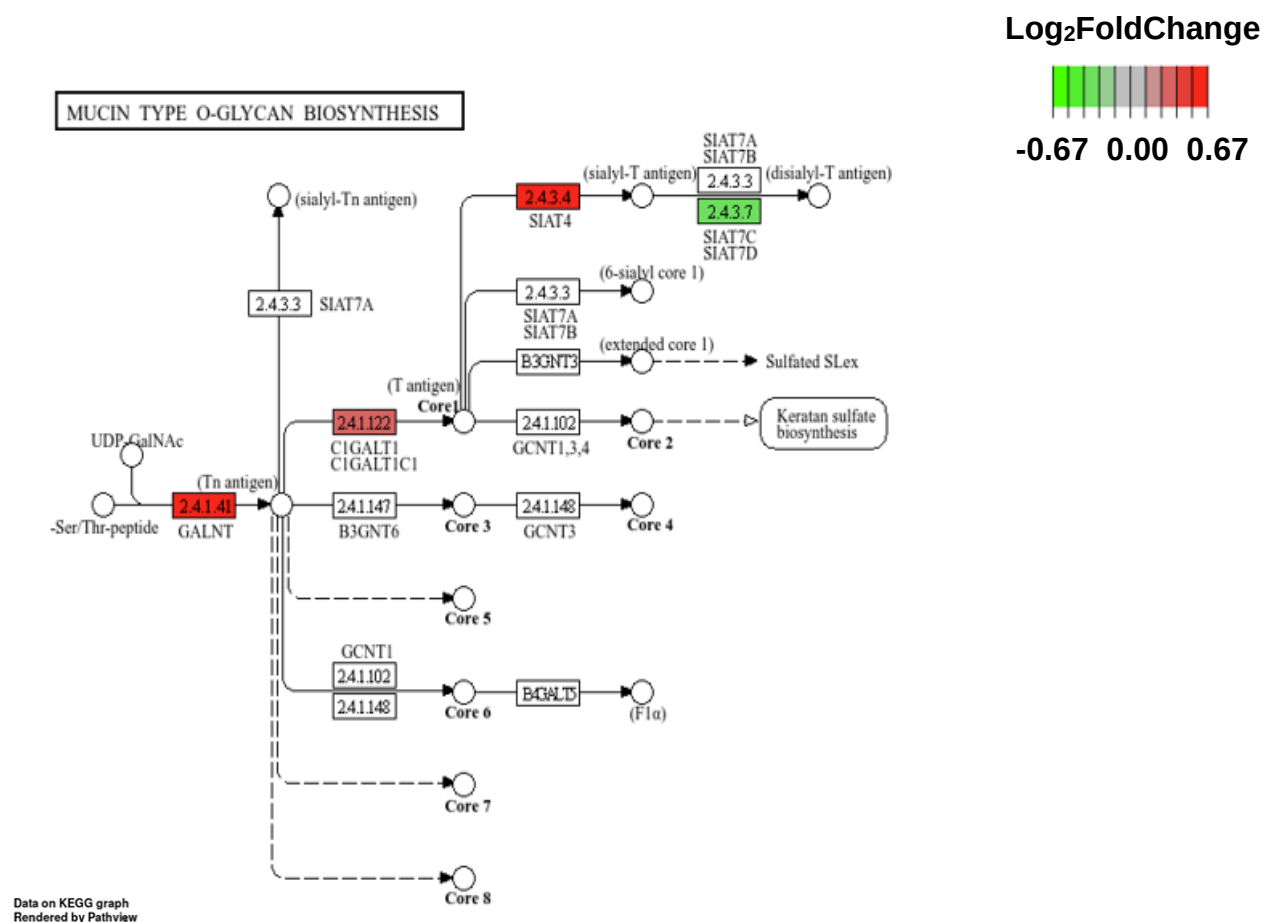


Figure 6.10 Mucin-type O-glycan biosynthesis KEGG (hsa00512) identified 6/2,043 differentially expressed genes (DEGs) ($p=0.81$) in MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells. Four biological replicates were conducted.

Table 6.4 MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells mucin-type O-glycan biosynthesis KEGG (hsa00512) differentially expressed genes (DEGs). Regulation direction is HPA-negative compared to HPA-positive. Four biological replicates were conducted (6/2,043 genes, $p=0.81$).

MCF-7 HPA-negative mucin-type O-glycan biosynthesis KEGG DEGs			
Gene name	Fold-change	Regulation direction	pvalue
GALNT7	1.31	Up	$p<0.025$
SAP30L-AS1	1.59	Up	$p<0.025$
C1GALT1	1.29	Up	$p<0.05$
ST3GAL1	1.24	Up	$p<0.025$
ST3GAL2	1.40	Up	$p<0.025$
ST6GALNAC4	1.32	Down	$p<0.025$

Table 6.5 T-47D *Helix pomatia* agglutinin (HPA)-negative cells mucin-type O-glycan biosynthesis KEGG (hsa00512) differentially expressed genes (DEGs). Regulation direction is HPA-negative compared to HPA-positive. Four biological replicates were conducted (6/577 genes, $p<0.05$).

T-47D HPA-negative mucin-type O-glycan biosynthesis KEGG DEGs			
Gene name	Fold-change	Regulation direction	pvalue
GALNT3	1.21	Up	$p<0.05$
GALNT7	1.25	Up	$p<0.025$
SAP30L-AS1	1.37	Up	$p<0.05$
C1GALT1	1.35	Up	$p<0.025$
GALNT18	1.27	Down	$p<0.05$
ST6GALNAC4	1.36	Down	$p<0.025$

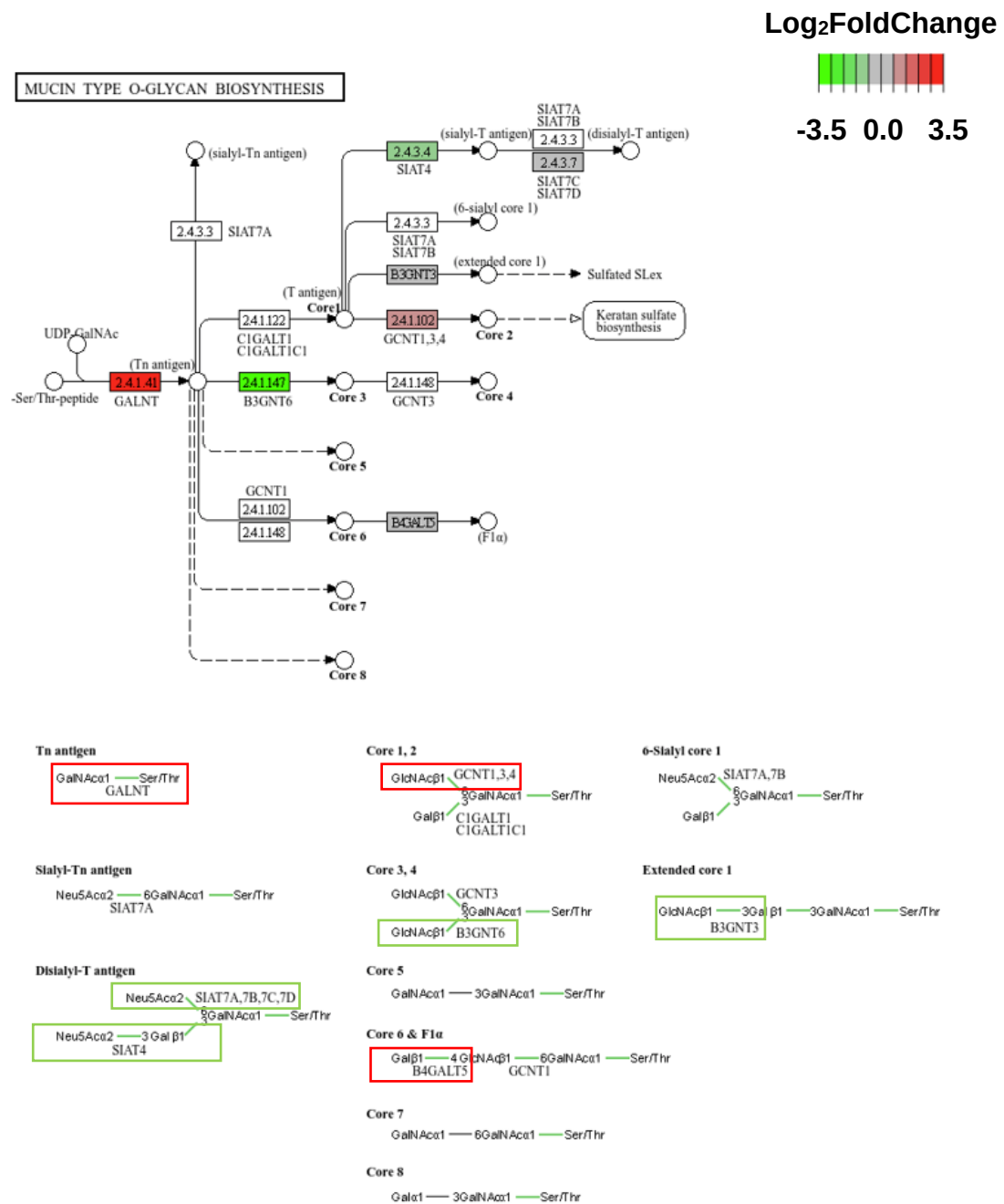


Figure 6.12 Mucin-type O-glycan biosynthesis KEGG (hsa00512) identified 16/3,589 differentially expressed genes (DEGs) ($p=0.29$) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 *Helix pomatia* agglutinin (HPA)-positive cells. Four biological replicates were conducted.

Table 6.6 MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cell mucin-type O-glycan biosynthesis KEGG (hsa00512) differentially expressed genes (DEGs). Regulation direction is TGF- β 1 treated compared to *Helix pomatia* agglutinin (HPA)-positive. Four biological replicates were conducted (16/3,589 genes, $p=0.29$).

MCF-7 TGF- β 1 treated mucin-type O-glycan biosynthesis KEGG DEGs			
Gene name	Fold-change	Regulation direction	pvalue
GALNT1	1.47	Up	$p<0.025$
GALNT7	1.26	Up	$p<0.025$
GALNT10	1.87	Up	$p<0.025$
GALNT11	1.43	Up	$p<0.025$
GALNT12	2.01	Up	$p<0.025$
GALNT16	2.76	Up	$p<0.025$
GCNT4	1.83	Up	$p<0.025$
B4GALT5	1.31	Up	$p<0.025$
GALNT6	1.13	Down	$p<0.025$
GALNT14	1.83	Down	$p<0.025$
GALNT18	1.21	Down	$p<0.025$
B3GNT3	1.42	Down	$p<0.025$
B3GNT6	10.97	Down	$p<0.025$
ST3GAL1	1.74	Down	$p<0.025$
ST3GAL2	1.38	Down	$p<0.025$
ST6GALNAC4	1.55	Down	$p<0.025$

6.4.11 EMT transcription factor DEGs

Many genes highlighted previously in KEGG functional enrichment analysis are implicit in the EMT process. Key EMT transcription factors are often dubbed the “master regulators of EMT” [354] and were therefore manually retrieved from DEGs lists for each comparison condition, MCF-7 HPA-negative compared to MCF-7 HPA-positive, T-47D HPA-positive compared to T-47D HPA-negative and MCF-7 TGF- β 1 treated compared to MCF-7 HPA-positive, see Table 6.7. SNAI2 (+1.61 fold-change, $p<0.025$) and ZEB1 (+1.45 fold-change, $p<0.025$) were significantly upregulated in MCF-7 HPA-negative cells. Whereas, in T-47D HPA-negative cells only ZEB1 was significantly upregulated (+1.54 fold-change, $p<0.025$). MCF-7 TGF- β 1 treated cells demonstrated significant upregulation of all Snail family members, SNAI1 (+1.41 fold-change, $p<0.05$), SNAI2 (+2.44 fold-change, $p<0.025$), and SNAI3 (+1.40 fold-change, $p<0.025$).

Table 6.7 Epithelial-mesenchymal transition (EMT)-related transcription factors differentially expressed genes (DEGs) identified in MCF-7 *Helix pomatia* agglutinin (HPA)-negative, T-47D HPA-negative and MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells. Regulation direction is HPA-negative or TGF- β 1 treated cells compared to relevant HPA-positive cell population. Values with ≥ 1.5 fold-change are identified by bold text; four biological replicates were conducted.

EMT transcription factor DEGs				
	Gene name	Fold-change	Regulation direction	pvalue
MCF-7 HPA-negative	SNAI2	1.61	Up	p<0.025
	ZEB1	1.45	Up	p<0.025
T-47D HPA-negative	ZEB1	1.54	Up	p<0.025
MCF-7 TGF- β 1	SNAI1	1.41	Up	p<0.05
	SNAI2	2.44	Up	p<0.025
	SNAI3	1.40	Up	p<0.025

6.5 Discussion

RNA sequencing was conducted to determine transcriptomic differences between HPA-positive and HPA-negative cells separated from MCF-7 and T-47D cell lines, and between TGF- β 1 treated and HPA-positive MCF-7 cells. These populations have distinct morphological and functional differences, associated with EMT and cell-surface glycosylation profiles as demonstrated in Chapters 3-5. RNA sequencing therefore allowed for comparisons between transcriptomes and identification of DEGs in mucin-type O-glycan biosynthetic pathways, which could explain the differing glycosylation profiles. It also highlighted DEGs associated with EMT, proteoglycans and stemness, which could explain the altered functionality demonstrated between cell populations. It shows the importance of the ECM, which is primarily comprised of proteoglycans, its interaction with signalling networks of cells, and glycosylation in the phenotypic switch from epithelial-like to mesenchymal-like (reviewed by [355]).

Adding to the body of evidence that MCF-7 and T-47D HPA-negative cells and MCF-7 TGF- β 1 treated cells demonstrate a mesenchymal-like phenotype induced by EMT, key transcription factors (Snail, TWIST and ZEB) that are deemed the “master regulators of EMT” [354], known to promote and induce EMT [356], were differentially expressed across all comparison groups, see Figure 1.16 for overview of transcriptional regulation of EMT. MCF-7 HPA-negative cells demonstrated significant upregulation of SNAI2 (+1.61 fold-change, p<0.025) and ZEB1 (+1.45 fold-change, p<0.025) compared to MCF-7 HPA-positive cells. Similarly, T-47D HPA-negative cells displayed significant upregulation of

ZEB1 compared to T-47D HPA-positive cells (+1.54 fold-change, $p < 0.025$). MCF-7 TGF- β 1 treated cells also demonstrated significant upregulation of all three of the Snail family transcription factors, SNAI1 (+1.41 fold-change, $p < 0.05$), SNAI2 (+2.44 fold-change, $p < 0.025$) and SNAI3 (+1.40 fold-change, $p < 0.025$). Single-cell sequencing of breast epithelial cells, MCF10A, undergoing EMT induced by TGF- β 1 stimulation (5ng/mL) have also previously demonstrated SNAI2 upregulation [357]. Snail transcription factors are widely regarded as playing an important role in TGF- β induced EMT, allowing cells to escape initial tumour suppressive effects of TGF- β and strongly repress E-cadherin (reviewed by [358]). These results show TGF- β 1 treatment used in this study was likely to activate the expected pathways to induce EMT.

6.5.1 MCF-7 HPA-negative DEGs associated with EMT

In the MCF-7 HPA-negative cell top ten most significant DEG list (fold-change of ≥ 1.5), Table 6.1, five have been explicitly linked to EMT in the literature previously. Due to phenotypic and functional differences of the HPA-negative cells, as previously discussed, this is unsurprising. DEGs of note included FN1, which encodes for fibronectin 1. It was the third most significant DEG in MCF-7 HPA-negative cells compared to MCF-7 HPA-positive cells (+1.87 fold-change, $p < 0.025$). Fibronectin is a glycoprotein that is a well-recognised mesenchymal marker of EMT, associated with focal adhesion, metastasis, increased cell migration and a component of the ECM that is not usually produced by normal adult breast tissue [334, 359, 360]. It was unsurprisingly, also identified in the proteoglycans in cancer KEGG, for which MCF-7 HPA-negative cells were significantly enriched compared to MCF-7 HPA-positive cells (64/2,043 DEGs, $p < 0.025$), see Figure 6.13 for overview across all comparison groups. Park and Schwarzbauer [361] demonstrated in MCF10A breast epithelial cells that interaction with exogenous fibronectin could induce an EMT response that resulted in increased cell migration; EMT was identified through upregulation of FN1, Snail, N-cadherin and MMP-2, and activation of Smad2. These effects could also be partially reversed when a TGF- β neutralising antibody was added, suggesting fibronectin has an enhancing effect of endogenous TGF- β in inducing EMT. Fibronectin has also previously been documented to be a key regulator of latent TGF- β -binding proteins and latent TGF- β 1 in the ECM [362]. It could therefore have a role in collating both signals, chemical, TGF- β 1, and mechanical force that can induce EMT [363]. Interestingly, FN1 was also significantly upregulated in MCF-7 TGF- β 1 treated cells compared to MCF-7 HPA-positive cells (+2.44 fold-change, $p < 0.025$).

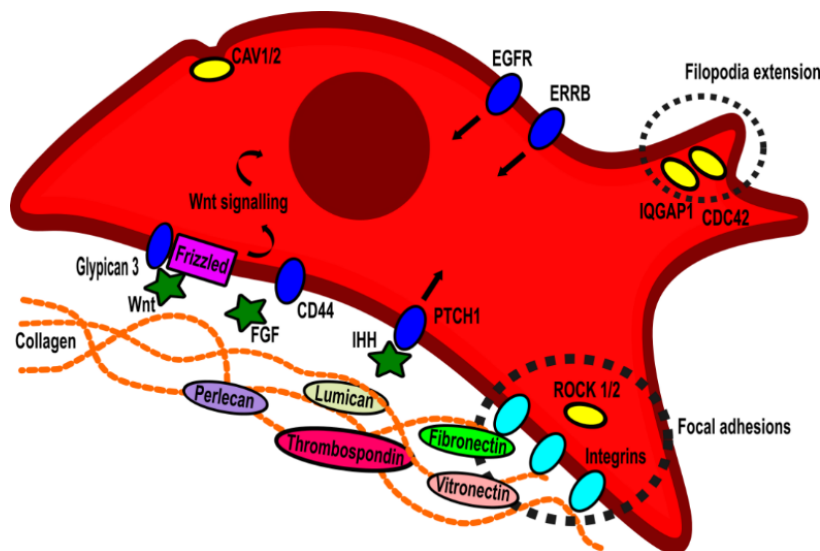


Figure 6.13 Significant differentially expressed genes (DEGs) associated with proteoglycans in cancer KEGG (hsa05205) highlighted from all comparison groups.

Another DEG of note, thrombospondin 1, encoded by THBS1, is a secreted extracellular matrix glycoprotein that is well-recognised to have a role in activating latent TGF- β 1 complexes, cell adhesion and migration [337, 364]. THBS1 was differentially expressed in MCF-7 HPA-negative cells compared to MCF-7 HPA-positive cells (+1.68 fold-change, $p < 0.025$), and was identified in the significantly enriched proteoglycans in cancer KEGG. Liu et al., [337] demonstrated that the siRNA knockdown of THBS1 in SW480 and LOVO colorectal cancer cell lines reduced cell migration in a wound-healing assay. Western Blot analysis of E-cadherin and vimentin of THBS1 knockdown cells revealed increased E-cadherin and decreased vimentin compared to control cells, indicating a role of THBS1 in EMT. Interestingly, THBS1 was also significantly upregulated in MCF-7 TGF- β 1 treated cells compared to MCF-7 HPA-positive cells (+2.25 fold-change, $p < 0.025$).

Cell migration inducing hyaluronidase 1, a hyaluronic acid binding protein, encoded for by CEMIP was significantly upregulated in MCF-7 HPA-negative cells compared to MCF-7 HPA-positive cells (+2.34 fold-change, $p < 0.025$). CEMIP has been linked to EMT through the EMT-related signalling pathways, such as the Wnt/ β -catenin/Snail pathway in many tumours [338, 365]. Liang et al., [366] demonstrated that knockdown of CEMIP using shRNA in colorectal cancer cell lines, HCT116 and SW480, resulted in reduced cell migration in a wound healing assay and also attenuated EMT and Wnt/ β -catenin/Snail signal transduction, as identified by upregulation of E-cadherin and decrease of Snail in Western Blotting. Colorectal cancer stem cells, identified by FACS for CD44⁺VE/CD45⁻VE cells from colorectal cancer metastases from the lymph node tissue, also demonstrated high levels of CEMIP compared to paired primary tumour tissue, identified by gene chip analysis and confirmed by quantitative RT-PCR [367].

Interestingly these findings are consistent with the results presented in this chapter, as MCF-7 HPA-negative cells which have DEG CEMIP also have significantly upregulated CD44 (+1.60 fold-change, $p < 0.025$) compared to HPA-positive cells. CD44 (+1.15 fold-change, $p < 0.025$) and CEMIP2 (+1.61 fold-change, $p < 0.025$) were also significantly upregulated in MCF-7 TGF- β 1 treated cells compared to MCF-7 HPA-positive cells. CD44, a transmembrane glycoprotein, considered a mesenchymal marker gene for cancer stem cells (reviewed by [368]), was identified in the significantly enriched signalling pathways regulating pluripotency of stem cells KEGG for MCF-7 HPA-negative versus MCF-7 HPA-positive cells (39/2,043 DEGs, $p < 0.05$) and MCF-7 TGF- β 1 treated cells compared to MCF-7 HPA-positive cells (62/3,589 DEGs, $p < 0.05$), see Figure 6.14 for overview across all comparison groups. The connection between EMT, which requires cells to demonstrate high levels of dynamic cellular plasticity, and stemness was first described by Mani et al., [369] in non-tumorigenic HMLE immortalised breast cells, whereby induction of EMT by ectopic expression of TWIST or Snail, resulted in a mesenchymal cell phenotype and also the expression of CD44.

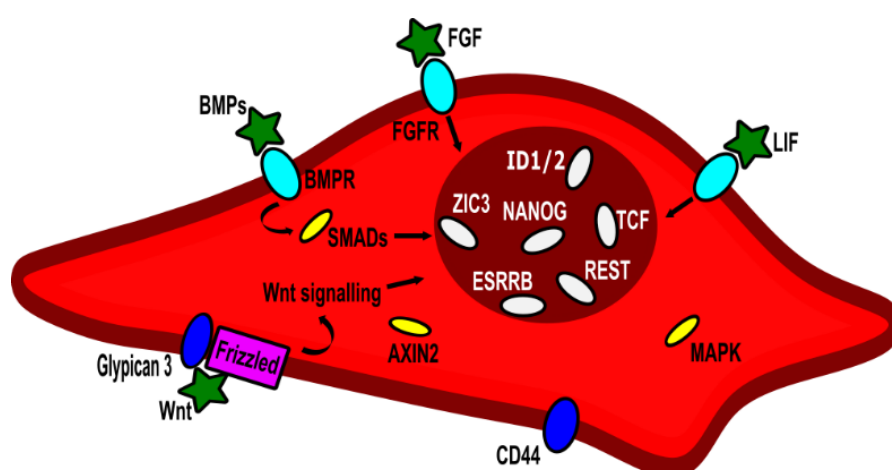


Figure 6.14 Significant differentially expressed genes (DEGs) associated with signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) highlighted from all comparison groups.

6.5.2 T-47D HPA-negative DEGs associated with EMT

Similarly, the top ten most significant DEG list (fold-change of ≥ 1.5) for T-47D HPA-negative cells also featured five genes that have been previously implicated with EMT in the literature, as described in Table 6.2. One DEG of note was ANXA1 which encodes for annexin A1, significantly upregulated in T-47D HPA-negative cells compared to T-47D HPA-positive cells (+2.97 fold-change, $p < 0.025$). Oshi et al., [339] utilised transcriptomes of pancreatic cancer samples from the Cancer Genome Atlas (TCGA) to show that those with elevated ANXA1

also demonstrated elevated EMT-related genes such as CDH1, CHD2, SNAI1, SNAI2 and TWIST. de Graauw et al., [370] demonstrated ANXA1 knockdown by shRNA in the MDA-MB-231 breast cancer cell line, which has high endogenous protein levels of ANXA1, resulted in a reduced mesenchymal-like phenotype and switch to an epithelial-like phenotype. Interestingly, luminal-like breast cancer cell lines, MCF-7 and T-47D, demonstrated very minimal protein levels of ANXA1. GFP-ANXA1 transfection of MCF-7 cells resulted in increased cell scattering and an EMT-like switch.

Another DEG of note was DSCAM-AS1, which encodes for long noncoding RNA DSCAM-AS1, demonstrated a significant upregulation in T-47D HPA-negative cells compared to T-47D HPA-positive cells (+7.13 fold-change, $p < 0.025$). Xu et al., [341] found DSCAM-AS1 to be significantly upregulated in HT29, LOVO, PKO and SW480 colorectal cancer cell lines compared to normal colon epithelial cell line NCM460. Knockdown of DSCAM-AS1 using siRNA resulted in the reduced migratory capability of cells in a transwell migration assay and also reduced vimentin and increased E-cadherin, assessed by immunofluorescence, indicating reduced EMT. DSCAM-AS1 was shown to interact with miR-317, which can suppress Notch-1 expression, confirming a potential role in EMT.

T-47D HPA-negative cells were significantly enriched for the signalling pathways regulating pluripotency of stem cells KEGG (17/577 DEGs, $p < 0.025$) in comparison to T-47D HPA-positive cells. Of note, the T-47D HPA-negative cells demonstrated significant upregulation of TCF7 (+1.76 fold-change, $p < 0.025$) compared to T-47D HPA-positive cells. TCF7 encodes for T-cell factor 1, a key transcription factor at the end of the Wnt/ β -catenin signalling pathway and therefore regulates target gene expression and ultimately cancer stemness [371].

6.5.3 TGF- β 1 treated MCF-7 DEGs associated with EMT

Unsurprisingly, treatment of MCF-7 cells with TGF- β 1, a known EMT-inducing cytokine, resulted in a large number of DEGs associated with EMT. All of the top ten most significant DEGs (fold-change of ≥ 1.5) in TGF- β 1 treated MCF-7 cells have been previously associated with EMT in the literature, as described in Table 6.3. The most significant DEG in MCF-7 TGF- β 1 treated cells compared to MCF-7 HPA-positive cells was FBN2 (+12.08 fold-change, $p < 0.025$), which encodes for the extracellular matrix protein fibrillin 2 which is involved in the formation of elastic fibres. Hong et al., [344] demonstrated that FBN2 mRNA and protein expression was markedly higher in PC-9 and H1640 lung cancer cell lines than in BEAS-2B normal lung epithelial cells. Knockdown of FBN2, by shRNA, resulted in significant inhibition of cell migration in a transwell migration assay, and reduced mRNA and protein levels of N-cadherin and vimentin, while E-cadherin was upregulated. These results

suggest a role of FBN2 in EMT and cell migration. Interestingly, FBN2 was also found to be significantly upregulated in MCF-7 HPA-negative cells compared to MCF-7 HPA-positive cells (+1.59 fold-change, $p < 0.05$).

CAV1, which encodes for caveolin-1, the main scaffolding protein component of caveolae found at the plasma membrane, was also significantly upregulated in TGF- β 1 treated MCF-7 cells compared to MCF-7 HPA-positive cells (+6.45 fold-change, $p < 0.025$). Caveolin-1 has been documented to be upregulated in many cancers, including breast, and implicated in influencing cell adhesion at focal adhesions, cell migration, EMT, and immune response suppression [372-374]. Bailey and Liu [350] treated a mouse mammary epithelial cell line, NMuMG, with TGF- β 1 (2ng/mL for 4 days) which induced EMT, assessed by cell morphology and Western Blotting for E-cadherin and N-cadherin. Caveolin-1 protein expression was also upregulated, detected by Western Blotting. The authors suggested that due to its synchronicity to morphology and EMT protein changes, the upregulation of caveolin-1 could be a product of EMT. Significant upregulation of CAV1, and also CAV2 (+2.42 fold-change, $p < 0.025$), encoding for caveolin-2 in the results presented in this chapter were identified in the proteoglycans in cancer KEGG analysis, which was significantly enriched in the MCF-7 TGF- β 1 treated cells (119/3,589 DEGs, $p < 0.025$). Vitronectin, another ECM glycoprotein differentially expressed in TGF- β 1 treated cells was identified in this KEGG. Vitronectin has a role in cell adhesion, acting through several ligands, including integrins, promoting focal adhesion sites and promoting metastasis [375, 376].

Several DEGs were also significantly downregulated in the MCF-7 TGF- β 1 treated cells compared to MCF-7 HPA-positive cells. For example, TSKU (-4.94 fold-change, $p < 0.025$), which encodes for Tsukushi, an ECM proteoglycan. Yamada et al., [346] overexpressed TSKU in H1975 non-small cell lung cancer cells and demonstrated a significant reduction of SNAI1, SNAI2 and vimentin compared to wild-type, detected at mRNA level using qRT-PCR and protein level using Western Blotting. When knocked out using CRISPR/Cas9, expression of SNAI1 and vimentin were significantly elevated, and knock-out cells also demonstrated higher invasion in a Matrigel invasion assay, suggesting TSKU may have a role in inhibiting EMT. The results of this chapter also demonstrated that SEMA3B was also significantly downregulated (-3.84 fold-change, $p < 0.025$) in TGF- β 1 treated MCF-7 cells. SEMA3B, which encodes for semaphorin 3B, is part of a large semaphorin family of 20 secreted, transmembrane and glycosylphosphatidylinositol (GPI)-linked glycoproteins, that can induce changes in cytoskeletal architecture, control integrin function and affect cell motility. They have also been associated with cell phenotype plasticity in EMT and MET (reviewed by [377]). SEMA3B has also been linked to a tumour-suppressive role, downregulating EMT transcription factors, and when lost or mutated can result in aggressive breast cancer development [349].

As previously mentioned, there was significant enrichment for signalling pathways regulating pluripotency of stem cells KEGG in TGF- β 1 treated cells compared to MCF-7 HPA-positive cells. This KEGG identified DEGs involved in the Wnt pathway, including six of the ten human frizzled genes that encode receptors for the Wnt signalling pathway [378] and also the BMP signalling pathway, see Supplementary Table 8 for an overview. These pathways have all been implicated in cancer stemness and EMT. Stemness is beneficial to motile, migratory cancer cells due to the association with immune suppression, treatment resistance and self-renewal properties, that can promote cancer metastasis (reviewed by [379]).

6.5.4 Mucin-type O-glycan biosynthesis DEGs

RNA sequencing also revealed that a variety of mucin-type O-glycan biosynthesis-associated genes were differentially expressed in MCF-7 HPA-negative, T-47D HPA-negative and MCF-7 TGF- β 1 treated cells. MCF-7 HPA-negative and T-47D HPA-negative cells demonstrated very similar DEGs. Two DEGs were significantly differentially expressed across all three groups compared to their respective HPA-positive cells, GALNT7 (MCF-7 HPA-negative +1.31 fold-change, $p < 0.025$, T-47D HPA-negative +1.25 fold-change, $p < 0.025$, MCF-7 TGF- β 1 treated +1.26 fold-change, $p < 0.025$) and ST6GALNAC4 (MCF-7 HPA-negative -1.32 fold-change, $p < 0.025$, T-47D HPA-negative -1.36 fold-change, $p < 0.025$, MCF-7 TGF- β 1 treated -1.55 fold-change, $p < 0.025$), see Table 6.4-6.6.

GALNT7 is a member of the GALNT family that initiates mucin-type O-linked glycosylation, known to be widely expressed across all tissues and categorised as a 'late' acting GALNT, preferring to act on areas of already high O-GalNAc glycosylation (reviewed by [54]). Mockl et al., [380] demonstrated that TGF- β 1 treatment (5ng/mL for one week) of MCF10A normal breast epithelial cells, EMT induction confirmed by morphology assessment, and Western Blotting of E-cadherin and vimentin, resulted in increased glycocalyx height compared to untreated MCF10A cells, inferring increased glycan length. Glycocalyx height was determined by using metabolically labelled GalNAc and sialic acid (on the basis that sialic acid is the distal point of the glycan) and super-resolution microscopy. Interestingly, when GALNT7 was knocked down, using siRNA in iKras pancreatic cells, the height of the glycocalyx was reduced. GALNT7 could be a potential mediator of cell surface glycosylation and glycocalyx height. Hence, correlating to the results of this study, overexpression could result in the elongation of the truncated O-GalNAc glycans seen in HPA-positive cell populations. Peng et al., [381] demonstrated that knockdown of GALNT7 using siRNA in HeLa and C33A cervical cancer cells significantly inhibited cell migration and invasion in transwell migration assays. GALNT7 upregulation could therefore play a key role in

enhancing the migratory ability of a cell, which could be linked to its role in determining glycan length in the glycocalyx.

ST6GALNAC4 adds sialic acid to galactose-containing substrates to create the α 2-6 sialyl-core 1 and disialyl-core 1 glycans, which are thought to play a role in immune suppression [382, 383]. Results presented in this thesis demonstrate downregulation of ST6GALNAC4 in MCF-7 HPA-negative, T-47D-HPA negative and MCF-7 TGF- β 1 treated cells compared to their relative HPA-positive comparison group, suggesting that there could potentially be more core 1 and α 2-3 sialyl-core 1 structures being created instead of the truncated O-GalNAc glycans seen in HPA-positive cell populations. Contrary to these results, Man et al., [384] demonstrated in HCCLM3 and MHCC97H hepatocellular carcinoma cells that ST6GALNAC4 knockdown using siRNA, reduced cell proliferation, migration in wound healing assays and invasion in transwell migration assays. It also resulted in reduced levels of TGF- β R2, N-cadherin and phosphorylated Smad2/3 protein levels, whilst protein levels of E-cadherin increased. ST6GALNAC4 overexpression resulted in the opposite effect and the authors, therefore, concluded that overexpression of ST6GALNAC4 could drive high levels of TGF- β R2 which promotes proliferation, migration and invasion. There are no reports of the functional effect of ST6GALNAC4 expression in breast cancer cells that could be found in the literature but, theoretically, ST6GALNAC4 could be competing for the Tn antigen with C1GALT1, which was significantly upregulated, discussed below.

Both MCF-7 HPA-negative (+1.29 fold-change, $p < 0.025$) and T-47D HPA-negative cells (+1.35 fold-change, $p < 0.025$) demonstrated significant upregulation of C1GALT1 (creates core 1 glycans) compared to their HPA-positive counterparts. Like the Tn antigen, core 1 (T antigen) overexpression is also regarded as a tumour-associated carbohydrate antigen [385]. C1GALT1 is frequently upregulated in many tumour types, including breast [386], often on glycoproteins such as CD44, MUC1 and integrins (reviewed by [385], [387] and [388]). C1GALT1 has been implicated in EMT by modifying O-GalNAc glycans on key EMT receptors such as FGFR2 (reviewed by [389]). Chou et al., [390] demonstrated that overexpression of C1GALT1 in MCF-7 cells resulted in increased cell migration in a transwell migration assay and invasion in a Matrigel assay. It also led to a decrease in VVA binding, assessed by flow cytometry, indicating less cell surface Tn antigen. When PNA pulldown was conducted there was also evidence of increased core 1 on MUC1.

Additionally, in MCF-7 HPA-negative cells, there was significant upregulation of the sialyltransferases ST3GAL1 (+1.24 fold-change, $p < 0.025$) and ST3GAL2 (+1.40 fold-change, $p < 0.025$) compared to MCF-7 HPA-positive cells. ST3GAL1 and ST3GAL2 catalyse the addition of sialic acid to galactose to form sialyl-core 1, and sialylation terminates chain extension. Hypersialylation is common in cancers and results in increased negatively charged sialylated glycans on cell surfaces, which has been shown to promote cancer metastasis through immune evasion and enhance cell migration and invasion

(reviewed by [391]). ST3GAL1 overexpression in A2780 ovarian cancer cells resulted in increased cell growth, migration and invasion, detected using transwell migration assays. EMT was induced using TGF- β 1 treatment (10ng/mL) and ST3GAL1 mRNA expression was found to be significantly higher than in untreated cells. EMT could also be inhibited by knockdown of ST3GAL1 using shRNA [267]. In glioblastoma cancer tissue samples, *in situ* hybridisation revealed that ST3GAL1 labelling was most extensive at the 'leading edge' of invasive tumours and in mesenchymal-like cells. mRNA expression of TGF- β and phosphorylated Smad2 was linearly correlated with ST3GAL1 expression [392]. Interestingly, work presented in this chapter demonstrates that TGF- β 1 treatment of MCF-7 cells resulted in the downregulation of ST3GAL1 (-1.74 fold-change, $p < 0.025$) and ST3GAL2 (-1.38 fold-change, $p < 0.025$) compared to MCF-7 HPA-positive cells.

TGF- β 1 treated MCF-7 cells demonstrated multiple DEGs compared to MCF-7 HPA-positive cells. In addition to elevated GALNT7 previously discussed, there was also significant upregulation of GALNT1, -T10, -T11, -T12 and -T16. Interestingly, TGF- β 1 treated cells also demonstrated significantly upregulated GCNT4 (+1.83 fold-change, $p < 0.025$), required for creating core 2 glycans. There is minimal literature on the role of GCNT4 in cancer progression [393]. However, Sun et al., [394] demonstrated that GCNT4 mRNA was significantly downregulated in gastric cancer tissue samples compared to their paired adjacent normal mucosa, and that GCNT4 overexpression could repress cell proliferation and cell cycle in SGC-7901 and MGC-803 gastric cancer cells. Hu et al., [395] subsequently demonstrated that miR-130a-3p could bind to and downregulate GCNT4, and enhance MKN45 and AGS gastric cell migration in wound healing assays and invasion in a transwell migration assay and also upregulate protein levels of TGF- β 1 and phosphorylated Smad. Lan et al., [396] identified that miR-485-3p could also bind to and downregulate GCNT4 in gastric cancer cells. Overexpression of GCNT4 reduced cell proliferation, migration and invasion. Studies into the other GCNT family members, such as GCNT2, have shown that overexpression in breast cancer cell lines is linked to increased cell migration, invasion and EMT [397].

The results presented in this chapter suggest that mucin-type O-linked glycan KEGG-related gene expression could be significantly different between HPA-binding populations in a mixed cell population and between TGF- β 1 treated cells and HPA-positive cells. Many of the DEGs have been associated with changes in cell behaviour, cancer progression and metastasis; and, notably, with increased sialylation. Elongation of Tn antigen in MCF-7 and T-47D HPA-positive cells into core 1 and sialyl-core 1 glycans is a fast, dynamic change that could provide cells with rapid plasticity that allows for alterations in cell behaviour, see Figure 6.15. TGF- β 1 treatment of MCF-7 cells also indicates through its DEGs an elongation of the glycan chain, but to core 2 structures instead, see Figure 6.16. Future work would be required to validate this, which could include qPCR validation of DEGs and lectin cytochemistry or mass spectrometry to reveal altered glycosylation profiles in HPA-negative

cells.

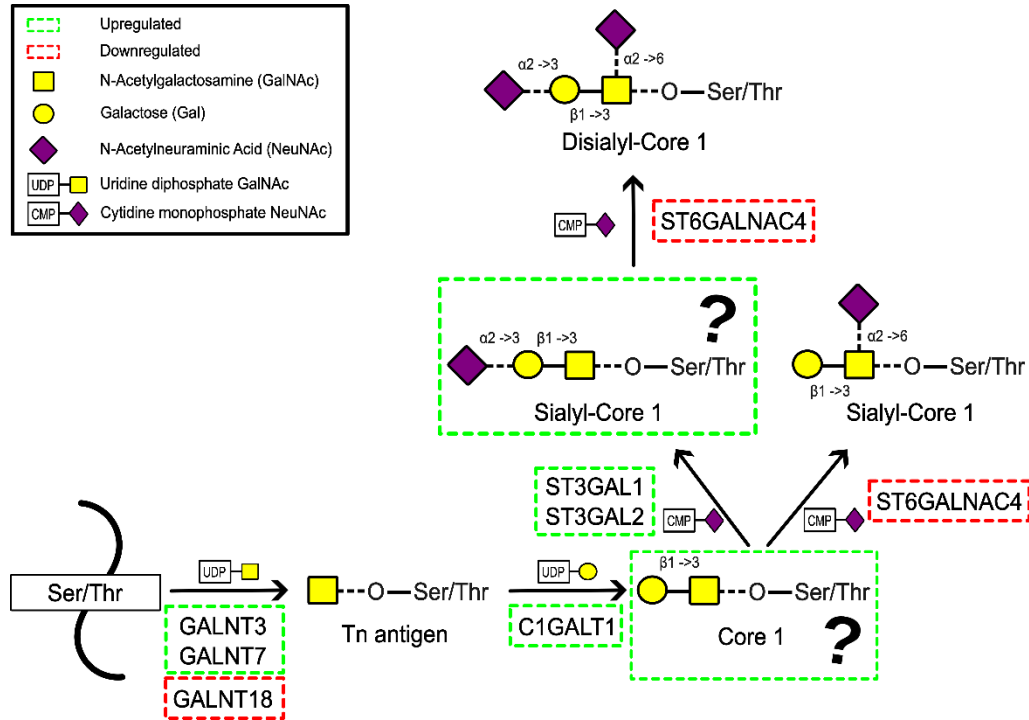


Figure 6.15 Differentially expressed genes (DEGs) in mucin-type O-glycan biosynthesis in MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells and T-47D HPA-negative cells compared to their HPA-positive counterpart.

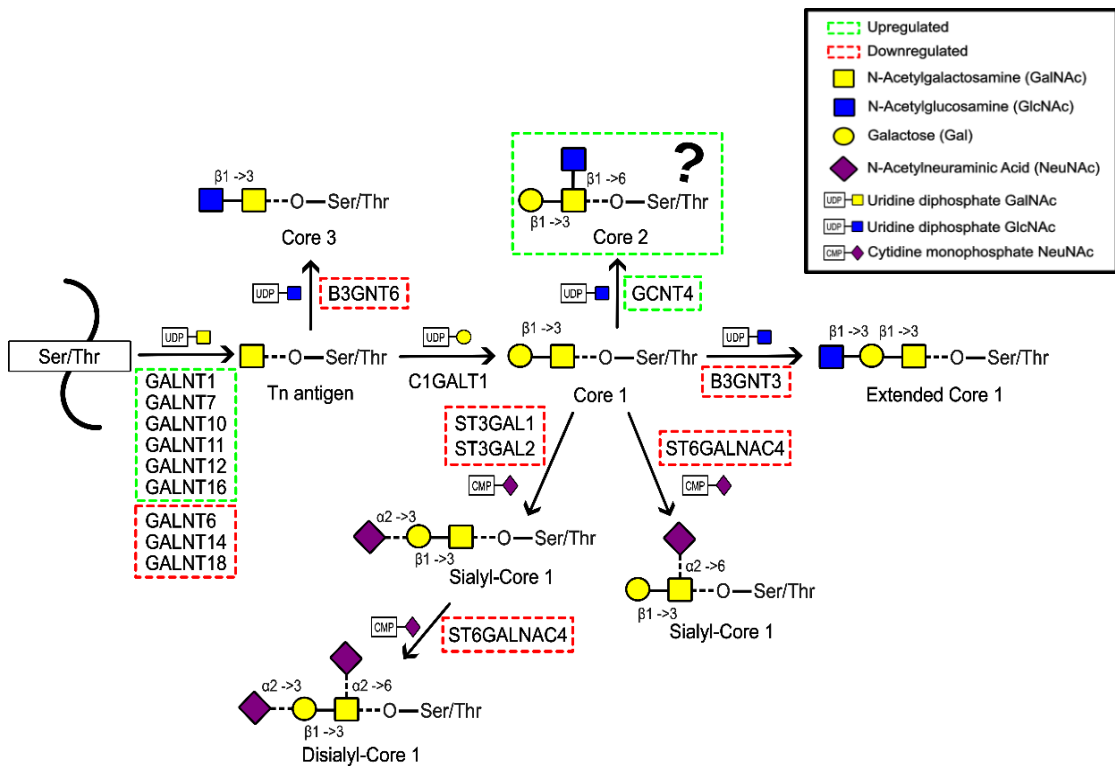


Figure 6.16 Differentially expressed genes (DEGs) in mucin-type O-glycan biosynthesis in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 *Helix pomatia* agglutinin (HPA)-positive cells.

6.6 Key findings

- EMT-related genes, relating to mesenchymal markers, EMT signalling pathways and EMT master regulator transcription factors, were differentially expressed in HPA-negative and TGF- β 1 treated cell populations compared to their HPA-positive counterparts.
- HPA-negative MCF-7 cell DEGs featured upregulation of C1GALT1, ST3GAL1 and ST3GAL2 which could indicate that Tn antigen is being elongated into core 1 and sialyl-core 1 glycans.
- HPA-negative T-47D cell DEGs featured upregulation of C1GALT1 which could indicate that Tn antigen is being elongated into core 1 glycans.
- TGF- β 1 treated MCF-7 cells DEGs featured upregulation of GCNT4 and down-regulation of B3GNT6, B3GNT3, ST3GAL1, ST3GAL2, ST6GALNAC4, which could indicate that Tn antigen is being elongated into core 2 glycans.
- DEG enrichment analysis demonstrates that HPA-negative and TGF- β 1 treated cells featured enrichment for signalling pathways associated with cancer stem cells and altered proteoglycans, it also highlights the importance of ECM interactions.

Chapter 7

Final discussion and future work

7. Final discussion

The glycome is critically important for a range of cellular functions including cell adhesion, receptor activation, signal transduction and molecular trafficking of proteins. Altered glycosylation is implicated in a range of pathologies, such as cancer, and is a feature of EMT, cancer stemness and metastasis. Typical glycan changes in cancer include increased exposure of incomplete or truncated glycans (e.g. Tn antigen), increased sialylation (e.g. sTn antigen and sialyl-core 1) and neo-synthesis of *de novo* glycans (e.g. SLe^a and SLe^x). However, despite significant research efforts, understanding of the regulation and functional implications of altered glycosylation remains incomplete. This is in part due to the complexity of glycosylation, involving a myriad of glycosylation-related machinery, with some enzymes demonstrating overlapping specificities and functional roles, and the wide range of potential regulatory mechanisms that can take effect from the transcription of glyco genes through to the enzyme localisation and the structure and environment of the Golgi apparatus.

In this thesis, the O-GalNAc glycosylation profiles were characterised, and the functional and morphological implications of glycosylation were explored, *in vitro* using a range of model breast epithelial cell lines. This analysis revealed that model breast epithelial cell lines demonstrated a heterogeneous population of cells with differing cell surface glycans profiles, which were associated with the morphology of the cell, EMT marker protein expression, and functionality of the cell, notably cell migration and adhesion capabilities, which are of importance in EMT and metastasis. The analysis also revealed that TGF- β 1 treatment, a known inducer of EMT *in vitro*, of two breast cancer epithelial cell lines resulted in similar cell morphology, EMT protein expression and reduced cell surface O-GalNAc glycans in a cell phenotype that was strikingly similar to cells from a mixed cell line population that did not bind with HPA. Preliminary RNA sequencing results indicated some potential alternative cell surface glycans for HPA-nonbinding cells and a further role, confirmation and explanation of how altered glycosylation profiles can impact the functionality of the cells. It revealed DEGs that were enriched for biological processes including cell adhesion, ECM modification, and stemness, all of which are of importance in EMT and metastasis. The work in this thesis suggests that epithelial breast cancer cells can potentially utilise, and access, differing cell surface glycan profiles, representative of differing cell phenotypes, in a dynamic and highly plastic manner, that also reflects altered transcriptomes, aiding them in becoming successful at each step of the metastatic cascade. This is of importance due to the differing functionality that is required to be successful at each step, see Figure 7.1.

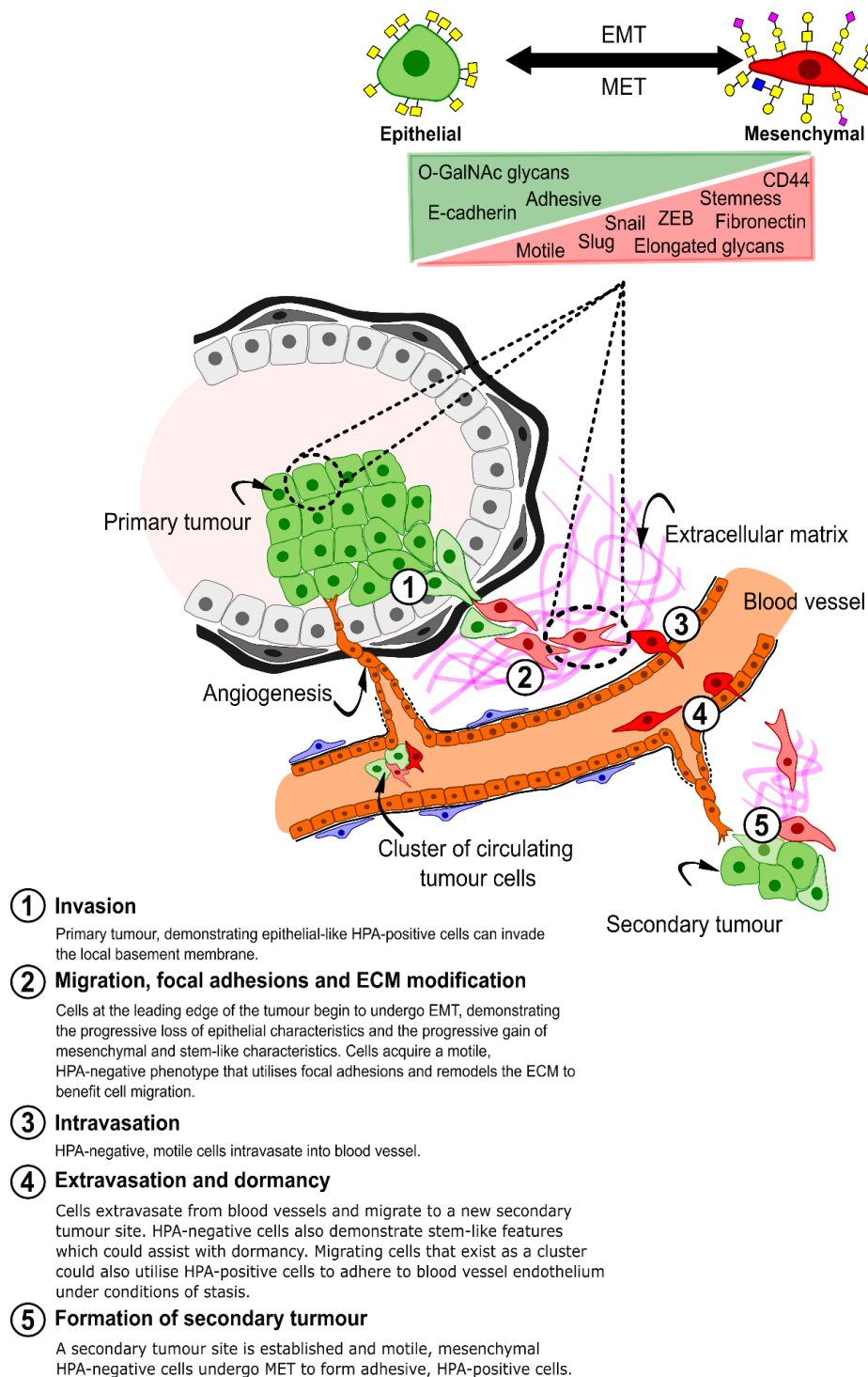


Figure 7.1 Metastatic cascade: proposed cell and glycan dynamics during the epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET). (1) Primary tumour, demonstrating epithelial-like *Helix pomatia* agglutinin (HPA)-positive cells. (2) Cells at the leading edge of the tumour begin to undergo EMT. Cells acquire a motile, HPA-negative phenotype that utilises focal adhesions and remodels the ECM to benefit cell migration. (3) HPA-negative, motile cells intravasate into blood vessel. (4) Cells extravasate from blood vessels and migrate to a new secondary tumour site. HPA-negative cells also demonstrate stem-like features. Migrating cells that exist as a cluster could also utilise HPA-positive cells to adhere to blood vessel endothelium under conditions of stasis. (5) A secondary tumour site is established and motile, mesenchymal HPA-negative cells undergo MET to form adhesive, HPA-positive cells.

7.1 O-GalNAc glycan profiles of model breast epithelial cell lines

As described in Chapter 3, it is well established that cancer cells feature altered glycosylation in the form of truncated glycans [179, 398]. Since the first descriptions by Leatham and Brooks [182] of HPA binding, which detects O-GalNAc glycans, predicting metastasis and poor prognosis in primary breast cancers, various breast cancer cell lines have also been characterised as synthesising HPA-binding glycoproteins, consistent with their origin [39, 164, 165]. The results presented in this thesis demonstrated HPA-binding profiles of model breast epithelial cell lines consistent with those reported for the same cell lines in the literature by Brooks et al., [164], Schumacher et al., [165], and unpublished work from our laboratory group. This includes the existence of a heterogeneous mixed cell population of HPA-binding and HPA-nonbinding cells within a single cell line, and that cancer cell lines, particularly those derived from cancer metastases, demonstrated significantly more HPA labelling than immortalised normal breast epithelial cells. A heterogeneous profile of HPA-binding cells has been shown previously in both the primary tumours and metastases originating from HPA-binding breast cancers [220, 221]. One explanation for this, suggests that the O-GalNAc glycosylation profile of cancer cells must be dynamic and highly plastic and that there could be a functional role as cancers have retained both phenotypes.

In addition to the exploration of cell glycosylation through HPA labelling, an alternative O-GalNAc binding lectin, VVA, was also employed in the experiments reported here. VVA demonstrates nominal specificity for the Tn antigen and demonstrated broadly comparable labelling to that of HPA. VVA labelling of the model cell lines was consistent with their origins [206, 207] and what has been reported previously in the literature [208]. Consistent with observations using HPA labelling, there appears to be a heterogeneous cell population of VVA-binding and -nonbinding cells that exist in both the primary tumour and metastases [204, 222]. As in the study reported here, the VVA and HPA labelling results were indistinguishable, it might be surmised that HPA is predominantly recognising common glycan structures to VVA, specifically, O-GalNAc in Tn antigen.

The mixed profile of cell surface glycans, indicated by the existence of both HPA/VVA-binding and -nonbinding cells in a cell line or clinical tumour, is an intriguing feature of tumour heterogeneity. Tumour heterogeneity often represents one of the most challenging behavioural aspects of cancers as it is associated with metastatic competence and drug-resistant behaviours (reviewed by [399]). Interestingly, tumour heterogeneity often arises from interactions with the tumour microenvironment and ECM. RNA sequencing results presented in this thesis highlighted that HPA-nonbinding cells demonstrated significant DEGs that were associated with ECM, proteoglycans and stem-like properties, which could play a role in providing a stimulus for increasing tumour heterogeneity, induction

of EMT and increasing migratory capabilities associated with aggressive metastatic behaviours.

TGF- β 1 treatment of MCF-7 and T-47D cells resulted in the significant reduction of cell surface O-GalNAc glycans, and an elongated mesenchymal-like phenotype. Interestingly, in both cell lines, the whole cell population did not become HPA-negative upon TGF- β 1 treatment; a small population of cells, or individual cells, retained cell surface O-GalNAc glycans. These results show for the first time that the glycosylation profile of breast cancer cells is affected by the TGF- β 1 cytokine. It also demonstrates the important link between an EMT state and changes in cell surface glycosylation, as these cells also showed reduced E-cadherin protein. This is important, as cells at the edge of a tumour are more likely to interact with cytokines such as TGF- β 1 in the ECM due to their positioning.

Previous literature suggests that glycosylation is also regulated by a wide range of factors ranging from transcriptional control [329], mutational events [75], posttranslational control [65], relocation of glycosyltransferases within the secretory pathway [71, 400] and changes in the fundamental structure and environment of the Golgi apparatus [84, 88]. RNA sequencing in this thesis revealed preliminary evidence that could suggest that differential expression of genes encoding for glycosyltransferases could lead to altered glycosylation profiles of the cells. MCF-7 HPA-negative cells demonstrated significant upregulation of C1GALT1, ST3GAL1 and ST3GAL2, which would be consistent with Tn antigen being elongated into core 1 and sialyl-core 1 glycans. T-47D HPA-negative cells showed upregulation of C1GALT1 which would be consistent with Tn antigen being elongated into core 1 glycans. Finally, TGF- β 1 treated MCF-7 cells DEGs showed upregulation of GCNT4 and down-regulation of B3GNT6, B3GNT3, ST3GAL1, ST3GAL2, ST6GALNAC4 which would be consistent with Tn antigen being elongated into core 2 glycans. Transcriptome changes in these cells could provide an interesting future direction for research, as it could provide transcriptional regulatory targets to prevent altered migratory capability of cells.

Interestingly, increased expression of mRNA encoding for enzymes associated with sialylated glycans, such as ST3GAL1 [267, 401-403], and core 2 glycans, such as GCNT3 [404], have previously been associated with EMT, increased cell migration and invasion capabilities. Altered sialylation has consistently been associated with metastatic cell behaviours including invasive capabilities, enhanced cell survival and cell migration, with most research investigating sTn and SLe^x [267, 405]. As sialic acid is negatively charged at physiologic pH it has the potential to alter protein conformation or oligomerization events. It can also regulate the retention of cell surface receptors and therefore, could play an important role in mediating EMT and cancer progression through influencing signal transduction (reviewed by [405]). As glycans are added to proteins during the biosynthetic pathway it is thought that their presence is relatively stable however, there is evidence that suggests soluble neuraminidases, which remove α -glycosidically linked sialic acid, can exist

at the cell surface membrane and can therefore alter glycans temporally [405-407]. This could provide a mechanism that allows sialylated glycans to be dynamically regulated on the cell surface and could explain how breast epithelial cancer cells in the experiments described in this thesis can access both cell glycosylation phenotypes in such a dynamic way during cell migration.

7.2 Functional implications of mucin-type O-linked glycosylation profiles

Morphological analysis revealed that both HPA and VVA-binding cells demonstrated a significantly more rounded, epithelial-like phenotype, while HPA and VVA-negative cells demonstrated significantly elongated mesenchymal-like phenotype morphology. Wound healing assays reported here revealed, for the first time, that HPA-negative cells were predominant at the 'leading edge' of the 'wound' as it was 'healing' through a series of timepoints. There was then a reversion back to a mixed population of HPA-positive and -negative cells when the 'wound' was 'fully healed'. This, for the first time, indicates that HPA-negative cells exhibit a more motile phenotype. These HPA-negative cells at the 'leading edge' of the healing 'wound' were also shown to demonstrate significantly less epithelial protein marker, E-cadherin, as the 'wound' was 'healing', and there was a reversion to no significant difference when the 'wound' was 'fully healed'.

This further confirms that HPA-negative, elongated, motile cells exhibit an EMT or pEMT state, as they feature loss of E-cadherin, a protein referred to as the "gatekeeper of the epithelial phenotype". Its loss, results in the dissolution of adherens junctions and the acquisition of an elongated, motile phenotype [121]. Preliminary RNA sequencing results, discussed in Chapter 6, provided additional evidence, as both HPA-negative T-47D and MCF-7 cells demonstrated significant upregulation of two key transcription factors, SNAI2 and ZEB1, that represent two families of the "master regulators of EMT" [354] and are known to promote EMT [356]. Both Snail and ZEB act to suppress transcription of E-cadherin and induce transcription of mesenchymal markers such as N-cadherin, vimentin and fibronectin [287, 288].

Although in the experiments reported in this thesis, mesenchymal markers N-cadherin and vimentin were not detectable at the protein level, RNA sequencing revealed that there was significant upregulation of fibronectin in both MCF-7 HPA-negative cells and MCF-7 TGF- β 1 treated cells. The glycoprotein fibronectin was not directly investigated in the experiments described in this thesis but, could be a target of future work. Importantly, the fibronectin glycoprotein is well-recognised as a marker for EMT and is involved in regulating focal adhesions and facilitating increased cell migration [334, 359, 360]. Similarly, it has also been

demonstrated to play a crucial role in wound healing (reviewed by [408]). Interestingly, O-GalNAc glycosylation has been reported to interfere with intracellular degradation of fibronectin, thus providing a stabilizing role for the glycoprotein in human breast cancer cell lines [241] and therefore promoting EMT [409]. The lack of detectable N-cadherin or vimentin in the experiments reported here could be a result of incomplete or partial activation of signalling pathways and associated transcription factors, which have been associated with a pEMT state. pEMT, which can involve cells existing in a wide spectrum of phenotypes, is more common than a full binary switch into a mesenchymal phenotype and allows for a highly dynamic switch in phenotype throughout EMT and MET (reviewed by [123] and [289]).

The pEMT state is also consistent with the type of cell migration modelled in the wound healing assays used in the experiments reported in this thesis, which model collective cell migration rather than single-cell migration. Interestingly, SNAI2, encoding for Slug, which was identified in the work described in this thesis as significantly upregulated in MCF-7 HPA-negative and MCF-7 TGF- β 1 treated cells, has been consistently associated with mediating a pEMT state and collective migration (reviewed by [410]). 2D collective cell migration allows for coordinated movement, whereby cells retain a level of intercellular connection, through maintenance of some cell-cell junctions, collective polarity and are led by select 'leader' cells. Collective migration is considered the predominant migration method of cancers *in vivo* [411, 412] and is thought to allow faster, and more directed cell migration (reviewed by [413]). Whereas, single-cell migration is a feature of a 'full EMT' state. Single-cell migration is also found in amoeboid migration which demonstrates blebbing and weak adhesions [414, 415]. Depending on environmental stimuli, it is believed that cancers can switch between single-cell and collective-cell migration [412, 416].

Importantly, collective cell migration is not limited to cancer, it also occurs in wound healing and embryonic morphogenesis (reviewed by [417]). These collective cell migration cell characteristics are visible in the wound healing assays reported in Chapter 4, whereby most cells did not fully lose E-cadherin protein however, the loss was most pronounced in 'leader' cells. HPA-negative cells, that predominate the leading edge of the healing 'wound', were significantly enriched for genes associated with focal adhesions and ECM components in the RNA sequencing reported in Chapter 6. Focal adhesions are documented to be present in the 'leader cells' and allow cells to generate traction forces to pull the 'follower cells' behind them [413].

pEMT has also been associated with aggressive disease phenotypes, acquisition of cancer stem-cell-like features and therapeutic resistance [225, 418]. Donnenberg et al., [419] demonstrated that breast cancer stem cell markers CD44 and CD90 were localised to the tumour periphery, indicating a potential link between the acquisition of stemness traits and 'leader cells' at the tumour edge. Papadaki et al., [420] showed that circulating tumour cells from metastatic breast cancer patients, demonstrated a pEMT phenotype that was also high

in cancer stem cell markers. Similarly, in the RNA sequencing reported in Chapter 6, MCF-7 HPA-negative cells and MCF-7 TGF- β 1 treated cells (which are predominantly HPA-negative) in the RNA sequencing also demonstrated significant upregulation of CD44. MCF-7 HPA-negative, T-47D HPA-negative and MCF-7 TGF- β 1 treated cells also demonstrated significant enrichment for the 'signalling pathways regulating pluripotency of stem cells' KEGG. Therefore, results reported in this thesis could provide potential additional evidence that links a pEMT state with cancer stem cell features, which in turn could aid cancer cells in metastasis.

The static adhesion assays reported in this thesis revealed that cell lines featuring a higher proportion of HPA-positive cells were generally more adherent to an endothelial monolayer than an epithelial cell line with low levels of HPA-binding glycans. This demonstrates a potentially important functional role of O-GalNAc glycans in cell adhesion when cells come to arrest in a capillary bed. The role of truncated O-GalNAc glycans has also previously been implicated in cancer cell adhesion to endothelium under conditions of sweeping flow, mimicking cells moving in the bloodstream. Masking O-GalNAc glycans by preincubating the cells with HPA lectin resulted in a significant decrease in cell adhesion to the endothelial monolayer [172]. The functional results are consistent with observable cell morphology and reveal a potential dynamic link between O-GalNAc glycosylation and functionality in both EMT and MET.

7.3 Study limitations

It is important to recognise the limitations of several approaches taken in the experiments reported in this thesis. Lectins can bind to a wide range of structures with differing binding affinities, so it is difficult to determine which structures are being identified definitively. However, HPA and VVA binding to cancer has been consistently associated with clinical outcomes in the literature. The experiments reported here were also only conducted using breast epithelial cell lines. Another limitation is that the experiments were conducted on 2D *in vitro* cell cultures, more advanced 3D, or co-culture cell culture techniques may provide results that are more representative of an *in vivo* tumour microenvironment. Finally, preliminary RNA sequencing data needs further validation using qPCR and changes to cell surface glycosylation would also need to be confirmed.

7.4 Future work

Future work to further these results should include the use of other epithelial cancer types to confirm if these findings translate beyond breast cancer. Validation of the RNA sequencing results and proposed alterations in cell surface glycans should also be investigated. This could be done through qPCR of chosen genes of interest and by utilising various lectins, such as MAA (detects α 2-3 linked sialic acid attached to galactose), SNA (detects α 2-6 linked sialic acid attached to galactose) and PNA (detects core 1 glycans) to assess binding patterns in lectin cytochemistry or flow cytometry. This could provide novel insights into what the cell surface glycosylation profile is for HPA-negative cells. It could also yield novel targets for the development of therapies aimed at inhibiting the migratory capabilities of cancers. There may also be aspects of this work that are translatable to wound healing or embryo morphogenesis research.

7.5 Novel contributions

The novel contributions of this work towards the understanding of mucin-type O-linked glycosylation in cancer, metastasis and EMT presented here are:

- HPA-positive cells, which exhibit cell surface O-GalNAc glycans, are significantly more rounded, suggestive of an epithelial-like phenotype. HPA-negative cells are significantly more elongated, suggestive of a mesenchymal-like phenotype.
- VVA-positive cells, which exhibit cell surface Tn antigen, are significantly more rounded, suggestive of an epithelial-like phenotype. VVA-negative cells are significantly more elongated, suggestive of a mesenchymal-like phenotype.
- HPA-positive MCF-7 and T-47D breast cancer cells feature significantly more E-cadherin (epithelial marker) protein than their HPA-negative cell counterparts. This difference approaches significance in hTERT-HME1 cells ($p=0.06$).
- There is no significant difference in vimentin protein expression between HPA-binding populations in MCF-7, T-47D and hTERT-HME1 cell lines.
- MCF-7 and T-47D HPA-negative cells are concentrated at the 'leading edge' of a healing 'wound' in a wound healing assay. There were significantly fewer O-GalNAc

glycans on cells at the 'leading edge' of the healing 'wound' at 8 and 24 hours in MCF-7 cells and 24 and 72 hours in T-47D cells, as detected by HPA binding. When the MCF-7 cell line 'wound' is fully healed there is a reversion to a mixed population of HPA-positive and -negative cells, with no significant difference in levels of O-GalNAc glycans.

- In a wound healing assay, MCF-7 and T-47D cells at the 'leading edge' of the healing 'wound' that feature significantly less O-GalNAc glycans also demonstrate significantly reduced E-cadherin protein expression. This was significant at 4, 8 and 24 hours in MCF-7 cells, and at 24 and 48 hours in T-47D cells.
- TGF- β 1 treatment of MCF-7 and T-47D cells resulted in a significant reduction of O-GalNAc glycans on the cell surface of cells at the 7-day timepoint. This effect was also visually apparent but not significant at all preceding timepoints investigated for each cell line.
- In RNA sequencing analysis, EMT-related genes, relating to mesenchymal markers and EMT signalling pathways, were differentially expressed in HPA-negative and TGF- β 1 treated cell populations compared to their HPA-positive counterparts. DEG enrichment analysis demonstrates that HPA-negative and TGF- β 1 treated cells show enrichment for signalling pathways associated with cancer stem cells and altered proteoglycans.
- MCF-7 HPA-negative cell DEGs show upregulation of C1GALT1, ST3GAL1 and ST3GAL2, consistent with the elongation of Tn antigen into core 1 and sialyl-core 1 glycans.
- T-47D HPA-negative cell DEGs show upregulation of C1GALT1, consistent with the elongation of Tn antigen into core 1 glycans.
- TGF- β 1 treated MCF-7 cells DEGs show upregulation of GCNT4 and down-regulation of B3GNT6, B3GNT3, ST3GAL1, ST3GAL2, ST6GALNAC4, consistent with the elongation of Tn antigen into core 2 glycans.

References

8. References

1. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 1999;1473(1):4-8.
2. Smith BAH, Bertozzi CR. The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans. *Nature Reviews Drug Discovery*. 2021;20(3):217-43.
3. Stanley P. Golgi glycosylation. *Cold Spring Harbor Perspectives in Biology*. 2011;3(4).
4. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell*. 2006;126(5):855-67.
5. Joshi HJ, Hansen L, Narimatsu Y, Freeze HH, Henrissat B, Bennett E, et al. Glycosyltransferase genes that cause monogenic congenital disorders of glycosylation are distinct from glycosyltransferase genes associated with complex diseases. *Glycobiology*. 2018;28(5):284-94.
6. Wuhler M, Deelder AM, Hokke CH. Protein glycosylation analysis by liquid chromatography-mass spectrometry. *Journal of Chromatography B Analyt Technol Biomed Life Sci*. 2005;825(2):124-33.
7. Freeze HH, Boyce M, Zachara NE, Hart GW, Schnaar RL. Glycosylation Precursors. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. *Essentials of Glycobiology*. 4th ed. Cold Spring Harbor (NY)2022. p. 53-66.
8. Conte F, van Buuringen N, Voermans NC, Lefeber DJ. Galactose in human metabolism, glycosylation and congenital metabolic diseases: Time for a closer look. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 2021;1865(8):129898.
9. Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Marth JD, et al. Symbol nomenclature for glycan representation. *Proteomics*. 2009;9(24):5398-9.
10. Brooks SA, Dwek MV, Schumacher U. *Functional and Molecular Glycobiology*. Oxford, New York: BIOS Scientific; Distributed in the U.S. by Springer-Verlag; 2002.
11. Brockhausen I, Wandall HH, Hagen KGT, Stanley P. O-GalNAc Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. *Essentials of Glycobiology*. 4th ed. Cold Spring Harbor (NY)2022. p. 117-28.
12. Zachara NE, Akimoto Y, Boyce M, Hart GW. The O-GlcNAc Modification. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. *Essentials of Glycobiology*. 4th ed. Cold Spring Harbor (NY)2022. p. 251-64.
13. Praissman JL, Wells L. Mammalian O-mannosylation pathway: glycan structures, enzymes, and protein substrates. *Biochemistry*. 2014;53(19):3066-78.
14. Holdener BC, Haltiwanger RS. Protein O-fucosylation: structure and function. *Current Opinion in Structural Biology*. 2019;56:78-86.
15. Yu H, Takeuchi H. Protein O-glucosylation: another essential role of glucose in biology. *Current Opinion in Structural Biology*. 2019;56:64-71.
16. Brooks SA, Carter TM, Royle L, Harvey DJ, Fry SA, Kinch C, et al. Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies. *Anti-Cancer Agents in Medicinal Chemistry*. 2008;8(1):2-21.
17. Snider MD, Rogers OC. Intracellular movement of cell surface receptors after endocytosis: resialylation of asialo-transferrin receptor in human erythroleukemia cells. *The Journal of Cell Biology*. 1985;100(3):826-34.
18. Litvinov SV, Hilken J. The epithelial sialomucin, episialin, is sialylated during recycling. *Journal of Biological Chemistry*. 1993;268(28):21364-71.
19. Breitling J, Aebi M. N-linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harbor Perspectives in Biology*. 2013;5(8):a013359.
20. Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. *Nature Reviews Nephrology*. 2019;15(6):346-66.
21. Rush JS. Role of Flippases in Protein Glycosylation in the Endoplasmic Reticulum. *Lipid Insights*. 2015;8(Suppl 1):45-53.

22. Stanley P, Moremen KW, Lewis NE, Taniguchi N, Aebi M. N-Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. *Essentials of Glycobiology*. 4th ed. Cold Spring Harbor (NY)2022. p. 103-16.
23. Van den Steen P, Rudd PM, Dwek RA, Opdenakker G. Concepts and principles of O-linked glycosylation. *Critical Reviews in Biochemistry and Molecular Biology*. 1998;33(3):151-208.
24. Wang S, Chen C, Gadi MR, Saikam V, Liu D, Zhu H, et al. Chemoenzymatic modular assembly of O-GalNAc glycans for functional glycomics. *Nature Communications*. 2021;12(1):3573.
25. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO Journal*. 2013;32(10):1478-88.
26. Weissmann B, Hinrichsen DF. Mammalian alpha-acetylgalactosaminidase. Occurrence, partial purification, and action on linkages in submaxillary mucins. *Biochemistry*. 1969;8(5):2034-43.
27. Dahr W, Uhlenbruck G, Gunson HH, Van Der Hart M. Molecular basis of Tn-polyagglutinability. *Vox Sanguinis*. 1975;29(1):36-50.
28. Roth J. Cytochemical localization of terminal N-acetyl-D-galactosamine residues in cellular compartments of intestinal goblet cells: implications for the topology of O-glycosylation. *The Journal of Cell Biology*. 1984;98(2):399-406.
29. Ten Hagen KG, Fritz TA, Tabak LA. All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology*. 2003;13(1):1R-16R.
30. Strous GJ. Initial glycosylation of proteins with acetylgalactosaminylserine linkages. *Proceedings of the National Academy of Sciences*. 1979;76(6):2694-8.
31. Karlsson NG, Herrmann A, Karlsson H, Johansson ME, Carlstedt I, Hansson GC. The glycosylation of rat intestinal Muc2 mucin varies between rat strains and the small and large intestine. A study of O-linked oligosaccharides by a mass spectrometric approach. *Journal of Biological Chemistry*. 1997;272(43):27025-34.
32. Iwai T, Inaba N, Naundorf A, Zhang Y, Gotoh M, Iwasaki H, et al. Molecular cloning and characterization of a novel UDP-GlcNAc:GalNAc-peptide beta1,3-N-acetylglucosaminyltransferase (beta 3Gn-T6), an enzyme synthesizing the core 3 structure of O-glycans. *Journal of Biological Chemistry*. 2002;277(15):12802-9.
33. Breg J, Van Halbeek H, Vliegthart JF, Klein A, Lamblin G, Roussel P. Primary structure of neutral oligosaccharides derived from respiratory-mucus glycoproteins of a patient suffering from bronchiectasis, determined by combination of 500-MHz 1H-NMR spectroscopy and quantitative sugar analysis. 2. Structure of 19 oligosaccharides having the GlcNAc beta(1----3)GalNAc-ol core (type 3) or the GlcNAc beta(1----3)[GlcNAc beta(1----6)]GalNAc-ol core (type 4). *European Journal of Biochemistry*. 1988;171(3):643-54.
34. Robbe C, Capon C, Coddeville B, Michalski JC. Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. *Biochemical Journal*. 2004;384(Pt 2):307-16.
35. Madunic K, Mayboroda OA, Zhang T, Weber J, Boons GJ, Morreau H, et al. Specific (sialyl-)Lewis core 2 O-glycans differentiate colorectal cancer from healthy colon epithelium. *Theranostics*. 2022;12(10):4498-512.
36. Hounsell EF, Lawson AM, Feeney J, Gooi HC, Pickering NJ, Stoll MS, et al. Structural analysis of the O-glycosidically linked core-region oligosaccharides of human meconium glycoproteins which express oncofoetal antigens. *European Journal of Biochemistry*. 1985;148(2):367-77.
37. Chai WG, Hounsell EF, Cashmore GC, Rosankiewicz JR, Bauer CJ, Feeney J, et al. Neutral oligosaccharides of bovine submaxillary mucin. A combined mass spectrometry and 1H-NMR study. *European Journal of Biochemistry*. 1992;203(1-2):257-68.
38. van Halbeek H, Strang AM, Lhermitte M, Rahmoune H, Lamblin G, Roussel P. Structures of monosialyl oligosaccharides isolated from the respiratory mucins of a non-secretor (O, Lea+b-) patient suffering from chronic bronchitis. Characterization of a novel type of mucin carbohydrate core structure. *Glycobiology*. 1994;4(2):203-19.
39. Brooks SA, Hall DM. Investigations into the potential role of aberrant N-acetylgalactosamine glycans in tumour cell interactions with basement membrane components. *Clinical & Experimental Metastasis*. 2002;19(6):487-93.

40. Wandall HH, Nielsen MAI, King-Smith S, de Haan N, Bagdonaite I. Global functions of O-glycosylation: promises and challenges in O-glycobiology. *The FEBS Journal*. 2021;288(24):7183-212.
41. Ju T, Cummings RD. A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. *Proceedings of the National Academy of Sciences*. 2002;99(26):16613-8.
42. Kudo T, Iwai T, Kubota T, Iwasaki H, Takayma Y, Hiruma T, et al. Molecular cloning and characterization of a novel UDP-Gal:GalNAc(alpha) peptide beta 1,3-galactosyltransferase (C1Gal-T2), an enzyme synthesizing a core 1 structure of O-glycan. *Journal of Biological Chemistry*. 2002;277(49):47724-31.
43. Ju T, Lanneau GS, Gautam T, Wang Y, Xia B, Stowell SR, et al. Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. *Cancer Research*. 2008;68(6):1636-46.
44. Sun X, Zhan M, Sun X, Liu W, Meng X. C1GALT1 in health and disease. *Oncology Letters*. 2021;22(2):589.
45. Hanisch FG. O-glycosylation of the mucin type. *Biological Chemistry*. 2001;382(2):143-9.
46. Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*. 2012;22(6):736-56.
47. de Las Rivas M, Lira-Navarrete E, Gerken TA, Hurtado-Guerrero R. Polypeptide GalNAc-Ts: from redundancy to specificity. *Current Opinion in Structural Biology*. 2019;56:87-96.
48. Fritz TA, Hurley JH, Trinh LB, Shiloach J, Tabak LA. The beginnings of mucin biosynthesis: the crystal structure of UDP-GalNAc:polypeptide alpha-N-acetylgalactosaminyltransferase-T1. *Proceedings of the National Academy of Sciences*. 2004;101(43):15307-12.
49. Rottger S, White J, Wandall HH, Olivo JC, Stark A, Bennett EP, et al. Localization of three human polypeptide GalNAc-transferases in HeLa cells suggests initiation of O-linked glycosylation throughout the Golgi apparatus. *Journal of Cell Science*. 1998;111 (Pt 1):45-60.
50. Roth J, Wang Y, Eckhardt AE, Hill RL. Subcellular localization of the UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-mediated O-glycosylation reaction in the submaxillary gland. *Proceedings of the National Academy of Sciences*. 1994;91(19):8935-9.
51. Paulson JC, Colley KJ. Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *Journal of Biological Chemistry*. 1989;264(30):17615-8.
52. Imberty A, Piller V, Piller F, Breton C. Fold recognition and molecular modeling of a lectin-like domain in UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Protein Engineering*. 1997;10(12):1353-6.
53. Kubota T, Shiba T, Sugioka S, Furukawa S, Sawaki H, Kato R, et al. Structural basis of carbohydrate transfer activity by human UDP-GalNAc: polypeptide alpha-N-acetylgalactosaminyltransferase (pp-GalNAc-T10). *Journal of Molecular Biology*. 2006;359(3):708-27.
54. Beaman EM, Carter DRF, Brooks SA. GALNTs: master regulators of metastasis-associated epithelial-mesenchymal transition (EMT)? *Glycobiology*. 2022;32(7):556-79.
55. Bennett EP, Weghuis DO, Merkx G, van Kessel AG, Eiberg H, Clausen H. Genomic organization and chromosomal localization of three members of the UDP-N-acetylgalactosamine: polypeptide N-acetylgalactosaminyltransferase family. *Glycobiology*. 1998;8(6):547-55.
56. Hagen FK, Nehrke K. cDNA cloning and expression of a family of UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase sequence homologs from *Caenorhabditis elegans*. *Journal of Biological Chemistry*. 1998;273(14):8268-77.
57. Ten Hagen KG, Tran DT. A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is essential for viability in *Drosophila melanogaster*. *Journal of Biological Chemistry*. 2002;277(25):22616-22.
58. Homa FL, Hollander T, Lehman DJ, Thomsen DR, Elhammer AP. Isolation and expression of a cDNA clone encoding a bovine UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase. *Journal of Biological Chemistry*. 1993;268(17):12609-16.

59. White T, Bennett EP, Takio K, Sorensen T, Bonding N, Clausen H. Purification and cDNA cloning of a human UDP-N-acetyl- α -D-galactosamine:polypeptide N-acetylgalactosaminyltransferase. *Journal of Biological Chemistry*. 1995;270(41):24156-65.
60. Yu X, Du Z, Sun X, Shi C, Zhang H, Hu T. Aberrant Cosmc genes result in Tn antigen expression in human colorectal carcinoma cell line HT-29. *International Journal of Clinical and Experimental Pathology*. 2015;8(3):2590-602.
61. Thomas D, Sagar S, Caffrey T, Grandgenett PM, Radhakrishnan P. Truncated O-glycans promote epithelial-to-mesenchymal transition and stemness properties of pancreatic cancer cells. *Journal of Cellular and Molecular Medicine*. 2019;23(10):6885-96.
62. Vavasseur F, Dole K, Yang J, Matta KL, Myerscough N, Corfield A, et al. O-glycan biosynthesis in human colorectal adenoma cells during progression to cancer. *European Journal of Biochemistry*. 1994;222(2):415-24.
63. An G, Wei B, Xia B, McDaniel JM, Ju T, Cummings RD, et al. Increased susceptibility to colitis and colorectal tumors in mice lacking core 3-derived O-glycans. *The Journal of Experimental Medicine*. 2007;204(6):1417-29.
64. Neelamegham S, Mahal LK. Multi-level regulation of cellular glycosylation: from genes to transcript to enzyme to structure. *Current Opinion in Structural Biology*. 2016;40:145-52.
65. Thu CT, Mahal LK. Sweet Control: MicroRNA Regulation of the Glycome. *Biochemistry*. 2020;59(34):3098-110.
66. Stewart N, Wisnovsky S. Bridging Glycomics and Genomics: New Uses of Functional Genetics in the Study of Cellular Glycosylation. *Frontiers in Molecular Biosciences*. 2022;9:934584.
67. Pangen RP, Channathodiyil P, Huen DS, Eagles LW, Johal BK, Pasha D, et al. The GALNT9, BNC1 and CCDC8 genes are frequently epigenetically dysregulated in breast tumours that metastasise to the brain. *Clinical Epigenetics*. 2015;7(1):57.
68. Vojta A, Samarzija I, Bockor L, Zoldos V. Glyco-genes change expression in cancer through aberrant methylation. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 2016;1860(8):1776-85.
69. Mi R, Song L, Wang Y, Ding X, Zeng J, Lehoux S, et al. Epigenetic silencing of the chaperone Cosmc in human leukocytes expressing tn antigen. *Journal of Biological Chemistry*. 2012;287(49):41523-33.
70. Shi C, Xu X, Yu X, Du Z, Luan X, Liu D, et al. CD3/CD28 dynabeads induce expression of tn antigen in human t cells accompanied by hypermethylation of the cosmc promoter. *Molecular Immunology*. 2017;90:98-105.
71. Gill DJ, Chia J, Senewiratne J, Bard F. Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes. *The Journal of Cell Biology*. 2010;189(5):843-58.
72. Gill DJ, Tham KM, Chia J, Wang SC, Steentoft C, Clausen H, et al. Initiation of GalNAc-type O-glycosylation in the endoplasmic reticulum promotes cancer cell invasiveness. *Proceedings of the National Academy of Sciences*. 2013;110(34):E3152-61.
73. Chia J, Tham KM, Gill DJ, Bard-Chapeau EA, Bard FA. ERK8 is a negative regulator of O-GalNAc glycosylation and cell migration. *Elife*. 2014;3:e01828.
74. Bard F, Chia J. Cracking the Glycome Encoder: Signaling, Trafficking, and Glycosylation. *Trends in Cell Biology*. 2016;26(5):379-88.
75. Chia J, Goh G, Bard F. Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 2016;1860(8):1623-39.
76. Gomes J, Mereiter S, Magalhaes A, Reis CA. Early GalNAc O-Glycosylation: Pushing the Tumor Boundaries. *Cancer Cell*. 2017;32(5):544-5.
77. Nguyen AT, Chia J, Ros M, Hui KM, Saltel F, Bard F. Organelle Specific O-Glycosylation Drives MMP14 Activation, Tumor Growth, and Metastasis. *Cancer Cell*. 2017;32(5):639-53 e6.
78. Chia J, Tay F, Bard F. The GalNAc-T Activation (GALA) Pathway: Drivers and markers. *PLoS One*. 2019;14(3):e0214118.
79. Ros M, Nguyen AT, Chia J, Le Tran S, Le Guezennec X, McDowall R, et al. ER-resident oxidoreductases are glycosylated and trafficked to the cell surface to promote matrix degradation by tumour cells. *Nature Cell Biology*. 2020;22(11):1371-81.

80. Kingsley DM, Kozarsky KF, Hobbie L, Krieger M. Reversible defects in O-linked glycosylation and LDL receptor expression in a UDP-Gal/UDP-GalNAc 4-epimerase deficient mutant. *Cell*. 1986;44(5):749-59.
81. Stanley P, Sundaram S, Sallustio S. A subclass of cell surface carbohydrates revealed by a CHO mutant with two glycosylation mutations. *Glycobiology*. 1991;1(3):307-14.
82. Kumamoto K, Goto Y, Sekikawa K, Takenoshita S, Ishida N, Kawakita M, et al. Increased expression of UDP-galactose transporter messenger RNA in human colon cancer tissues and its implication in synthesis of Thomsen-Friedenreich antigen and sialyl Lewis A/X determinants. *Cancer Research*. 2001;61(11):4620-7.
83. Koike T, Kimura N, Miyazaki K, Yabuta T, Kumamoto K, Takenoshita S, et al. Hypoxia induces adhesion molecules on cancer cells: A missing link between Warburg effect and induction of selectin-ligand carbohydrates. *Proceedings of the National Academy of Sciences*. 2004;101(21):8132-7.
84. Zhang X, Wang Y. Glycosylation Quality Control by the Golgi Structure. *Journal of Molecular Biology*. 2016;428(16):3183-93.
85. Bui S, Mejia I, Diaz B, Wang Y. Adaptation of the Golgi Apparatus in Cancer Cell Invasion and Metastasis. *Frontiers in Cell and Developmental Biology*. 2021;9:806482.
86. Fullekrug J, Nilsson T. Protein sorting in the Golgi complex. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 1998;1404(1-2):77-84.
87. Weiss M, Nilsson T. Protein sorting in the Golgi apparatus: a consequence of maturation and triggered sorting. *FEBS Letters*. 2000;486(1):2-9.
88. Axelsson MA, Karlsson NG, Steel DM, Ouwendijk J, Nilsson T, Hansson GC. Neutralization of pH in the Golgi apparatus causes redistribution of glycosyltransferases and changes in the O-glycosylation of mucins. *Glycobiology*. 2001;11(8):633-44.
89. Hassinen A, Pujol FM, Kokkonen N, Pieters C, Kihlstrom M, Korhonen K, et al. Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells. *Journal of Biological Chemistry*. 2011;286(44):38329-40.
90. Goldstein IJ, Hughes RC, Monsigny M, Osawa T, Sharon N. What should be called a lectin? *Nature*. 1980;285(5760):66-.
91. Ambrosi M, Cameron NR, Davis BG. Lectins: tools for the molecular understanding of the glycode. *Organic & Biomolecular Chemistry*. 2005;3(9):1593-608.
92. Roth J. Lectins for histochemical demonstration of glycans. *Histochemistry and Cell Biology*. 2011;136(2):117-30.
93. Stillmark H. Über Ricin, ein giftiges Ferment aus den Samen von Ricinus comm. L. und einigen anderen Euphorbiaceen. Thesis, Medical Faculty, University of Dorpat (Tartu, Estonia)1888.
94. Radhakrishnan A, Park K, Kwak I-S, Jaabir M, Sivakamavalli J. Classification of Lectins. In: Elumalai P, Lakshmi S, editors. *Lectins: Innate immune defense and Therapeutics*. Singapore: Springer Singapore; 2021. p. 51-72.
95. Brooks S. Lectins as versatile tools to explore cellular glycosylation. *European Journal of Histochemistry*. 2024;68(1).
96. Tsaneva M, Van Damme EJM. 130 years of Plant Lectin Research. *Glycoconjugate Journal*. 2020;37(5):533-51.
97. Biswas SK, Banerjee S, Baker GW, Kuo CY, Chowdhury I. The Mammary Gland: Basic Structure and Molecular Signaling during Development. *International Journal of Molecular Sciences*. 2022;23(7).
98. Cancer Research UK. Breast cancer statistics 2019 [Cited 06.08.2024]. Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer>.
99. Ju T, Otto VI, Cummings RD. The Tn antigen-structural simplicity and biological complexity. *Angewandte Chemie International Edition*. 2011;50(8):1770-91.
100. Prokop O, Uhlenbruck G. N-acetyl-D-galactosamine in tumor cell membranes: demonstration by means of Helix agglutinins. *Med Welt*. 1969;46:2515-9.
101. Springer GF. Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. *Journal of Molecular Medicine*. 1997;75(8):594-602.
102. Cornelissen LA, Van Vliet SJ. A Bitter Sweet Symphony: Immune Responses to Altered O-glycan Epitopes in Cancer. *Biomolecules*. 2016;6(2).

103. Rømer TB, Aasted MKM, Dabelsteen S, Groen A, Schnabel J, Tan E, et al. Mapping of truncated O-glycans in cancers of epithelial and non-epithelial origin. *British Journal of Cancer*. 2021;125(9):1239-50.
104. Brooks SA. The involvement of Helix pomatia lectin (HPA) binding N-acetylgalactosamine glycans in cancer progression. *Histology and Histopathology*. 2000;15(1):143-58.
105. Hanisch FG, Baldus SE. The Thomsen-Friedenreich (TF) antigen: a critical review on the structural, biosynthetic and histochemical aspects of a pancarcinoma-associated antigen. *Histology and Histopathology*. 1997;12(1):263-81.
106. Kontomanolis EN, Koutras A, Syllaios A, Schizas D, Mastoraki A, Garmpis N, et al. Role of Oncogenes and Tumor-suppressor Genes in Carcinogenesis: A Review. *Anticancer Research*. 2020;40(11):6009-15.
107. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.
108. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-74.
109. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discovery*. 2022;12(1):31-46.
110. Fares J, Fares MY, Khachfe HH, Salhab HA, Fares Y. Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduction and Targeted Therapy*. 2020;5(1):28.
111. Gerstberger S, Jiang Q, Ganesh K. Metastasis. *Cell*. 2023;186(8):1564-79.
112. Hynes RO, Naba A. Overview of the matrisome--an inventory of extracellular matrix constituents and functions. *Cold Spring Harbor Perspectives in Biology*. 2012;4(1):a004903.
113. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Molecular & Cellular Proteomics*. 2012;11(4):M111 014647.
114. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nature Communications*. 2020;11(1):5120.
115. Cabral-Pacheco GA, Garza-Veloz I, Castruita-De la Rosa C, Ramirez-Acuna JM, Perez-Romero BA, Guerrero-Rodriguez JF, et al. The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases. *International Journal of Molecular Sciences*. 2020;21(24).
116. Greenburg G, Hay ED. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. *The Journal of Cell Biology*. 1982;95(1):333-9.
117. Lee JM, Dedhar S, Kalluri R, Thompson EW. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *The Journal of Cell Biology*. 2006;172(7):973-81.
118. Hugo H, Ackland ML, Blick T, Lawrence MG, Clements JA, Williams ED, et al. Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. *Journal of Cellular Physiology*. 2007;213(2):374-83.
119. Micalizzi DS, Farabaugh SM, Ford HL. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. *Journal of Mammary Gland Biology and Neoplasia*. 2010;15(2):117-34.
120. Nakajima S, Doi R, Toyoda E, Tsuji S, Wada M, Koizumi M, et al. N-cadherin expression and epithelial-mesenchymal transition in pancreatic carcinoma. *Clinical Cancer Research*. 2004;10(12 Pt 1):4125-33.
121. Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Developmental cell*. 2008;14(6):818-29.
122. Frisch SM, Schaller M, Cieply B. Mechanisms that link the oncogenic epithelial-mesenchymal transition to suppression of anoikis. *Journal of Cell Science*. 2013;126(Pt 1):21-9.
123. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nature Reviews Molecular Cell Biology*. 2019;20(2):69-84.
124. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation*. 2009;119(6):1420-8.

125. Marconi GD, Fonticoli L, Rajan TS, Pierdomenico SD, Trubiani O, Pizzicannella J, et al. Epithelial-Mesenchymal Transition (EMT): The Type-2 EMT in Wound Healing, Tissue Regeneration and Organ Fibrosis. *Cells*. 2021;10(7).
126. Luond F, Sugiyama N, Bill R, Bornes L, Hager C, Tang F, et al. Distinct contributions of partial and full EMT to breast cancer malignancy. *Developmental Cell*. 2021;56(23):3203-21 e11.
127. Bakir B, Chiarella AM, Pitarresi JR, Rustgi AK. EMT, MET, Plasticity, and Tumor Metastasis. *Trends in cell biology*. 2020;30(10):764-76.
128. Liao C, Wang Q, An J, Long Q, Wang H, Xiang M, et al. Partial EMT in Squamous Cell Carcinoma: A Snapshot. *International Journal of Biological Sciences*. 2021;17(12):3036-47.
129. Jordan NV, Johnson GL, Abell AN. Tracking the intermediate stages of epithelial-mesenchymal transition in epithelial stem cells and cancer. *Cell Cycle*. 2011;10(17):2865-73.
130. Hong T, Watanabe K, Ta CH, Villarreal-Ponce A, Nie Q, Dai X. An *Ovol2-Zeb1* Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States. *PLoS Computational Biology*. 2015;11(11):e1004569.
131. Lundgren K, Nordenskjold B, Landberg G. Hypoxia, Snail and incomplete epithelial-mesenchymal transition in breast cancer. *British Journal of Cancer*. 2009;101(10):1769-81.
132. Umbreit C, Flanjak J, Weiss C, Erben P, Aderhold C, Faber A, et al. Incomplete epithelial-mesenchymal transition in p16-positive squamous cell carcinoma cells correlates with beta-catenin expression. *Anticancer Research*. 2014;34(12):7061-9.
133. Lu M, Jolly MK, Levine H, Onuchic JN, Ben-Jacob E. MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proceedings of the National Academy of Sciences*. 2013;110(45):18144-9.
134. Sampson VB, David JM, Puig I, Patil PU, de Herreros AG, Thomas GV, et al. Wilms' tumor protein induces an epithelial-mesenchymal hybrid differentiation state in clear cell renal cell carcinoma. *PLoS One*. 2014;9(7):e102041.
135. Jia D, Jolly MK, Boareto M, Parsana P, Mooney SM, Pienta KJ, et al. OVOL guides the epithelial-hybrid-mesenchymal transition. *Oncotarget*. 2015;6(17):15436-48.
136. Schliekelman MJ, Taguchi A, Zhu J, Dai X, Rodriguez J, Celiktas M, et al. Molecular portraits of epithelial, mesenchymal, and hybrid States in lung adenocarcinoma and their relevance to survival. *Cancer Research*. 2015;75(9):1789-800.
137. Steinway SN, Zanudo JGT, Michel PJ, Feith DJ, Loughran TP, Albert R. Combinatorial interventions inhibit TGFbeta-driven epithelial-to-mesenchymal transition and support hybrid cellular phenotypes. *NPJ systems biology and applications*. 2015;1:15014.
138. Sinha D, Saha P, Samanta A, Bishayee A. Emerging Concepts of Hybrid Epithelial-to-Mesenchymal Transition in Cancer Progression. *Biomolecules*. 2020;10(11).
139. Bednarz-Knoll N, Alix-Panabieres C, Pantel K. Plasticity of disseminating cancer cells in patients with epithelial malignancies. *Cancer and Metastasis Reviews*. 2012;31(3-4):673-87.
140. Chebouti I, Kasimir-Bauer S, Buderath P, Wimberger P, Hauch S, Kimmig R, et al. EMT-like circulating tumor cells in ovarian cancer patients are enriched by platinum-based chemotherapy. *Oncotarget*. 2017;8(30):48820-31.
141. Nam MW, Kim CW, Choi KC. Epithelial-Mesenchymal Transition-Inducing Factors Involved in the Progression of Lung Cancers. *Biomolecules & Therapeutics*. 2022;30(3):213-20.
142. Krakhmal NV, Zavyalova MV, Denisov EV, Vtorushin SV, Perelmuter VM. Cancer Invasion: Patterns and Mechanisms. *Acta Naturae*. 2015;7(2):17-28.
143. Wu JS, Jiang J, Chen BJ, Wang K, Tang YL, Liang XH. Plasticity of cancer cell invasion: Patterns and mechanisms. *Translational Oncology*. 2021;14(1):100899.
144. Friedl P. Prespecification and plasticity: shifting mechanisms of cell migration. *Current Opinion in Cell Biology*. 2004;16(1):14-23.
145. Paňková K, Rösel D, Novotný M, Brábek J. The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cellular and Molecular Life Sciences*. 2010;67(1):63-71.
146. Lorentzen A, Bamber J, Sadok A, Elson-Schwab I, Marshall CJ. An ezrin-rich, rigid uropod-like structure directs movement of amoeboid blebbing cells. *Journal of Cell Science*. 2011;124(Pt 8):1256-67.

147. Tozluoğlu M, Tournier AL, Jenkins RP, Hooper S, Bates PA, Sahai E. Matrix geometry determines optimal cancer cell migration strategy and modulates response to interventions. *Nature Cell Biology*. 2013;15(7):751-62.
148. Palecek SP, Loftus JC, Ginsberg MH, Lauffenburger DA, Horwitz AF. Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness. *Nature*. 1997;385(6616):537-40.
149. Lauffenburger DA, Horwitz AF. Cell migration: a physically integrated molecular process. *Cell*. 1996;84(3):359-69.
150. Graziani V, Rodriguez-Hernandez I, Maiques O, Sanz-Moreno V. The amoeboid state as part of the epithelial-to-mesenchymal transition programme. *Trends in Cell Biology*. 2022;32(3):228-42.
151. Friedl P, Locker J, Sahai E, Segall JE. Classifying collective cancer cell invasion. *Nature Cell Biology*. 2012;14(8):777-83.
152. Haeger A, Wolf K, Zegers MM, Friedl P. Collective cell migration: guidance principles and hierarchies. *Trends in Cell Biology*. 2015;25(9):556-66.
153. Bronsert P, Enderle-Ammour K, Bader M, Timme S, Kuehs M, Csanadi A, et al. Cancer cell invasion and EMT marker expression: a three-dimensional study of the human cancer-host interface. *The Journal of Pathology*. 2014;234(3):410-22.
154. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences*. 2003;100(7):3983-8.
155. Bisht S, Nigam M, Kunjwal SS, Sergey P, Mishra AP, Sharifi-Rad J. Cancer Stem Cells: From an Insight into the Basics to Recent Advances and Therapeutic Targeting. *Stem Cells International*. 2022;2022:9653244.
156. Velasco-Velazquez MA, Homsí N, De La Fuente M, Pestell RG. Breast cancer stem cells. *The International Journal of Biochemistry & Cell Biology*. 2012;44(4):573-7.
157. Crabtree JS, Miele L. Breast Cancer Stem Cells. *Biomedicines*. 2018;6(3).
158. Walcher L, Kistenmacher AK, Suo H, Kitte R, Dluczek S, Strauss A, et al. Cancer Stem Cells-Origins and Biomarkers: Perspectives for Targeted Personalized Therapies. *Frontiers in immunology*. 2020;11:1280.
159. Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, et al. Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Research*. 2009;69(4):1302-13.
160. Liu H, Patel MR, Prescher JA, Patsialou A, Qian D, Lin J, et al. Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. *Proceedings of the National Academy of Sciences*. 2010;107(42):18115-20.
161. Volovetz J, Berezovsky AD, Alban T, Chen Y, Lauko A, Aranjuez GF, et al. Identifying conserved molecular targets required for cell migration of glioblastoma cancer stem cells. *Cell Death & Disease*. 2020;11(2):152.
162. Setoguchi T, Taga T, Kondo T. Cancer stem cells persist in many cancer cell lines. *Cell Cycle*. 2004;3(4):414-5.
163. Dwek MV, Ross HA, Streets AJ, Brooks SA, Adam E, Titcomb A, et al. Helix pomatia agglutinin lectin-binding oligosaccharides of aggressive breast cancer. *International Journal of Cancer*. 2001;95(2):79-85.
164. Brooks SA, Hall DM, Buley I. GalNAc glycoprotein expression by breast cell lines, primary breast cancer and normal breast epithelial membrane. *British Journal of Cancer*. 2001;85(7):1014-22.
165. Schumacher U, Adam E, Brooks SA, Leathem AJ. Lectin-binding properties of human breast cancer cell lines and human milk with particular reference to Helix pomatia agglutinin. *Journal of Histochemistry & Cytochemistry*. 1995;43(3):275-81.
166. Shay JW, Wright WE, Brasiskyte D, Van der Haegen BA. E6 of human papillomavirus type 16 can overcome the M1 stage of immortalization in human mammary epithelial cells but not in human fibroblasts. *Oncogene*. 1993;8(6):1407-13.
167. Lasfargues EY, Coutinho WG, Redfield ES. Isolation of two human tumor epithelial cell lines from solid breast carcinomas. *Journal of the National Cancer Institute*. 1978;61(4):967-78.
168. Soule HD, Vazquez J, Long A, Albert S, Brennan M. A human cell line from a pleural effusion derived from a breast carcinoma. *Journal of the National Cancer Institute*. 1973;51(5):1409-16.

169. Keydar I, Chen L, Karby S, Weiss FR, Delarea J, Radu M, et al. Establishment and characterization of a cell line of human breast carcinoma origin. *European Journal of Cancer* (1965). 1979;15(5):659-70.
170. Engel LW, Young NA, Tralka TS, Lippman ME, O'Brien SJ, Joyce MJ. Establishment and characterization of three new continuous cell lines derived from human breast carcinomas. *Cancer Research*. 1978;38(10):3352-64.
171. Pantazi P, Carollo E, Carter DRF, Brooks SA. A practical toolkit to study aspects of the metastatic cascade in vitro. *Acta Histochemica*. 2020;122(8):151654.
172. Bapu D, Runions J, Kadhim M, Brooks SA. N-acetylgalactosamine glycans function in cancer cell adhesion to endothelial cells: A role for truncated O-glycans in metastatic mechanisms. *Cancer Letters*. 2016;375(2):367-74.
173. Suarez-Arnedo A, Torres Figueroa F, Clavijo C, Arbelaez P, Cruz JC, Munoz-Camargo C. An image J plugin for the high throughput image analysis of in vitro scratch wound healing assays. *PLoS One*. 2020;15(7):e0232565.
174. Beaman E-M, Carter DRF, Brooks SA. Isolation of Viable Glycosylation-Specific Cell Populations for Further In Vitro or In Vivo Analysis Using Lectin-Coated Magnetic Beads. In: Pellicciari C, Biggiogera M, editors. *Histochemistry of Single Molecules: Methods and Protocols*. New York, NY: Springer New York; 2017. p. 109-19.
175. Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nature Reviews Cancer*. 2005;5(7):526-42.
176. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. *Nature Reviews Cancer*. 2015;15(9):540-55.
177. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. *Annual Review of Pathology: Mechanisms of Disease*. 2015;10:473-510.
178. Bagdonaite I, Pallesen EMH, Nielsen MI, Bennett EP, Wandall HH. Mucin-Type O-GalNAc Glycosylation in Health and Disease. In: Lauc G, Trbojević-Akmačić I, editors. *The Role of Glycosylation in Health and Disease*. Cham: Springer International Publishing; 2021. p. 25-60.
179. Hakomori S, Kannagi R. Glycosphingolipids as tumor-associated and differentiation markers. *Journal of the National Cancer Institute*. 1983;71(2):231-51.
180. Sanchez JF, Lescar J, Chazalet V, Audfray A, Gagnon J, Alvarez R, et al. Biochemical and structural analysis of Helix pomatia agglutinin. A hexameric lectin with a novel fold. *Journal of Biological Chemistry*. 2006;281(29):20171-80.
181. Lescar J, Sanchez JF, Audfray A, Coll JL, Breton C, Mitchell EP, et al. Structural basis for recognition of breast and colon cancer epitopes Tn antigen and Forssman disaccharide by Helix pomatia lectin. *Glycobiology*. 2007;17(10):1077-83.
182. Leathem AJ, Brooks SA. Predictive value of lectin binding on breast-cancer recurrence and survival. *The Lancet*. 1987;1(8541):1054-6.
183. Fenlon S, Ellis IO, Bell J, Todd JH, Elston CW, Blamey RW. Helix pomatia and Ulex europeus lectin binding in human breast carcinoma. *The Journal of Pathology*. 1987;152(3):169-76.
184. Fukutomi T, Itabashi M, Tsugane S, Yamamoto H, Nanasawa T, Hirota T. Prognostic contributions of Helix pomatia and carcinoembryonic antigen staining using histochemical techniques in breast carcinomas. *Japanese Journal of Clinical Oncology*. 1989;19(2):127-34.
185. Alam SM, Whitford P, Cushley W, George WD, Campbell AM. Flow cytometric analysis of cell surface carbohydrates in metastatic human breast cancer. *British Journal of Cancer*. 1990;62(2):238-42.
186. Brooks SA, Leathem AJ. Prediction of lymph node involvement in breast cancer by detection of altered glycosylation in the primary tumour. *The Lancet*. 1991;338(8759):71-4.
187. Noguchi M, Thomas M, Kitagawa H, Kinoshita K, Ohta N, Nagamori M, et al. Further analysis of predictive value of Helix pomatia lectin binding to primary breast cancer for axillary and internal mammary lymph node metastases. *British Journal of Cancer*. 1993;67(6):1368-71.
188. Thomas M, Noguchi M, Fonseca L, Kitagawa H, Kinoshita K, Miyazaki I. Prognostic significance of Helix pomatia lectin and c-erbB-2 oncoprotein in human breast cancer. *British Journal of Cancer*. 1993;68(3):621-6.

189. Kakeji Y, Tsujitani S, Mori M, Maehara Y, Sugimachi K. Helix pomatia agglutinin binding activity is a predictor of survival time for patients with gastric carcinoma. *Cancer*. 1991;68(11):2438-42.
190. Shiraishi T, Atsumi S-i, Yatani R. Comparative Study of Prostatic Carcinoma Bone Metastasis among Japanese in Japan and Japanese Americans and Whites in Hawaii. In: Karr JP, Yamanaka H, editors. *Prostate Cancer and Bone Metastasis*. Boston, MA: Springer US; 1992. p. 7-16.
191. Schumacher U, Adam E, Flavell DJ, Boehm D, Brooks SA, Leathem AJ. Glycosylation patterns of the human colon cancer cell line HT-29 detected by Helix pomatia agglutinin and other lectins in culture, in primary tumours and in metastases in SCID mice. *Clinical & Experimental Metastasis*. 1994;12(6):398-404.
192. Ikeda Y, Mori M, Adachi Y, Matsushima T, Sugimachi K. Prognostic value of the histochemical expression of helix pomatia agglutinin in advanced colorectal cancer. A univariate and multivariate analysis. *Diseases of the Colon & Rectum*. 1994;37(2):181-4.
193. Yoshida Y, Okamura T, Yano K, Ezaki T. Silver stained nucleolar organizer region proteins and Helix pomatia agglutinin immunostaining in esophageal carcinoma: correlated prognostic factors. *Journal of Surgical Oncology*. 1994;56(2):116-21.
194. Parameswaran R, Sadler G, Brooks S. Helix pomatia agglutinin binding glycoproteins in thyroid tumors. *World Journal of Surgery*. 2011;35(10):2219-27.
195. Laack E, Nikbakht H, Peters A, Kugler C, Jasiewicz Y, Edler L, et al. Lectin histochemistry of resected adenocarcinoma of the lung: helix pomatia agglutinin binding is an independent prognostic factor. *The American Journal of Pathology*. 2002;160(3):1001-8.
196. Brooks SA, Lymboura M, Schumacher U, Leathem AJ. Histochemistry to detect Helix pomatia lectin binding in breast cancer: methodology makes a difference. *Journal of Histochemistry & Cytochemistry*. 1996;44(5):519-24.
197. Tollefsen SE, Kornfeld R. The B4 lectin from Vicia villosa seeds interacts with N-acetylgalactosamine residues alpha-linked to serine or threonine residues in cell surface glycoproteins. *Journal of Biological Chemistry*. 1983;258(8):5172-6.
198. Tollefsen SE, Kornfeld R. Isolation and characterization of lectins from Vicia villosa. Two distinct carbohydrate binding activities are present in seed extracts. *Journal of Biological Chemistry*. 1983;258(8):5165-71.
199. Kawaguchi T, Takazawa H, Imai S, Morimoto J, Watanabe T, Kanno M, et al. Expression of Vicia villosa agglutinin (VVA)-binding glycoprotein in primary breast cancer cells in relation to lymphatic metastasis: is atypical MUC1 bearing Tn antigen a receptor of VVA? *Breast Cancer Research and Treatment*. 2006;98(1):31-43.
200. Terasawa K, Furumoto H, Kamada M, Aono T. Expression of Tn and sialyl-Tn antigens in the neoplastic transformation of uterine cervical epithelial cells. *Cancer Research*. 1996;56(9):2229-32.
201. Nishiyama T, Matsumoto Y, Watanabe H, Fujiwara M, Sato S. Detection of Tn antigen with Vicia villosa agglutinin in urinary bladder cancer: its relevance to the patient's clinical course. *Journal of the National Cancer Institute*. 1987;78(6):1113-8.
202. Hamada S, Furumoto H, Kamada M, Hirao T, Aono T. High expression rate of Tn antigen in metastatic lesions of uterine cervical cancers. *Cancer Letters*. 1993;74(3):167-73.
203. Hirao T, Sakamoto Y, Kamada M, Hamada S, Aono T. Tn antigen, a marker of potential for metastasis of uterine cervix cancer cells. *Cancer*. 1993;72(1):154-9.
204. Yamashita K, Kuno A, Matsuda A, Ikehata Y, Katada N, Hirabayashi J, et al. Lectin microarray technology identifies specific lectins related to lymph node metastasis of advanced gastric cancer. *Gastric Cancer*. 2016;19(2):531-42.
205. Konno A, Hoshino Y, Terashima S, Motoki R, Kawaguchi T. Carbohydrate expression profile of colorectal cancer cells is relevant to metastatic pattern and prognosis. *Clinical & Experimental Metastasis*. 2002;19(1):61-70.
206. Kanska G, Guerry M, Caldefie-Chezet F, De Latour M, Guillot J. Study of the expression of Tn antigen in different types of human breast cancer cells using VVA-B4 lectin. *Oncology Reports*. 2006;15(2):305-10.
207. Korourian S, Siegel E, Kieber-Emmons T, Monzavi-Karbassi B. Expression analysis of carbohydrate antigens in ductal carcinoma in situ of the breast by lectin histochemistry. *BMC Cancer*. 2008;8:136.
208. Huang T, Wu Q, Huang H, Zhang C, Wang L, Wang L, et al. Expression of GALNT8 and O-glycosylation of BMP receptor 1A suppress breast cancer cell proliferation by

- upregulating ERalpha levels. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 2022;1866(1):130046.
209. Baker DA, Sugii S, Kabat EA, Ratcliffe RM, Hermentin P, Lemieux RU. Immunochemical studies on the combining sites of Forssman hapten reactive hemagglutinins from *Dolichos biflorus*, *Helix pomatia*, and *Wistaria floribunda*. *Biochemistry*. 1983;22(11):2741-50.
 210. Goldstein IJ, Poretz RD. Isolation, physiochemical characterization and carbohydrate-binding specificity of lectins. In: Liener IE SN, and Goldstein IJ editor. *The lectins - properties, functions and applications in biology and medicine*. Academic Press. New York 1986. p. 35-250.
 211. Brooks SA, Hall DMS. Lectin Histochemistry to Detect Altered Glycosylation in Cells and Tissues. In: Dwek M, Brooks SA, Schumacher U, editors. *Metastasis Research Protocols*. Totowa, NJ: Humana Press; 2012. p. 31-50.
 212. Brooks SA, Leathem AJ. Expression of alpha-GalNAc glycoproteins by breast cancers. *British Journal of Cancer*. 1995;71(5):1033-8.
 213. Osinaga E, Bay S, Tello D, Babino A, Pritsch O, Assemet K, et al. Analysis of the fine specificity of Tn-binding proteins using synthetic glycopeptide epitopes and a biosensor based on surface plasmon resonance spectroscopy. *FEBS Letters*. 2000;469(1):24-8.
 214. Iskratsch T, Braun A, Paschinger K, Wilson IB. Specificity analysis of lectins and antibodies using remodeled glycoproteins. *Analytical Biochemistry*. 2009;386(2):133-46.
 215. Kim SW, Roh J, Park CS. Immunohistochemistry for Pathologists: Protocols, Pitfalls, and Tips. *Journal of Pathology and Translational Medicine*. 2016;50(6):411-8.
 216. Leathem A, Dokal I, Atkins N. Lectin binding to normal and malignant breast tissue. *Diagnostic Histopathology*. 1983;6(3-4):171-80.
 217. Fazekas J, Grunt TW, Jensen-Jarolim E, Singer J. Long term storage in liquid nitrogen leads to only minor phenotypic and gene expression changes in the mammary carcinoma model cell line BT474. *Oncotarget*. 2017;8(21):35076-87.
 218. Piller V, Piller F, Cartron JP. Comparison of the carbohydrate-binding specificities of seven N-acetyl-D-galactosamine-recognizing lectins. *European Journal of Biochemistry*. 1990;191(2):461-6.
 219. Elliott AD. Confocal Microscopy: Principles and Modern Practices. *Current Protocols in Cytometry*. 2020;92(1):e68.
 220. Schumacher U, Kretzschmar H, Brooks S, Leathem A. *Helix pomatia* lectin binding pattern of brain metastases originating from breast cancers. *Pathology-Research and Practice*. 1992;188(3):284-6.
 221. Brooks SA, Leathem AJ. Expression of N-acetyl galactosaminylated and sialylated glycans by metastases arising from primary breast cancer. *Invasion and Metastasis*. 1999;18(3):115-21.
 222. Bresalier RS, Ho SB, Schoeppner HL, Kim YS, Sleisenger MH, Brodt P, et al. Enhanced sialylation of mucin-associated carbohydrate structures in human colon cancer metastasis. *Gastroenterology*. 1996;110(5):1354-67.
 223. Zhang J, Ten Dijke P, Wuhrer M, Zhang T. Role of glycosylation in TGF-beta signaling and epithelial-to-mesenchymal transition in cancer. *Protein & Cell*. 2021;12(2):89-106.
 224. Tse JC, Kalluri R. Mechanisms of metastasis: epithelial-to-mesenchymal transition and contribution of tumor microenvironment. *Journal of Cellular Biochemistry*. 2007;101(4):816-29.
 225. Saitoh M. Involvement of partial EMT in cancer progression. *The Journal of Biochemistry*. 2018;164(4):257-64.
 226. Kramer RH, Nicolson GL. Interactions of tumor cells with vascular endothelial cell monolayers: a model for metastatic invasion. *Proceedings of the National Academy of Sciences*. 1979;76(11):5704-8.
 227. Lawrence MB, Springer TA. Leukocytes roll on a selectin at physiologic flow rates: distinction from and prerequisite for adhesion through integrins. *Cell*. 1991;65(5):859-73.
 228. Lawrence MB, Kansas GS, Kunkel EJ, Ley K. Threshold levels of fluid shear promote leukocyte adhesion through selectins (CD62L,P,E). *The Journal of Cell Biology*. 1997;136(3):717-27.
 229. Ashby WJ, Zijlstra A. Established and novel methods of interrogating two-dimensional cell migration. *Integrative Biology (Camb)*. 2012;4(11):1338-50.

230. Kramer N, Walzl A, Unger C, Rosner M, Krupitza G, Hengstschlager M, et al. In vitro cell migration and invasion assays. *Mutation Research/Reviews in Mutation Research*. 2013;752(1):10-24.
231. Boyden S. The chemotactic effect of mixtures of antibody and antigen on polymorphonuclear leucocytes. *The Journal of Experimental Medicine*. 1962;115(3):453-66.
232. Marshall J. Transwell® Invasion Assays. In: Wells CM, Parsons M, editors. *Cell Migration: Developmental Methods and Protocols*. Totowa, NJ: Humana Press; 2011. p. 97-110.
233. Lowe DJ, Raj K. Quantitation of Endothelial Cell Adhesiveness In Vitro. *JoVE (Journal of Visualized Experiments)*. 2015(100):e52924.
234. Babu D, Khadim M, Brooks SA. Rocking Adhesion Assay System to Study Adhesion and Transendothelial Migration of Cancer Cells. In: Dwek M, Schumacher U, Brooks SA, editors. *Metastasis Research Protocols*. New York, NY: Springer New York; 2014. p. 37-45.
235. Orr FW, Wang HH, Lafrenie RM, Scherbarth S, Nance DM. Interactions between cancer cells and the endothelium in metastasis. *The Journal of Pathology*. 2000;190(3):310-29.
236. Lane WO, Jantzen AE, Carlon TA, Jamiolkowski RM, Grenet JE, Ley MM, et al. Parallel-plate flow chamber and continuous flow circuit to evaluate endothelial progenitor cells under laminar flow shear stress. *JoVE (Journal of Visualized Experiments)*. 2012(59).
237. Pucci M, Malagolini N, Dall'Olio F. Glycobiology of the Epithelial to Mesenchymal Transition. *Biomedicines*. 2021;9(7).
238. Harder JL, Whiteman EL, Pieczynski JN, Liu CJ, Margolis B. Snail destabilizes cell surface Crumbs3a. *Traffic*. 2012;13(8):1170-85.
239. Carvalho S, Reis CA, Pinho SS. Cadherins Glycans in Cancer: Sweet Players in a Bitter Process. *Trends in Cancer*. 2016;2(9):519-31.
240. Gupta R, Ponangi R, Indresh KG. Role of glycosylation in breast cancer progression and metastasis: implications for miRNA, EMT and multidrug resistance. *Glycobiology*. 2023;33(7):545-55.
241. Park JH, Katagiri T, Chung S, Kijima K, Nakamura Y. Polypeptide N-acetylgalactosaminyltransferase 6 disrupts mammary acinar morphogenesis through O-glycosylation of fibronectin. *Neoplasia*. 2011;13(4):320-6.
242. Wu Q, Zhang C, Zhang K, Chen Q, Wu S, Huang H, et al. ppGalNAc-T4-catalyzed O-Glycosylation of TGF-beta type II receptor regulates breast cancer cells metastasis potential. *Journal of Biological Chemistry*. 2021;296:100119.
243. Wang S, Mao Y, Narimatsu Y, Ye Z, Tian W, Goth CK, et al. Site-specific O-glycosylation of members of the low-density lipoprotein receptor superfamily enhances ligand interactions. *Journal of Biological Chemistry*. 2018;293(19):7408-22.
244. Ludyga N, Anastasov N, Gonzalez-Vasconcellos I, Ram M, Hofler H, Aubele M. Impact of protein tyrosine kinase 6 (PTK6) on human epidermal growth factor receptor (HER) signalling in breast cancer. *Molecular BioSystems*. 2011;7(5):1603-12.
245. Luo Y, Radice GL. N-cadherin acts upstream of VE-cadherin in controlling vascular morphogenesis. *The Journal of Cell Biology*. 2005;169(1):29-34.
246. Liu X, Huang H, Remmers N, Hollingsworth MA. Loss of E-cadherin and epithelial to mesenchymal transition is not required for cell motility in tissues or for metastasis. *Tissue Barriers*. 2014;2(4):e969112.
247. Bruner HC, Derksen PWB. Loss of E-Cadherin-Dependent Cell-Cell Adhesion and the Development and Progression of Cancer. *Cold Spring Harbor Perspectives in Biology*. 2018;10(3).
248. Sommers CL, Byers SW, Thompson EW, Torri JA, Gelmann EP. Differentiation state and invasiveness of human breast cancer cell lines. *Breast Cancer Research and Treatment*. 1994;31(2-3):325-35.
249. Kapoor P, Saunders MM, Li Z, Zhou Z, Sheaffer N, Kunze EL, et al. Breast cancer metastatic potential: correlation with increased heterotypic gap junctional intercellular communication between breast cancer cells and osteoblastic cells. *International Journal of Cancer*. 2004;111(5):693-7.
250. Beaudin SG, Robilotto S, Welsh J. Comparative regulation of gene expression by 1,25-dihydroxyvitamin D3 in cells derived from normal mammary tissue and breast cancer. *The Journal of Steroid Biochemistry and Molecular Biology*. 2015;148:96-102.

251. Dine JL, O'Sullivan CC, Voeller D, Greer YE, Chavez KJ, Conway CM, et al. The TRAIL receptor agonist drozitumab targets basal B triple-negative breast cancer cells that express vimentin and Axl. *Breast Cancer Research and Treatment*. 2016;155(2):235-51.
252. Prasad RR, Mishra DK, Kumar M, Yadava PK. Human telomerase reverse transcriptase promotes the epithelial to mesenchymal transition in lung cancer cells by enhancing c-MET upregulation. *Heliyon*. 2022;8(1):e08673.
253. Polioudaki H, Agelaki S, Chiotaki R, Politaki E, Mavroudis D, Matikas A, et al. Variable expression levels of keratin and vimentin reveal differential EMT status of circulating tumor cells and correlation with clinical characteristics and outcome of patients with metastatic breast cancer. *BMC Cancer*. 2015;15:399.
254. Raghu D, Mobley RJ, Shendy NAM, Perry CH, Abell AN. GALNT3 Maintains the Epithelial State in Trophoblast Stem Cells. *Cell Reports*. 2019;26(13):3684-97 e7.
255. Wang Y, Liu J, Ying X, Lin PC, Zhou BP. Twist-mediated Epithelial-mesenchymal Transition Promotes Breast Tumor Cell Invasion via Inhibition of Hippo Pathway. *Scientific Reports*. 2016;6:24606.
256. Groen RW, de Rooij MF, Kocemba KA, Reijmers RM, de Haan-Kramer A, Overdijk MB, et al. N-cadherin-mediated interaction with multiple myeloma cells inhibits osteoblast differentiation. *Haematologica*. 2011;96(11):1653-61.
257. Liwosz A, Lei T, Kukuruzinska MA. N-glycosylation affects the molecular organization and stability of E-cadherin junctions. *Journal of Biological Chemistry*. 2006;281(32):23138-49.
258. Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Lavery SB, et al. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. *Proceedings of the National Academy of Sciences*. 2013;110(52):21018-23.
259. Kam Y, Guess C, Estrada L, Weidow B, Quaranta V. A novel circular invasion assay mimics in vivo invasive behavior of cancer cell lines and distinguishes single-cell motility in vitro. *BMC Cancer*. 2008;8:198.
260. Chen H, Paradies NE, Fedor-Chaiken M, Brackenbury R. E-cadherin mediates adhesion and suppresses cell motility via distinct mechanisms. *Journal of Cell Science*. 1997;110 (Pt 3):345-56.
261. Liang G, Ding M, Lu H, Cao NA, Niu Y, Gao Y, et al. Metformin upregulates E-cadherin and inhibits B16F10 cell motility, invasion and migration. *Oncology Letters*. 2015;10(3):1527-32.
262. McInroy L, Maatta A. Down-regulation of vimentin expression inhibits carcinoma cell migration and adhesion. *Biochemical and Biophysical Research Communications*. 2007;360(1):109-14.
263. Mendez MG, Kojima S, Goldman RD. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. *FASEB Journal*. 2010;24(6):1838-51.
264. Sun J, Fu X, Wang Y, Liu Y, Zhang Y, Hao T, et al. Erianin inhibits the proliferation of T47D cells by inhibiting cell cycles, inducing apoptosis and suppressing migration. *American Journal of Translational Research*. 2016;8(7):3077-86.
265. Ferreira JA, Videira PA, Lima L, Pereira S, Silva M, Carrascal M, et al. Overexpression of tumour-associated carbohydrate antigen sialyl-Tn in advanced bladder tumours. *Molecular Oncology*. 2013;7(3):719-31.
266. Kolasinska E, Przybylo M, Janik M, Litynska A. Towards understanding the role of sialylation in melanoma progression. *Acta Biochimica Polonica*. 2016;63(3):533-41.
267. Wu X, Zhao J, Ruan Y, Sun L, Xu C, Jiang H. Sialyltransferase ST3GAL1 promotes cell migration, invasion, and TGF-beta1-induced EMT and confers paclitaxel resistance in ovarian cancer. *Cell Death & Disease*. 2018;9(11):1102.
268. Hofmann BT, Schluter L, Lange P, Mercanoglu B, Ewald F, Folster A, et al. COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer. *Molecular Cancer*. 2015;14:109.
269. Markowski AR, Ustymowicz K, Markowska AJ, Romanczyk W, Guzinska-Ustymowicz K. E-Cadherin Expression Varies Depending on the Location within the Primary Tumor and Is Higher in Colorectal Cancer with Lymphoid Follicles. *Cancers (Basel)*. 2023;15(12).
270. Christofori G, Semb H. The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene. *Trends in Biochemical Sciences*. 1999;24(2):73-6.

271. Wager LJ, Guanzon D, Leavesley DI, Thompson EW, Murray RZ. Partial Epithelial-Mesenchymal Transition: Reduced miR-4792 and miR-146b-5p Inversely Correlated with SIAH2 in Migrating Keratinocytes in Vitro. *Experimental Dermatology*. 2021;30(12):1838-9.
272. Ritch SJ, Noman ASM, Goyeneche AA, Telleria CM. The metastatic capacity of high-grade serous ovarian cancer cells changes along disease progression: inhibition by mifepristone. *Cancer Cell International*. 2022;22(1):397.
273. Gilles C, Polette M, Zahm JM, Tournier JM, Volders L, Foidart JM, et al. Vimentin contributes to human mammary epithelial cell migration. *Journal of Cell Science*. 1999;112 (Pt 24):4615-25.
274. Yan C, Grimm WA, Garner WL, Qin L, Travis T, Tan N, et al. Epithelial to mesenchymal transition in human skin wound healing is induced by tumor necrosis factor- α through bone morphogenic protein-2. *The American Journal of Pathology*. 2010;176(5):2247-58.
275. Strazza M, Azoulay-Alfaguter I, Pedoeem A, Mor A. Static adhesion assay for the study of integrin activation in T lymphocytes. *JoVE (Journal of Visualized Experiments)*. 2014(88).
276. Stevens JM. Platelet Adhesion Assays Performed Under Static Conditions. In: Gibbins JM, Mahaut-Smith MP, editors. *Platelets and Megakaryocytes: Volume 1: Functional Assays*. Totowa, NJ: Humana Press; 2004. p. 145-51.
277. Glinsky VV, Glinsky GV, Rittenhouse-Olson K, Huflejt ME, Glinskii OV, Deutscher SL, et al. The role of Thomsen-Friedenreich antigen in adhesion of human breast and prostate cancer cells to the endothelium. *Cancer Research*. 2001;61(12):4851-7.
278. Danussi C, Coslovi A, Campa C, Mucignat MT, Spessotto P, Uggeri F, et al. A newly generated functional antibody identifies Tn antigen as a novel determinant in the cancer cell-lymphatic endothelium interaction. *Glycobiology*. 2009;19(10):1056-67.
279. Kishikawa T, Ghazizadeh M, Sasaki Y, Springer GF. Specific role of T and Tn tumor-associated antigens in adhesion between a human breast carcinoma cell line and a normal human breast epithelial cell line. *Japanese Journal of Cancer Research*. 1999;90(3):326-32.
280. Huber MA, Kraut N, Beug H. Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Current Opinion in Cell Biology*. 2005;17(5):548-58.
281. Sullivan NJ, Sasser AK, Axel AE, Vesuna F, Raman V, Ramirez N, et al. Interleukin-6 induces an epithelial-mesenchymal transition phenotype in human breast cancer cells. *Oncogene*. 2009;28(33):2940-7.
282. Fernando RI, Castillo MD, Litzinger M, Hamilton DH, Palena C. IL-8 signaling plays a critical role in the epithelial-mesenchymal transition of human carcinoma cells. *Cancer Research*. 2011;71(15):5296-306.
283. Wu Y, Deng J, Rychahou PG, Qiu S, Evers BM, Zhou BP. Stabilization of snail by NF- κ B is required for inflammation-induced cell migration and invasion. *Cancer Cell*. 2009;15(5):416-28.
284. Xu J, Lamouille S, Derynck R. TGF- β -induced epithelial to mesenchymal transition. *Cell Research*. 2009;19(2):156-72.
285. Hao Y, Baker D, Ten Dijke P. TGF- β -Mediated Epithelial-Mesenchymal Transition and Cancer Metastasis. *International Journal of Molecular Sciences*. 2019;20(11).
286. Miettinen PJ, Ebner R, Lopez AR, Derynck R. TGF- β induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *The Journal of Cell Biology*. 1994;127(6 Pt 2):2021-36.
287. Vincent T, Neve EP, Johnson JR, Kukalev A, Rojo F, Albanell J, et al. A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF- β mediated epithelial-mesenchymal transition. *Nature Cell Biology*. 2009;11(8):943-50.
288. Buyuk B, Jin S, Ye K. Epithelial-to-Mesenchymal Transition Signaling Pathways Responsible for Breast Cancer Metastasis. *Cellular and Molecular Bioengineering*. 2022;15(1):1-13.
289. Ray I, Michael A, Meira LB, Ellis PE. The Role of Cytokines in Epithelial-Mesenchymal Transition in Gynaecological Cancers: A Systematic Review. *Cells*. 2023;12(3).
290. Heldin CH, Landstrom M, Moustakas A. Mechanism of TGF- β signaling to growth arrest, apoptosis, and epithelial-mesenchymal transition. *Current Opinion in Cell Biology*. 2009;21(2):166-76.

291. Wendt MK, Tian M, Schiemann WP. Deconstructing the mechanisms and consequences of TGF-beta-induced EMT during cancer progression. *Cell and Tissue Research*. 2012;347(1):85-101.
292. Morrison CD, Parvani JG, Schiemann WP. The relevance of the TGF-beta Paradox to EMT-MET programs. *Cancer Letters*. 2013;341(1):30-40.
293. Bierie B, Moses HL. TGF-beta and cancer. *Cytokine & Growth Factor Reviews*. 2006;17(1-2):29-40.
294. Kubiczko L, Sedlarikova L, Hajek R, Sevcikova S. TGF-beta - an excellent servant but a bad master. *Journal of Translational Medicine*. 2012;10:183.
295. Yu Q, Stamenkovic I. Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. *Genes & Development*. 2000;14(2):163-76.
296. Schultz-Cherry S, Murphy-Ullrich JE. Thrombospondin causes activation of latent transforming growth factor-beta secreted by endothelial cells by a novel mechanism. *The Journal of Cell Biology*. 1993;122(4):923-32.
297. Munger JS, Huang X, Kawakatsu H, Griffiths MJ, Dalton SL, Wu J, et al. The integrin alpha v beta 6 binds and activates latent TGF beta 1: a mechanism for regulating pulmonary inflammation and fibrosis. *Cell*. 1999;96(3):319-28.
298. Mu D, Cambier S, Fjellbirkeland L, Baron JL, Munger JS, Kawakatsu H, et al. The integrin alpha(v)beta8 mediates epithelial homeostasis through MT1-MMP-dependent activation of TGF-beta1. *The Journal of Cell Biology*. 2002;157(3):493-507.
299. Barcellos-Hoff MH, Dix TA. Redox-mediated activation of latent transforming growth factor-beta 1. *Molecular Endocrinology*. 1996;10(9):1077-83.
300. Lyons RM, Keski-Oja J, Moses HL. Proteolytic activation of latent transforming growth factor-beta from fibroblast-conditioned medium. *The Journal of Cell Biology*. 1988;106(5):1659-65.
301. Gatza CE, Oh SY, Blobe GC. Roles for the type III TGF-beta receptor in human cancer. *Cellular Signalling*. 2010;22(8):1163-74.
302. Zhang YE. Non-Smad Signaling Pathways of the TGF-beta Family. *Cold Spring Harbor Perspectives in Biology*. 2017;9(2).
303. Liu S, Ren J, Ten Dijke P. Targeting TGFbeta signal transduction for cancer therapy. *Signal Transduction and Targeted Therapy*. 2021;6(1):8.
304. Zhang M, Zhang YY, Chen Y, Wang J, Wang Q, Lu H. TGF-beta Signaling and Resistance to Cancer Therapy. *Frontiers in Cell and Developmental Biology*. 2021;9:786728.
305. Derynck R, Weinberg RA. EMT and Cancer: More Than Meets the Eye. *Developmental Cell*. 2019;49(3):313-6.
306. Zhang X, Li Y, Zhang Y, Song J, Wang Q, Zheng L, et al. Beta-elemene blocks epithelial-mesenchymal transition in human breast cancer cell line MCF-7 through Smad3-mediated down-regulation of nuclear transcription factors. *PLoS One*. 2013;8(3):e58719.
307. Mahdi SH, Cheng H, Li J, Feng R. The effect of TGF-beta-induced epithelial-mesenchymal transition on the expression of intracellular calcium-handling proteins in T47D and MCF-7 human breast cancer cells. *Archives of Biochemistry and Biophysics*. 2015;583:18-26.
308. Zhang C, Lu Y, Li Q, Mao J, Hou Z, Yu X, et al. Salinomycin suppresses TGF-beta1-induced epithelial-to-mesenchymal transition in MCF-7 human breast cancer cells. *Chemico-Biological Interactions*. 2016;248:74-81.
309. Katsuno Y, Meyer DS, Zhang Z, Shokat KM, Akhurst RJ, Miyazono K, et al. Chronic TGF-beta exposure drives stabilized EMT, tumor stemness, and cancer drug resistance with vulnerability to bitopic mTOR inhibition. *Science Signaling*. 2019;12(570).
310. Kurcon T, Liu Z, Paradkar AV, Vaiana CA, Koppolu S, Agrawal P, et al. miRNA proxy approach reveals hidden functions of glycosylation. *Proceedings of the National Academy of Sciences*. 2015;112(23):7327-32.
311. Du J, Hong S, Dong L, Cheng B, Lin L, Zhao B, et al. Dynamic Sialylation in Transforming Growth Factor-beta (TGF-beta)-induced Epithelial to Mesenchymal Transition. *Journal of Biological Chemistry*. 2015;290(19):12000-13.
312. Tian M, Schiemann WP. TGF-beta Stimulation of EMT Programs Elicits Non-genomic ER-alpha Activity and Anti-estrogen Resistance in Breast Cancer Cells. *Journal of Cancer Metastasis and Treatment*. 2017;3:150-60.

313. Wang L, Xu C, Liu X, Yang Y, Cao L, Xiang G, et al. TGF-beta1 stimulates epithelial-mesenchymal transition and cancer-associated myoepithelial cell during the progression from in situ to invasive breast cancer. *Cancer Cell International*. 2019;19:343.
314. Pecina-Slaus N. Tumor suppressor gene E-cadherin and its role in normal and malignant cells. *Cancer Cell International*. 2003;3(1):17.
315. Janda E, Nevolo M, Lehmann K, Downward J, Beug H, Grieco M. Raf plus TGFbeta-dependent EMT is initiated by endocytosis and lysosomal degradation of E-cadherin. *Oncogene*. 2006;25(54):7117-30.
316. Aban CE, Lombardi A, Neiman G, Biani MC, La Greca A, Waisman A, et al. Downregulation of E-cadherin in pluripotent stem cells triggers partial EMT. *Scientific Reports*. 2021;11(1):2048.
317. Haerinck J, Berx G. Partial EMT takes the lead in cancer metastasis. *Developmental Cell*. 2021;56(23):3174-6.
318. Aiello NM, Kang Y. Context-dependent EMT programs in cancer metastasis. *The Journal of Experimental Medicine*. 2019;216(5):1016-26.
319. Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell*. 2017;171(7):1611-24 e24.
320. Cheung KJ, Gabrielson E, Werb Z, Ewald AJ. Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell*. 2013;155(7):1639-51.
321. Hazan RB, Phillips GR, Qiao RF, Norton L, Aaronson SA. Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis. *The Journal of Cell Biology*. 2000;148(4):779-90.
322. Rees-Milton KJ, Terry D, Anastassiades TP. Hyperglycosylation of fibronectin by TGF-beta1-stimulated chondrocytes. *Biochemical and Biophysical Research Communications*. 2004;317(3):844-50.
323. Yang X, Yip J, Anastassiades T, Harrison M, Brockhausen I. The action of TNFalpha and TGFbeta include specific alterations of the glycosylation of bovine and human chondrocytes. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*. 2007;1773(2):264-72.
324. Lucena MC, Carvalho-Cruz P, Donadio JL, Oliveira IA, de Queiroz RM, Marinho-Carvalho MM, et al. Epithelial Mesenchymal Transition Induces Aberrant Glycosylation through Hexosamine Biosynthetic Pathway Activation. *Journal of Biological Chemistry*. 2016;291(25):12917-29.
325. Delmotte P, Degroote S, Lafitte JJ, Lamblin G, Perini JM, Roussel P. Tumor necrosis factor alpha increases the expression of glycosyltransferases and sulfotransferases responsible for the biosynthesis of sialylated and/or sulfated Lewis x epitopes in the human bronchial mucosa. *Journal of Biological Chemistry*. 2002;277(1):424-31.
326. Hanasaki K, Varki A, Stamenkovic I, Bevilacqua MP. Cytokine-induced beta-galactoside alpha-2,6-sialyltransferase in human endothelial cells mediates alpha 2,6-sialylation of adhesion molecules and CD22 ligands. *Journal of Biological Chemistry*. 1994;269(14):10637-43.
327. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nature Reviews Genetics*. 2019;20(11):631-56.
328. Kukurba KR, Montgomery SB. RNA Sequencing and Analysis. *Cold Spring Harbor Perspectives in Biology*. 2015;2015(11):951-69.
329. Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. *Nature Reviews Molecular Cell Biology*. 2020;21(12):729-49.
330. Nairn AV, York WS, Harris K, Hall EM, Pierce JM, Moremen KW. Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *Journal of Biological Chemistry*. 2008;283(25):17298-313.
331. Nairn AV, Aoki K, dela Rosa M, Porterfield M, Lim JM, Kulik M, et al. Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis. *Journal of Biological Chemistry*. 2012;287(45):37835-56.
332. Maier T, Guell M, Serrano L. Correlation of mRNA and protein in complex biological samples. *FEBS Letters*. 2009;583(24):3966-73.

333. Huang YF, Aoki K, Akase S, Ishihara M, Liu YS, Yang G, et al. Global mapping of glycosylation pathways in human-derived cells. *Developmental Cell*. 2021;56(8):1195-209 e7.
334. Yang X, Hu Q, Hu LX, Lin XR, Liu JQ, Lin X, et al. miR-200b regulates epithelial-mesenchymal transition of chemo-resistant breast cancer cells by targeting FN1. *Discovery Medicine*. 2017;24(131):75-85.
335. Nagaishi M, Kim YH, Mittelbronn M, Giangaspero F, Paulus W, Brokinkel B, et al. Amplification of the STOML3, FREM2, and LHFP genes is associated with mesenchymal differentiation in gliosarcoma. *The American Journal of Pathology*. 2012;180(5):1816-23.
336. Li H, Batth IS, Qu X, Xu L, Song N, Wang R, et al. IGF-IR signaling in epithelial to mesenchymal transition and targeting IGF-IR therapy: overview and new insights. *Molecular Cancer*. 2017;16(1):6.
337. Liu X, Xu D, Liu Z, Li Y, Zhang C, Gong Y, et al. THBS1 facilitates colorectal liver metastasis through enhancing epithelial-mesenchymal transition. *Clinical and Translational Oncology*. 2020;22(10):1730-40.
338. Chen Y, Zhou H, Jiang WJ, Wang JF, Tian Y, Jiang Y, et al. The role of CEMIP in tumors: An update based on cellular and molecular insights. *Biomedicine & Pharmacotherapy*. 2022;146:112504.
339. Oshi M, Tokumaru Y, Mukhopadhyay S, Yan L, Matsuyama R, Endo I, et al. Annexin A1 Expression Is Associated with Epithelial-Mesenchymal Transition (EMT), Cell Proliferation, Prognosis, and Drug Response in Pancreatic Cancer. *Cells*. 2021;10(3).
340. Jin Y, Li JL. Olfactomedin-like 3: possible functions in embryonic development and tumorigenesis. *Chinese Medical Journal*. 2019;132(14):1733-8.
341. Xu J, Wu G, Zhao Y, Han Y, Zhang S, Li C, et al. Long Noncoding RNA DSCAM-AS1 Facilitates Colorectal Cancer Cell Proliferation and Migration via miR-137/Notch1 Axis. *Journal of Cancer*. 2020;11(22):6623-32.
342. Dou C, Xu Q, Liu J, Wang Y, Zhou Z, Yao W, et al. SHMT1 inhibits the metastasis of HCC by repressing NOX1-mediated ROS production. *Journal of Experimental & Clinical Cancer Research*. 2019;38(1):70.
343. Appunni S, Rubens M, Ramamoorthy V, Anand V, Khandelwal M, Saxena A, et al. Lumican, pro-tumorigenic or anti-tumorigenic: A conundrum. *Clinica Chimica Acta*. 2021;514:1-7.
344. Hong Q, Li R, Zhang Y, Gu K. Fibrillin 2 gene knockdown inhibits invasion and migration of lung cancer cells. *Cellular and Molecular Biology*. 2020;66(7):190-6.
345. Ketterer S, Mitschke J, Ketscher A, Schlimpert M, Reichardt W, Baeuerle N, et al. Cathepsin D deficiency in mammary epithelium transiently stalls breast cancer by interference with mTORC1 signaling. *Nature Communications*. 2020;11(1):5133.
346. Yamada T, Ohta K, Motooka Y, Fujino K, Kudoh S, Tenjin Y, et al. Significance of Tsukushi in lung cancer. *Lung Cancer*. 2019;131:104-11.
347. Chen JH, Kuo KT, Bamodu OA, Lin YC, Yang RB, Yeh CT, et al. Upregulated SCUBE2 expression in breast cancer stem cells enhances triple negative breast cancer aggression through modulation of notch signaling and epithelial-to-mesenchymal transition. *Experimental Cell Research*. 2018;370(2):444-53.
348. Chang J, Hu X, Nan J, Zhang X, Jin X. HOXD9-induced SCNN1A upregulation promotes pancreatic cancer cell proliferation, migration and predicts prognosis by regulating epithelial-mesenchymal transformation. *Molecular Medicine Reports*. 2021;24(5).
349. Shahi P, Wang CY, Chou J, Hagerling C, Gonzalez Velozo H, Ruderisch A, et al. GATA3 targets semaphorin 3B in mammary epithelial cells to suppress breast cancer progression and metastasis. *Oncogene*. 2017;36(40):5567-75.
350. Bailey KM, Liu J. Caveolin-1 up-regulation during epithelial to mesenchymal transition is mediated by focal adhesion kinase. *Journal of Biological Chemistry*. 2008;283(20):13714-24.
351. Guan Y, Wang Y, Bhandari A, Xia E, Wang O. IGSF1: A novel oncogene regulates the thyroid cancer progression. *Cell Biochemistry and Function*. 2019;37(7):516-24.
352. Ran F, Zhang Y, Shi Y, Liu J, Li H, Ding L, et al. miR-1224-3p Promotes Breast Cancer Cell Proliferation and Migration through PGM5-Mediated Aerobic Glycolysis. *Journal of Oncology*. 2021;2021:5529770.
353. Yusufu A, Shayimu P, Tuerdi R, Fang C, Wang F, Wang H. TFF3 and TFF1 expression levels are elevated in colorectal cancer and promote the malignant behavior of

colon cancer by activating the EMT process. *International Journal of Oncology*. 2019;55(4):789-804.

354. Zheng H, Kang Y. Multilayer control of the EMT master regulators. *Oncogene*. 2014;33(14):1755-63.

355. Karagiorgou Z, Fountas PN, Manou D, Knutsen E, Theocharis AD. Proteoglycans Determine the Dynamic Landscape of EMT and Cancer Cell Stemness. *Cancers (Basel)*. 2022;14(21).

356. Debnath P, Huiem RS, Dutta P, Palchaudhuri S. Epithelial-mesenchymal transition and its transcription factors. *Bioscience Reports*. 2022;42(1).

357. Deshmukh AP, Vasaikar SV, Tomczak K, Tripathi S, den Hollander P, Arslan E, et al. Identification of EMT signaling cross-talk and gene regulatory networks by single-cell RNA sequencing. *Proceedings of the National Academy of Sciences*. 2021;118(19).

358. Wang Y, Shi J, Chai K, Ying X, Zhou BP. The Role of Snail in EMT and Tumorigenesis. *Current Cancer Drug Targets*. 2013;13(9):963-72.

359. Li CL, Yang D, Cao X, Wang F, Hong DY, Wang J, et al. Fibronectin induces epithelial-mesenchymal transition in human breast cancer MCF-7 cells via activation of calpain. *Oncology Letters*. 2017;13(5):3889-95.

360. Li B, Shen W, Peng H, Li Y, Chen F, Zheng L, et al. Fibronectin 1 promotes melanoma proliferation and metastasis by inhibiting apoptosis and regulating EMT. *OncoTargets and Therapy*. 2019;12:3207-21.

361. Park J, Schwarzbauer JE. Mammary epithelial cell interactions with fibronectin stimulate epithelial-mesenchymal transition. *Oncogene*. 2014;33(13):1649-57.

362. Dallas SL, Sivakumar P, Jones CJ, Chen Q, Peters DM, Mosher DF, et al. Fibronectin regulates latent transforming growth factor-beta (TGF beta) by controlling matrix assembly of latent TGF beta-binding protein-1. *Journal of Biological Chemistry*. 2005;280(19):18871-80.

363. Griggs LA, Hassan NT, Malik RS, Griffin BP, Martinez BA, Elmore LW, et al. Fibronectin fibrils regulate TGF-beta1-induced Epithelial-Mesenchymal Transition. *Matrix Biology*. 2017;60-61:157-75.

364. Matsumura K, Hayashi H, Uemura N, Ogata Y, Zhao L, Sato H, et al. Thrombospondin-1 overexpression stimulates loss of Smad4 and accelerates malignant behavior via TGF-beta signal activation in pancreatic ductal adenocarcinoma. *Translational Oncology*. 2022;26:101533.

365. Tang Z, Ding Y, Shen Q, Zhang C, Li J, Nazar M, et al. KIAA1199 promotes invasion and migration in non-small-cell lung cancer (NSCLC) via PI3K-Akt mediated EMT. *Journal of Molecular Medicine*. 2019;97(1):127-40.

366. Liang G, Fang X, Yang Y, Song Y. Silencing of CEMIP suppresses Wnt/beta-catenin/Snail signaling transduction and inhibits EMT program of colorectal cancer cells. *Acta Histochemica*. 2018;120(1):56-63.

367. Zhang D, Zhao L, Shen Q, Lv Q, Jin M, Ma H, et al. Down-regulation of KIAA1199/CEMIP by miR-216a suppresses tumor invasion and metastasis in colorectal cancer. *International Journal of Cancer*. 2017;140(10):2298-309.

368. Kleiser S, Nystrom A. Interplay between Cell-Surface Receptors and Extracellular Matrix in Skin. *Biomolecules*. 2020;10(8).

369. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 2008;133(4):704-15.

370. de Graauw M, van Miltenburg MH, Schmidt MK, Pont C, Lalai R, Kartopawiro J, et al. Annexin A1 regulates TGF-beta signaling and promotes metastasis formation of basal-like breast cancer cells. *Proceedings of the National Academy of Sciences*. 2010;107(14):6340-5.

371. Cadigan KM, Waterman ML. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harbor Perspectives in Biology*. 2012;4(11).

372. Beardsley A, Fang K, Mertz H, Castranova V, Friend S, Liu J. Loss of caveolin-1 polarity impedes endothelial cell polarization and directional movement. *Journal of Biological Chemistry*. 2005;280(5):3541-7.

373. Urra H, Torres VA, Ortiz RJ, Lobos L, Diaz MI, Diaz N, et al. Caveolin-1-enhanced motility and focal adhesion turnover require tyrosine-14 but not accumulation to the rear in metastatic cancer cells. *PLoS One*. 2012;7(4):e33085.

374. Godina C, Khazaei S, Belting M, Vallon-Christersson J, Nodin B, Jirstrom K, et al. High Caveolin-1 mRNA expression in triple-negative breast cancer is associated with an aggressive tumor microenvironment, chemoresistance, and poor clinical outcome. *PLoS One*. 2024;19(7):e0305222.
375. Burgos-Panadero R, Noguera I, Canete A, Navarro S, Noguera R. Vitronectin as a molecular player of the tumor microenvironment in neuroblastoma. *BMC Cancer*. 2019;19(1):479.
376. Sen A, Ta M. Vitronectin acts as a key regulator of adhesion and migration in human umbilical cord-derived MSCs under different stress conditions. *Experimental Cell Research*. 2023;423(2):113467.
377. Gurrapu S, Tamagnone L. Semaphorins as Regulators of Phenotypic Plasticity and Functional Reprogramming of Cancer Cells. *Trends in Molecular Medicine*. 2019;25(4):303-14.
378. Huang HC, Klein PS. The Frizzled family: receptors for multiple signal transduction pathways. *Genome Biology*. 2004;5(7):234.
379. Chu X, Tian W, Ning J, Xiao G, Zhou Y, Wang Z, et al. Cancer stem cells: advances in knowledge and implications for cancer therapy. *Signal Transduction and Targeted Therapy*. 2024;9(1):170.
380. Mockl L, Pedram K, Roy AR, Krishnan V, Gustavsson AK, Dorigo O, et al. Quantitative Super-Resolution Microscopy of the Mammalian Glycocalyx. *Developmental Cell*. 2019;50(1):57-72 e6.
381. Peng RQ, Wan HY, Li HF, Liu M, Li X, Tang H. MicroRNA-214 suppresses growth and invasiveness of cervical cancer cells by targeting UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7. *Journal of Biological Chemistry*. 2012;287(17):14301-9.
382. Chang LY, Liang SY, Lu SC, Tseng HC, Tsai HY, Tang CJ, et al. Molecular Basis and Role of Siglec-7 Ligand Expression on Chronic Lymphocytic Leukemia B Cells. *Frontiers in Immunology*. 2022;13:840388.
383. Fuseya S, Izumi H, Hamano A, Murakami Y, Suzuki R, Koizumi R, et al. Reduction in disialyl-T antigen levels in mice deficient for both St6galnac3 and St6galnac4 results in blood filling of lymph nodes. *Scientific Reports*. 2023;13(1):10582.
384. Man D, Jiang Y, Zhang D, Wu J, Ding B, Liu H, et al. ST6GALNAC4 promotes hepatocellular carcinogenesis by inducing abnormal glycosylation. *Journal of Translational Medicine*. 2023;21(1):420.
385. Kurtenkov O. Profiling of Naturally Occurring Antibodies to the Thomsen-Friedenreich Antigen in Health and Cancer: The Diversity and Clinical Potential. *BioMed Research International*. 2020;2020:9747040.
386. Roy R, Baek MG. Glycodendrimers: novel glycotope isosteres unmasking sugar coding. case study with T-antigen markers from breast cancer MUC1 glycoprotein. *Reviews in Molecular Biotechnology*. 2002;90(3-4):291-309.
387. Gonzalez-Ramirez AM, Grosso AS, Yang Z, Companon I, Coelho H, Narimatsu Y, et al. Structural basis for the synthesis of the core 1 structure by C1GALT1. *Nature Communications*. 2022;13(1):2398.
388. Xia T, Xiang T, Xie H. Update on the role of C1GALT1 in cancer. *Oncology Letters*. 2022;23(3):97.
389. Tian H, Yu JL, Chu X, Guan Q, Liu J, Liu Y. Unraveling the role of C1GALT1 in abnormal glycosylation and colorectal cancer progression. *Frontiers in Oncology*. 2024;14:1389713.
390. Chou CH, Huang MJ, Chen CH, Shyu MK, Huang J, Hung JS, et al. Up-regulation of C1GALT1 promotes breast cancer cell growth through MUC1-C signaling pathway. *Oncotarget*. 2015;6(8):6123-35.
391. Dobie C, Skropeta D. Insights into the role of sialylation in cancer progression and metastasis. *British Journal of Cancer*. 2021;124(1):76-90.
392. Chong YK, Sandanaraj E, Koh LW, Thangaveloo M, Tan MS, Koh GR, et al. ST3GAL1-Associated Transcriptomic Program in Glioblastoma Tumor Growth, Invasion, and Prognosis. *Journal of the National Cancer Institute*. 2016;108(2).
393. Milde-Langosch K, Karn T, Schmidt M, zu Eulenburg C, Oliveira-Ferrer L, Wirtz RM, et al. Prognostic relevance of glycosylation-associated genes in breast cancer. *Breast Cancer Research and Treatment*. 2014;145(2):295-305.

394. Sun H, Chang J, Ye M, Weng W, Zhang M, Ni S, et al. GCNT4 is Associated with Prognosis and Suppress Cell Proliferation in Gastric Cancer. *OncoTargets and Therapy*. 2020;13:8601-13.
395. Hu W, Zheng X, Liu J, Zhang M, Liang Y, Song M. MicroRNA MiR-130a-3p promotes gastric cancer by targeting Glucosaminyl N-acetyl transferase 4 (GCNT4) to regulate the TGF-beta1/SMAD3 pathway. *Bioengineered*. 2021;12(2):11634-47.
396. Lan ZZ, Sun FH, Chen C, Niu L, Shi JD, Zhang WY. CircPRDM5 inhibits the proliferation, migration, invasion, and glucose metabolism of gastric cancer cells by reducing GCNT4 expression in a miR-485-3p-dependent manner. *The Kaohsiung Journal of Medical Sciences*. 2024;40(3):231-43.
397. Zhang H, Meng F, Wu S, Kreike B, Sethi S, Chen W, et al. Engagement of I-branching beta-1, 6-N-acetylglucosaminyltransferase 2 in breast cancer metastasis and TGF-beta signaling. *Cancer Research*. 2011;71(14):4846-56.
398. Thomas D, Rathinavel AK, Radhakrishnan P. Altered glycosylation in cancer: A promising target for biomarkers and therapeutics. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*. 2021;1875(1):188464.
399. Proietto M, Crippa M, Damiani C, Pasquale V, Sacco E, Vanoni M, et al. Tumor heterogeneity: preclinical models, emerging technologies, and future applications. *Frontiers in Oncology*. 2023;13:1164535.
400. Gill DJ, Clausen H, Bard F. Location, location, location: new insights into O-GalNAc protein glycosylation. *Trends in Cell Biology*. 2011;21(3):149-58.
401. Wu H, Shi XL, Zhang HJ, Song QJ, Yang XB, Hu WD, et al. Overexpression of ST3Gal-I promotes migration and invasion of HCCLM3 in vitro and poor prognosis in human hepatocellular carcinoma. *OncoTargets and Therapy*. 2016;9:2227-36.
402. Wen KC, Sung PL, Hsieh SL, Chou YT, Lee OK, Wu CW, et al. alpha2,3-sialyltransferase type I regulates migration and peritoneal dissemination of ovarian cancer cells. *Oncotarget*. 2017;8(17):29013-27.
403. Bai R, Luan X, Zhang Y, Robbe-Masselot C, Brockhausen I, Gao Y. The expression and functional analysis of the sialyl-T antigen in prostate cancer. *Glycoconjugate Journal*. 2020;37(4):423-33.
404. Li Q, Ran P, Zhang X, Guo X, Yuan Y, Dong T, et al. Downregulation of N-Acetylglucosaminyltransferase GCNT3 by miR-302b-3p Decreases Non-Small Cell Lung Cancer (NSCLC) Cell Proliferation, Migration and Invasion. *Cellular Physiology and Biochemistry*. 2018;50(3):987-1004.
405. Schultz MJ, Swindall AF, Bellis SL. Regulation of the metastatic cell phenotype by sialylated glycans. *Cancer and Metastasis Reviews*. 2012;31(3-4):501-18.
406. Uemura T, Shiozaki K, Yamaguchi K, Miyazaki S, Satomi S, Kato K, et al. Contribution of sialidase NEU1 to suppression of metastasis of human colon cancer cells through desialylation of integrin beta4. *Oncogene*. 2009;28(9):1218-29.
407. Miyagi T, Yamaguchi K. Mammalian sialidases: physiological and pathological roles in cellular functions. *Glycobiology*. 2012;22(7):880-96.
408. Patten J, Wang K. Fibronectin in development and wound healing. *Advanced Drug Delivery Reviews*. 2021;170:353-68.
409. Freire-de-Lima L, Gelfenbeyn K, Ding Y, Mandel U, Clausen H, Handa K, et al. Involvement of O-glycosylation defining oncofetal fibronectin in epithelial-mesenchymal transition process. *Proceedings of the National Academy of Sciences*. 2011;108(43):17690-5.
410. Jolly MK, Boareto M, Huang B, Jia D, Lu M, Ben-Jacob E, et al. Implications of the Hybrid Epithelial/Mesenchymal Phenotype in Metastasis. *Frontiers in Oncology*. 2015;5:155.
411. Cheung KJ, Ewald AJ. A collective route to metastasis: Seeding by tumor cell clusters. *Science*. 2016;352(6282):167-9.
412. Saxena K, Jolly MK, Balamurugan K. Hypoxia, partial EMT and collective migration: Emerging culprits in metastasis. *Translational Oncology*. 2020;13(11):100845.
413. Mayor R, Etienne-Manneville S. The front and rear of collective cell migration. *Nature Reviews Molecular Cell Biology*. 2016;17(2):97-109.
414. Grada A, Otero-Vinas M, Prieto-Castrillo F, Obagi Z, Falanga V. Research Techniques Made Simple: Analysis of Collective Cell Migration Using the Wound Healing Assay. *Journal of Investigative Dermatology*. 2017;137(2):e11-e6.

415. Lintz M, Munoz A, Reinhart-King CA. The Mechanics of Single Cell and Collective Migration of Tumor Cells. *Journal of Biomechanical Engineering*. 2017;139(2):0210051-9.
416. Barriga EH, Mayor R. Adjustable viscoelasticity allows for efficient collective cell migration. *Seminars in Cell & Developmental Biology*. 2019;93:55-68.
417. Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology*. 2009;10(7):445-57.
418. Aggarwal V, Montoya CA, Donnenberg VS, Sant S. Interplay between tumor microenvironment and partial EMT as the driver of tumor progression. *iScience*. 2021;24(2):102113.
419. Donnenberg VS, Donnenberg AD, Zimmerlin L, Landreneau RJ, Bhargava R, Wetzel RA, et al. Localization of CD44 and CD90 positive cells to the invasive front of breast tumors. *Cytometry Part B: Clinical Cytometry*. 2010;78(5):287-301.
420. Papadaki MA, Stoupis G, Theodoropoulos PA, Mavroudis D, Georgoulas V, Agelaki S. Circulating Tumor Cells with Stemness and Epithelial-to-Mesenchymal Transition Features Are Chemoresistant and Predictive of Poor Outcome in Metastatic Breast Cancer. *Molecular Cancer Therapeutics*. 2019;18(2):437-47.

Appendix

9. Appendix – Supplementary Materials

Supplementary Table 1 List of items and reagents with accompanying catalogue number and company supplier.

Item	Catalogue Number	Company
Human umbilical vein endothelial cells (HUVECs)	C-12203 Lot No: 479Z016	PromoCell®
Dulbecco's Modified Eagle Medium (DMEM) / Nutrient Mixture F12 (F12) / GlutaMAX™	11514436	Gibco™
Roswell Park Memorial Institute Medium-1640 (RPMI-1640) with L-glutamine	10-040-CV	Corning®
Heat inactivated foetal calf serum (FCS)	11550356	Gibco™
Epidermal growth factor (EGF)	AF-100-15	PeptoTech®
Insulin	I9278-5ML	Sigma-Aldrich®
Hydrocortisone	H0888	Sigma-Aldrich®
Sodium pyruvate	11360070	Gibco™
Endothelial Cell Growth Medium 2	C-22011	PromoCell®
Supplement mix / endothelial cell growth medium 2	C-39216	PromoCell®
1.5mL microcentrifuge tube	10154671	Fisher Scientific
15mL falcon tubes	11849650	Fisher Scientific
50mL falcon tubes	11899640	Fisher Scientific
Cryovials	11777939	Fisher Scientific
T-25cm ² cell culture flask	11830765	Fisher Scientific
T-75cm ² cell culture flask	11884235	Fisher Scientific
Phosphate-buffered saline (PBS)	14190144	Fisher Scientific
Trypsin/ethylenediaminetetraacetic acid (EDTA)	15400054	Gibco™
Dimethyl sulfoxide (DMSO)	15303671	Corning™
MycoAlert® mycoplasma detection kit	LT07-701	Lonza
Trypan Blue Dye	1450021	Bio-Rad
Cell counting slide	1450011	Bio-Rad
13mm diameter coverslips	11588492	Thermo Fisher
24-well cell culture plate	CLS3337	Corning™
Paraformaldehyde (PFA)	P6148	Sigma-Aldrich®
Triton X-100	X100	Sigma-Aldrich®
Hydrogen peroxide	216763	Sigma-Aldrich®
Tris-buffered saline (TBS)	133756	Thermo Fisher
Avidin/Biotin blocking kit	SP-2001	Vector Laboratories
Biotinylated- <i>Helix pomatia</i> agglutinin (HPA)	L3382	Sigma-Aldrich®
Biotinylated- <i>Vicia villosa</i> agglutinin (VVA)	B-1235	Vector Laboratories
Bovine serum albumin (BSA)	A2153	Sigma-Aldrich®
Avidin-peroxidase	S5512	Sigma-Aldrich®
3,3'-diaminobenzidine tetrahydrochloride hydrate (DAB)	SK-4100	Vector Labs
Harris haematoxylin	MHS16	Sigma-Aldrich®
DPX mounting medium	03817340	Sigma-Aldrich®
CellMask™ deep red	C10046	Thermo Fisher

Supplementary Table 1 (Continued).

Alexa Fluor™ 488 HPA	L11271	Thermo Fisher
ProLong Gold antifade with DAPI	P36931	Thermo Fisher
α-GalNAc	A2795	Sigma-Aldrich®
70µm cell-strainer	10788201	Fisher Scientific
40µm cell-strainer	10737821	Fisher Scientific
Alexa Fluor™ 647 HPA	L32454	Invitrogen™
Bovine gelatine	G9391	Sigma-Aldrich®
TNF-α	H8916	Sigma-Aldrich®
8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS)	H1529	Sigma-Aldrich®
Fluoromount™ aqueous mounting medium	F4680	Sigma-Aldrich®
Concanavalin A (ConA)	L-1000	Vector Laboratories
12-well cell culture plate	10098870	Fisher Scientific
Mouse anti-E-cadherin monoclonal antibody	324102	Biologend
Mouse anti-vimentin monoclonal antibody	677802	Biologend
Secondary antibody Alexa Fluor™ 594 goat polyclonal anti-mouse IgG	A11005	Thermo Fisher
96-well cell culture plate	167008	Fisher Scientific
3-(4,5-dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide (MTT)	M5655-1G	Sigma-Aldrich®
IGEPAL CA-630®	I7771	Sigma-Aldrich®
Mouse monoclonal anti-N-cadherin	350802	Biologend
Radioimmunoprecipitation assay (RIPA) buffer	20-188	Merck
Protease inhibitor cocktail	P8340	Sigma-Aldrich®
Pierce bicinchoninic acid (BCA) protein assay	23225	Thermo Fisher
Dithiothreitol (DTT)	R0861	Thermo Fisher
4x Laemmli sample buffer	161-0747	Bio-Rad
8-16% Mini-PROTEAN® TGX Stain-Free™ Precast Gels	4568103	Bio-Rad
Precision Plus Protein Dual Colour Standards	1610374	Bio-Rad
Tris Base	BP152-1	Thermo Fisher
Glycine	BP381-1	Thermo Fisher
Sodium dodecyl sulphate (SDS)	080917	Thermo Fisher
Trans-Blot® Turbo™ Mini-size low fluorescence PVDF Membrane (0.45µm pore)	1704274	Bio-Rad
Trans-Blot® Turbo™ Mini-size Transfer Stacks	1704274	Bio-Rad
Trans-Blot® Turbo™ 1X Transfer Buffer	1704274	Bio-Rad
Skimmed milk powder	N/A	Marvel
Tween® 20	P1379	Sigma-Aldrich®
Mouse monoclonal anti-β-actin	66009-1	proteintech®
Goat polyclonal anti-mouse IgG horseradish peroxidase (HRP)-Conjugate	W4028	Promega
Clarity™ Western Enhanced Chemiluminescence (ECL) Substrate	170-5061	Bio-Rad

Supplementary Table 1 (Continued).

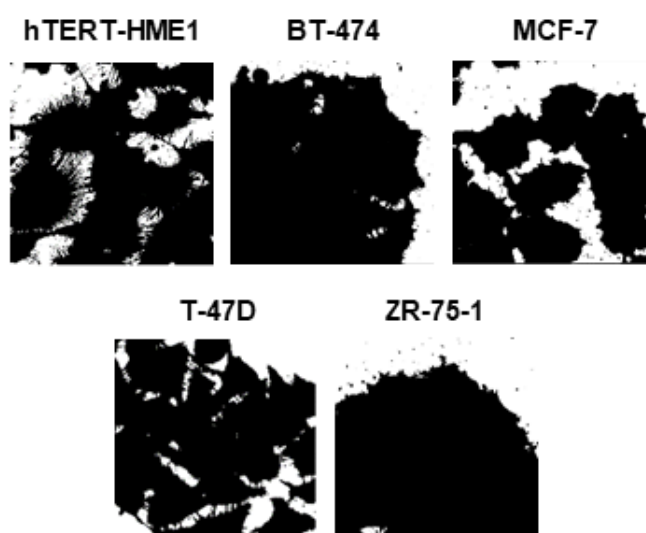
Transforming growth factor- β 1 (TGF- β 1)	100-21C	Thermo Fisher
6-well cell culture plate	CC7682-7506	Starlab
Sterile cell scraper	541070	Greiner Bio-One
Streptavidin-Dynabeads®	11205D	Sigma-Aldrich®
RNeasy® Mini Kit (50)	74104	QIAGEN GmbH
Ambion™ nuclease-free water	AM9938	Invitrogen™
RNase-free DNase set (50)	79254	QIAGEN GmbH
RNA ScreenTape	5067-5576	Agilent Technologies
RNA ScreenTape sample buffer	5067-5577	Agilent Technologies
RNA ScreenTape ladder	5067-5578	Agilent Technologies
Mr Frosty™ freezing container	5100-0001	Nalgene®

```

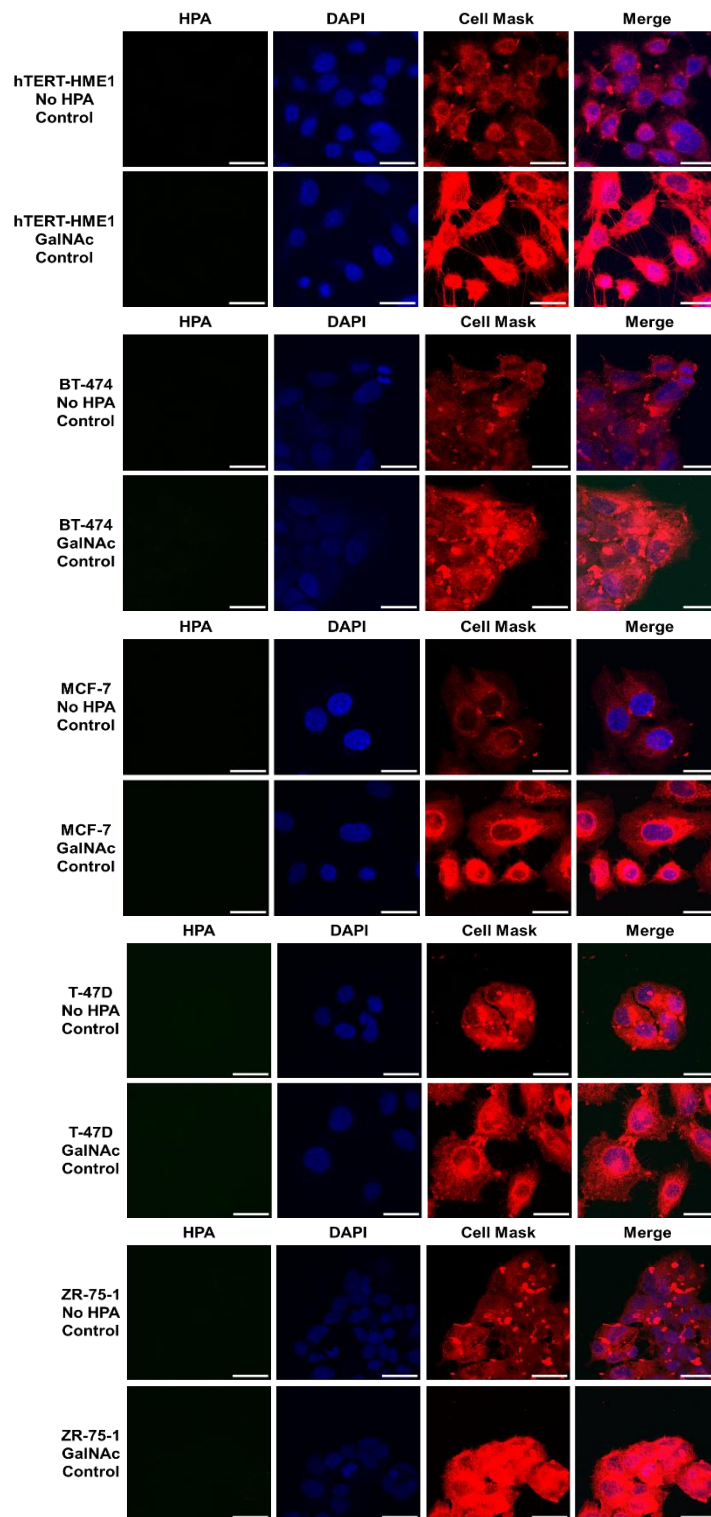
1  a=getTitle();
2  run("Z Project...", "projection=[Sum Slices]");
3  //run("Channels Tool...");
4  run("Split Channels");
5  selectWindow("C2-SUM_" +a);
6  setOption("ScaleConversions", true);
7  run("8-bit");
8  run("Auto Threshold", "method=MinError(I) white");
9  run("Watershed");
10 run("Analyze Particles...", "size=200-Infinity show=Overlay overlay add");
11 selectWindow("C1-SUM_" +a);
12 roiManager("Show None");
13 roiManager("Show All");
14
15

```

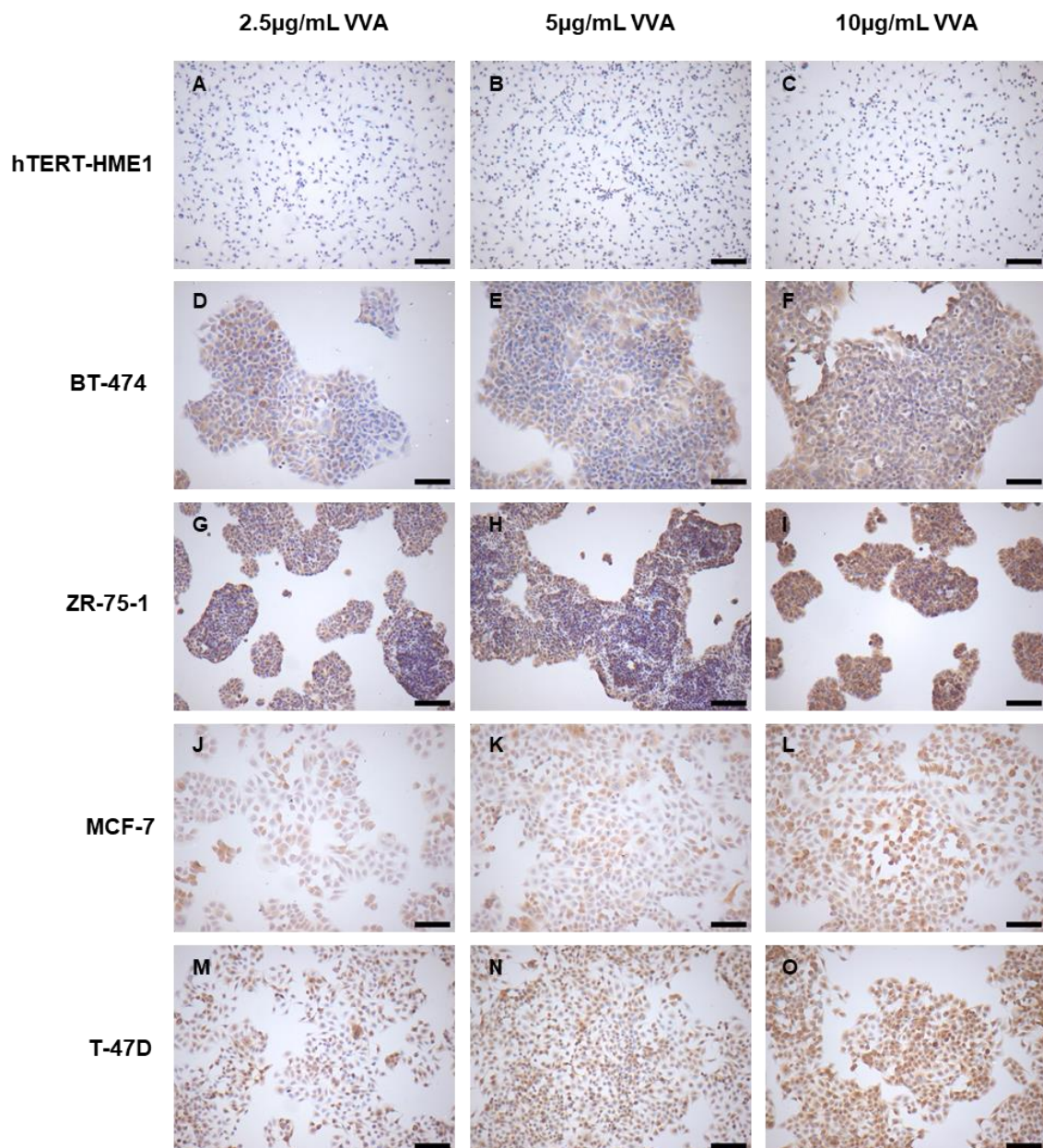
Supplementary Figure 1 FIJI (NIH, Bethesda, MD) ImageJ macro script used to aid the analysis of fluorescent *Helix pomatia* agglutinin (HPA) labelling.



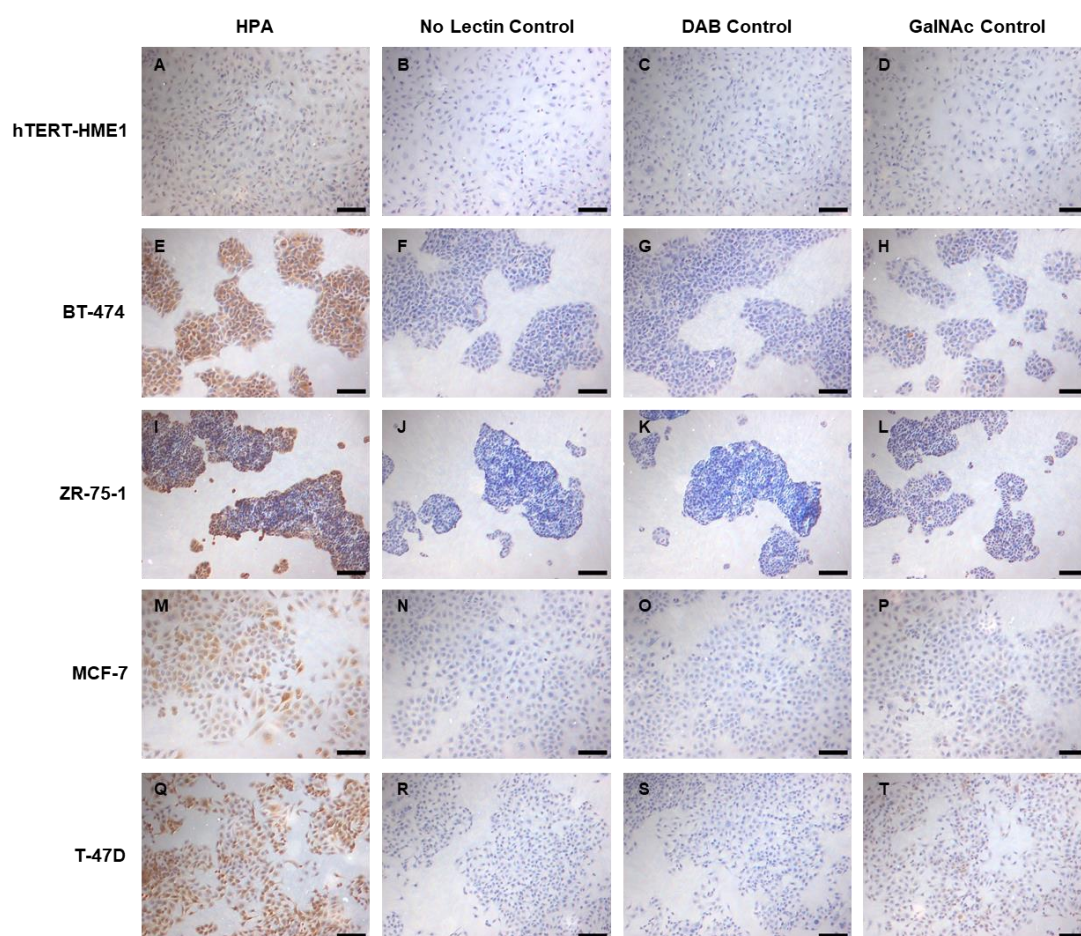
Supplementary Figure 2 Representative images of MinError(I) thresholding of the model breast epithelial cell lines used to identify cells from the background.



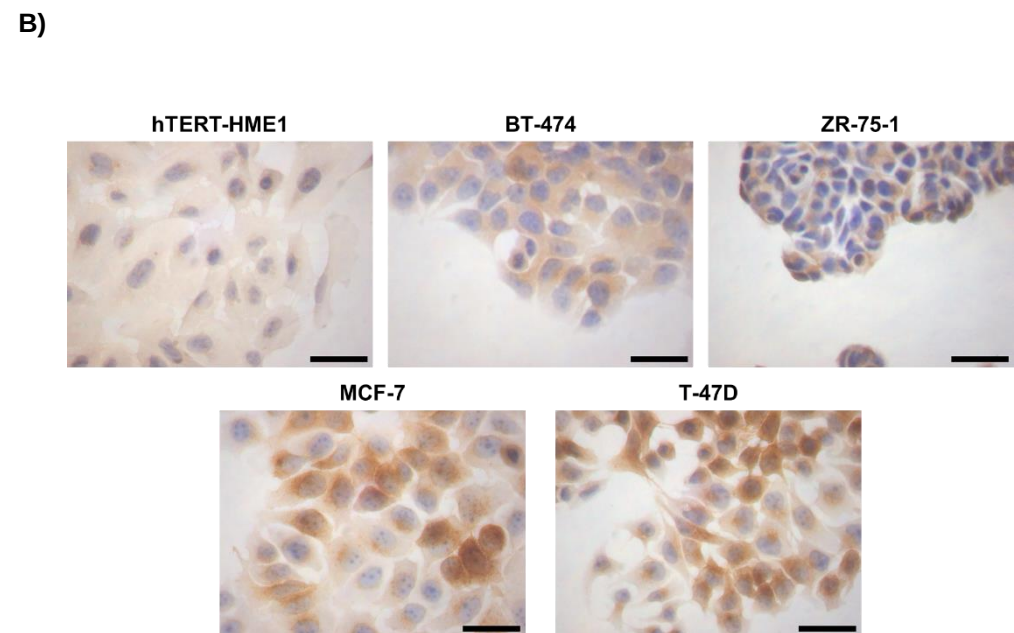
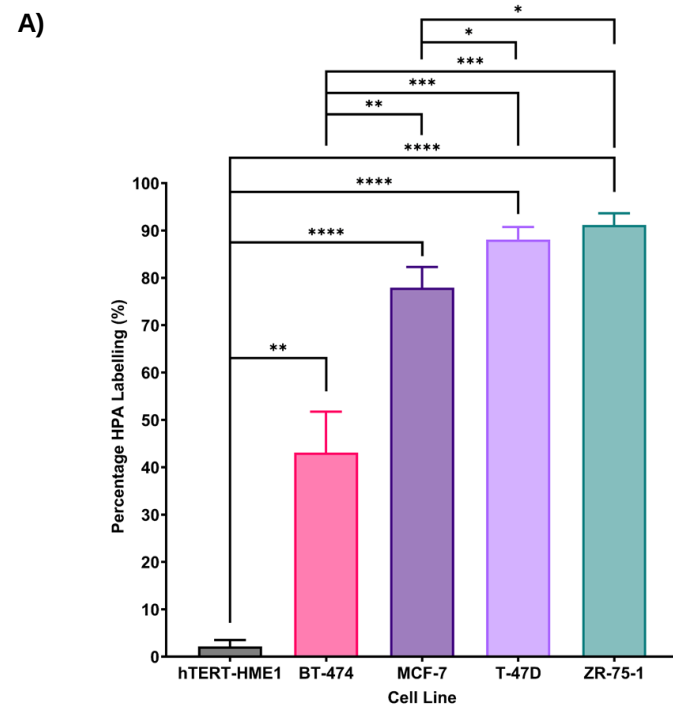
Supplementary Figure 3 Fluorescent *Helix pomatia* agglutinin (HPA) labelling controls. Representative Max Intensity Z-stack projections of no HPA control and N-acetylgalactosamine (GalNAc) control in hTERT-HME1, BT-474, MCF-7, T-47D, and ZR-75-1 cell lines. HPA (green), CellMask™ (red) and DAPI (blue). Scale bars = 25µm.



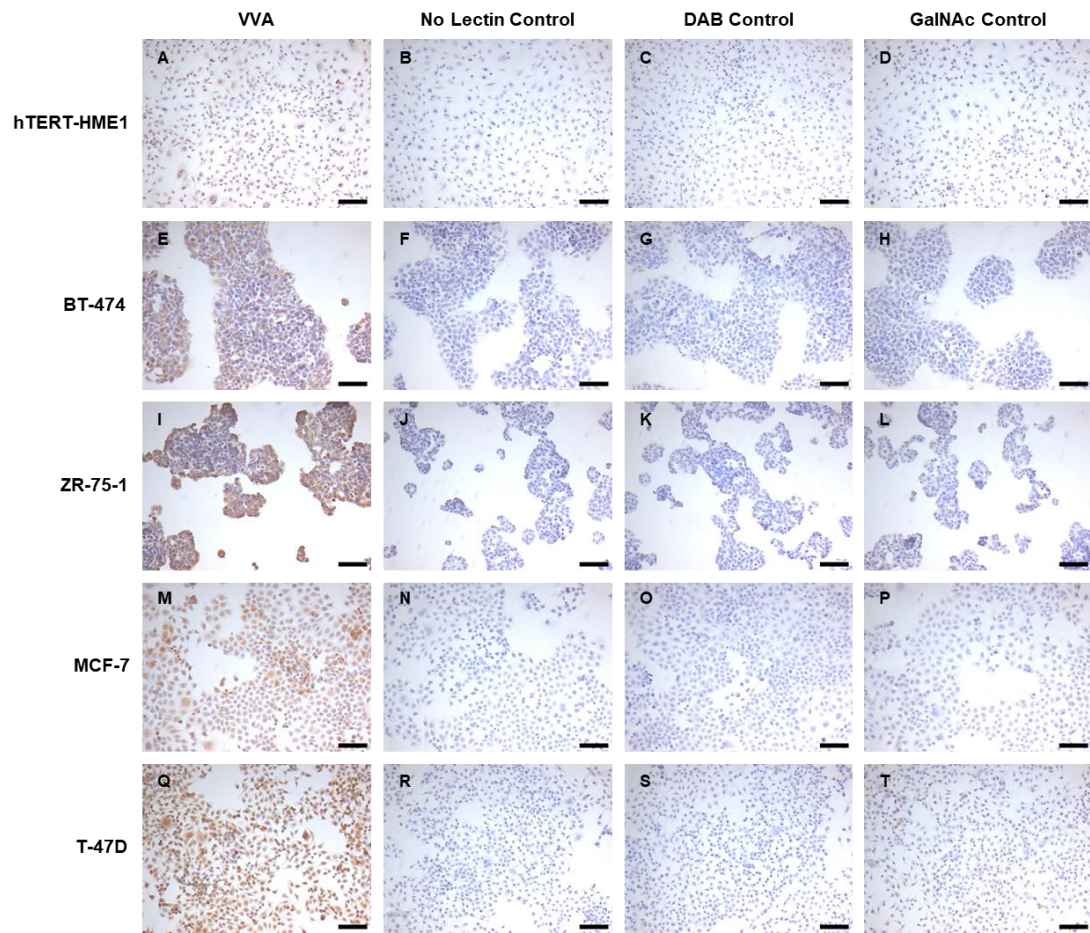
Supplementary Figure 4 *Vicia villosa* agglutinin (VVA) lectin cytochemistry concentration optimisation. hTERT-HME1 (A-C), BT-474 (D-F), ZR-75-1 (G-I), MCF-7 (J-L), T-47D (M-O) cells incubated with either 2.5µg/mL, 5µg/mL or 10µg/mL VVA for 5 minutes. Images are representatives of three biological replicates, each replicate consisting of three technical replicates per cell line. Scale bar = 10µm.



Supplementary Figure 5 *Helix pomatia* agglutinin (HPA) lectin cytochemistry of non-permeabilized cells. hTERT-HME1 (A-D), BT-474 (E-H), ZR-75-1 (I-L), MCF-7 (M-P), T-47D (Q-T) cells incubated with 10µg/mL HPA for 5 minutes. Images are representatives of three biological replicates, each replicate consists of three technical replicates per cell line. The HPA lectin was omitted in the no lectin control. Both HPA and avidin peroxidase were omitted in the 3,3'-Diaminobenzidine tetrahydrochloride hydrate (DAB) control. Finally, a N-acetylgalactosamine (GalNAc) control was conducted as a specificity control. Scale bar = 10µm.

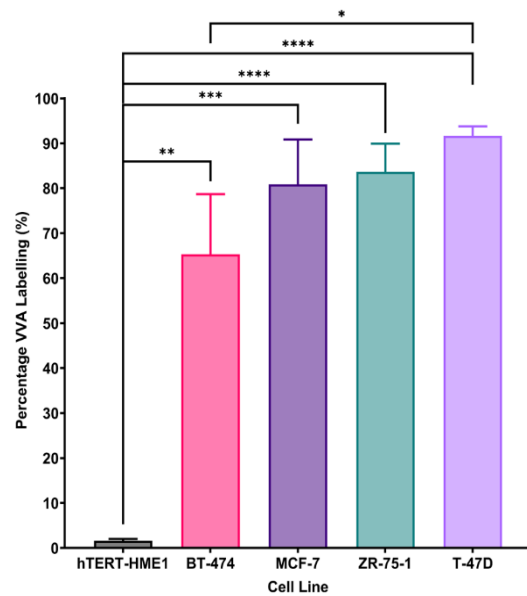


Supplementary Figure 6 Quantification of *Helix pomatia* agglutinin (HPA) lectin cytochemistry in non-permeabilized cells. A) hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D cell lines were investigated, with three biological replicates, each consisting of three technical replicates ($\pm$ SD). HPA-positive cells were identified as any cell that displayed brown cytoplasmic labelling of intensity equal to or greater than ++ (i.e. all those deemed ++, +++ or ++++). Those deemed to be + or less were considered to be HPA-negative. The percentage of HPA-positive cells was calculated by dividing the number of HPA-positive cells by the total number of cells in the image. hTERT-HME1 2.1% (SD 1.4%), BT-474 43.1% (SD 8.6%), MCF-7 77.9% (SD 4.4%), T-47D 88.1% (SD 2.6%), ZR-75-1 91.2% (SD 2.4%). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. B) Representative images of hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D HPA lectin cytochemistry of non-permeabilized cells. Scale bar = $2\mu\text{m}$.

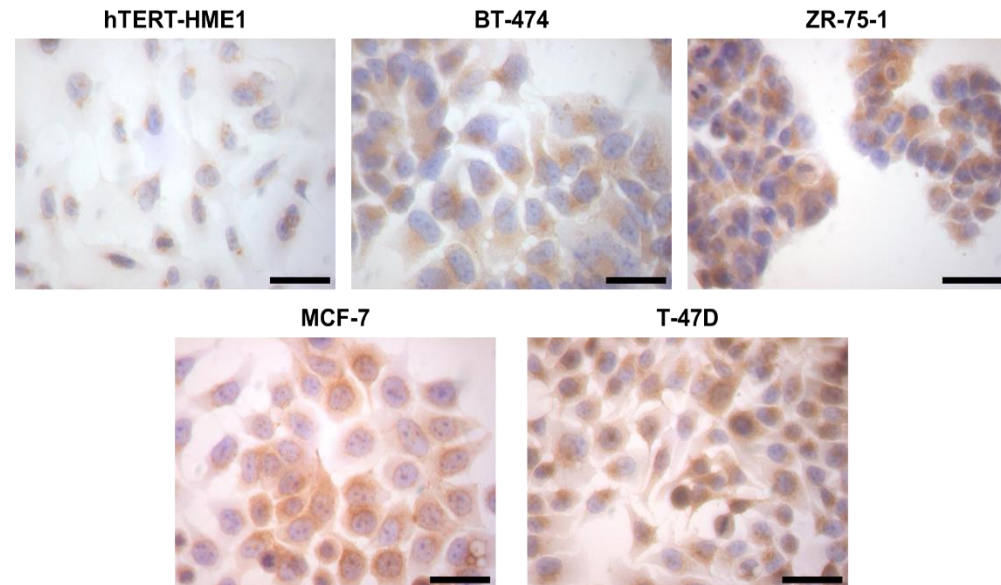


Supplementary Figure 7 *Vicia villosa* agglutinin (VVA) lectin cytochemistry of non-permeabilized cells. hTERT-HME1 (A-D), BT-474 (E-H), ZR-75-1 (I-L), MCF-7 (M-P), T-47D (Q-T) cells incubated with 5µg/mL VVA for 5 minutes. Images are representatives of three biological replicates, each replicate consisting of three technical replicates per cell line. The VVA lectin was omitted in the no lectin control. Both VVA and avidin peroxidase were omitted in the 3,3'-Diaminobenzidine tetrahydrochloride hydrate (DAB) control. Finally, a N-acetylgalactosamine (GalNAc) control was conducted as a specificity control. Scale bar = 10µm.

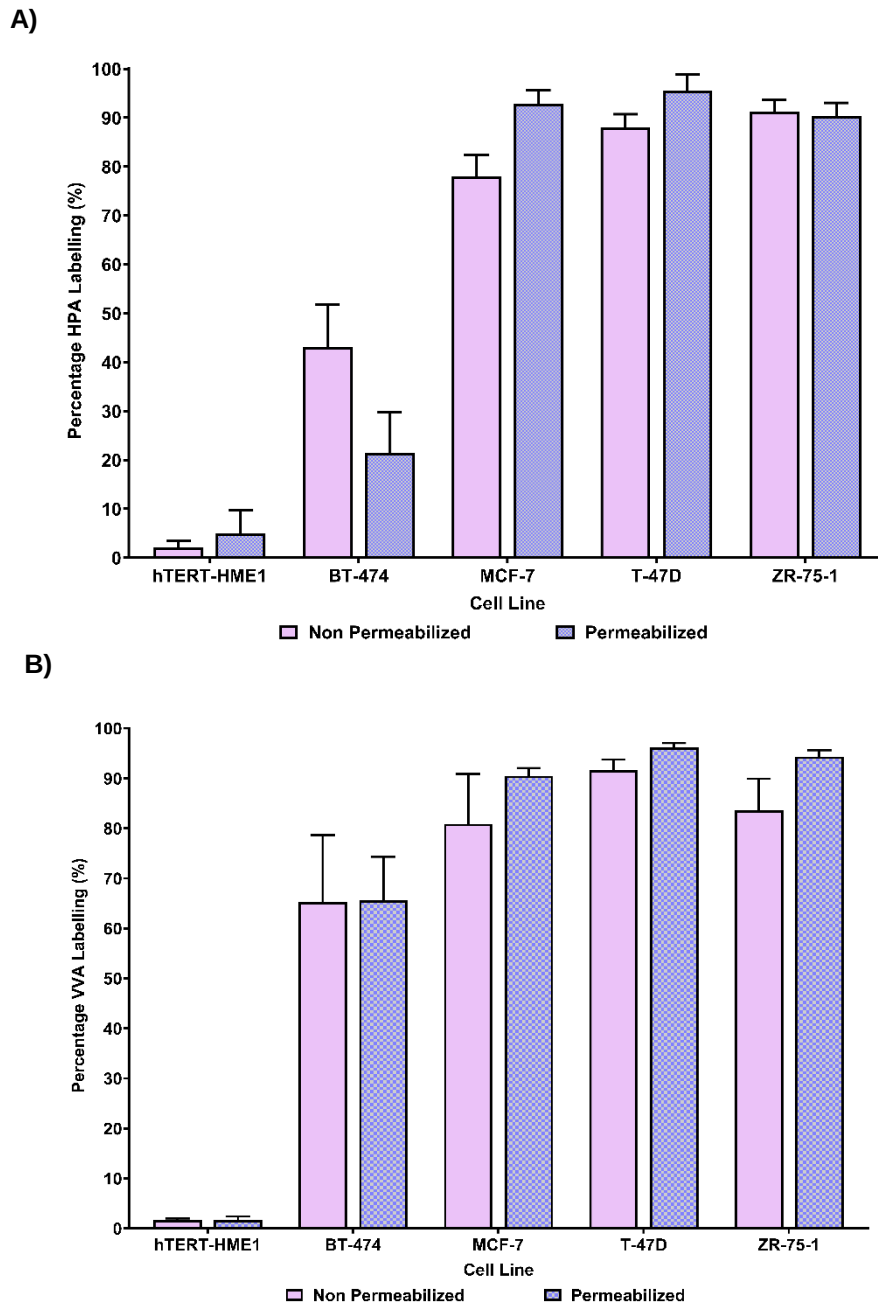
A)



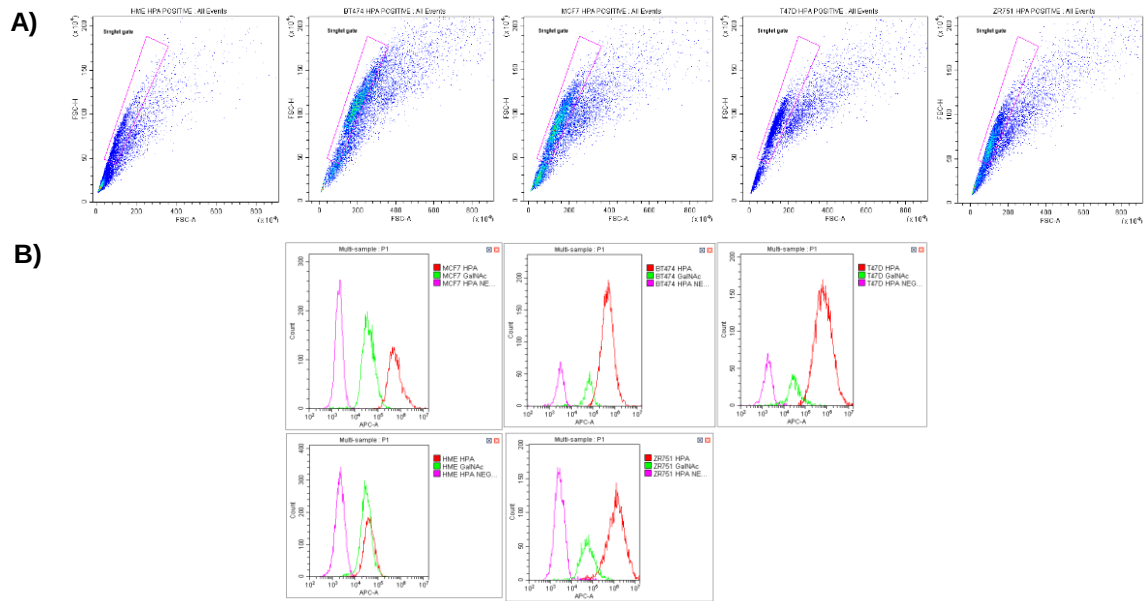
B)



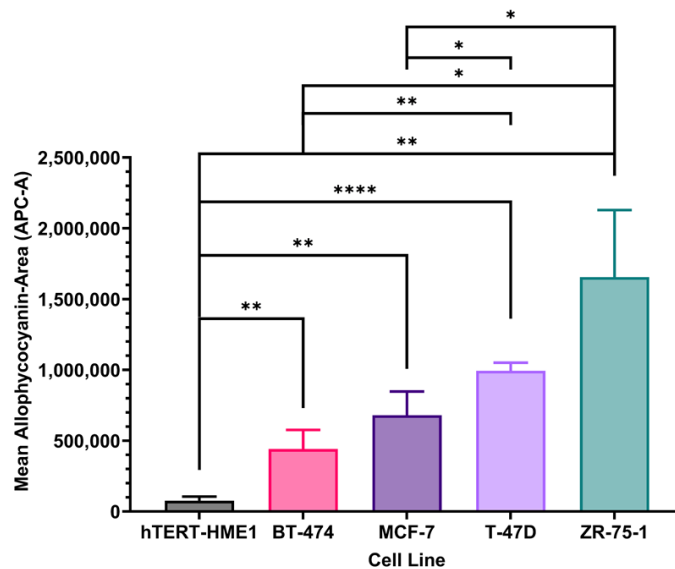
Supplementary Figure 8 Quantification of *Vicia villosa* agglutinin (VVA) lectin cytochemistry of non-permeabilized cells. A) Average percentage of VVA lectin labelling of hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D cell lines, with three biological replicates, each consisting of three technical replicates ($\pm$ SD). VVA-positive cells were identified as any cell that displayed brown cytoplasmic labelling of intensity equal to or greater than ++ (i.e. all those deemed ++, +++ or ++++). Those deemed to be + or less were considered to be VVA-negative. The percentage of VVA-positive cells was calculated by dividing the number of VVA-positive cells by the total number of cells in the image. hTERT-HME1 1.6% (SD 0.4%), BT-474 65.3% (SD 13.4%), MCF-7 80.9% (SD 10.0%), ZR-75-1 83.6% (SD 6.3%), T-47D 91.6% (SD 2.1%). Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. B) Representative images of hTERT-HME1, BT-474, ZR-75-1, MCF-7 and T-47D VVA lectin cytochemistry of non-permeabilized cells. Scale bar = 2 μ m.



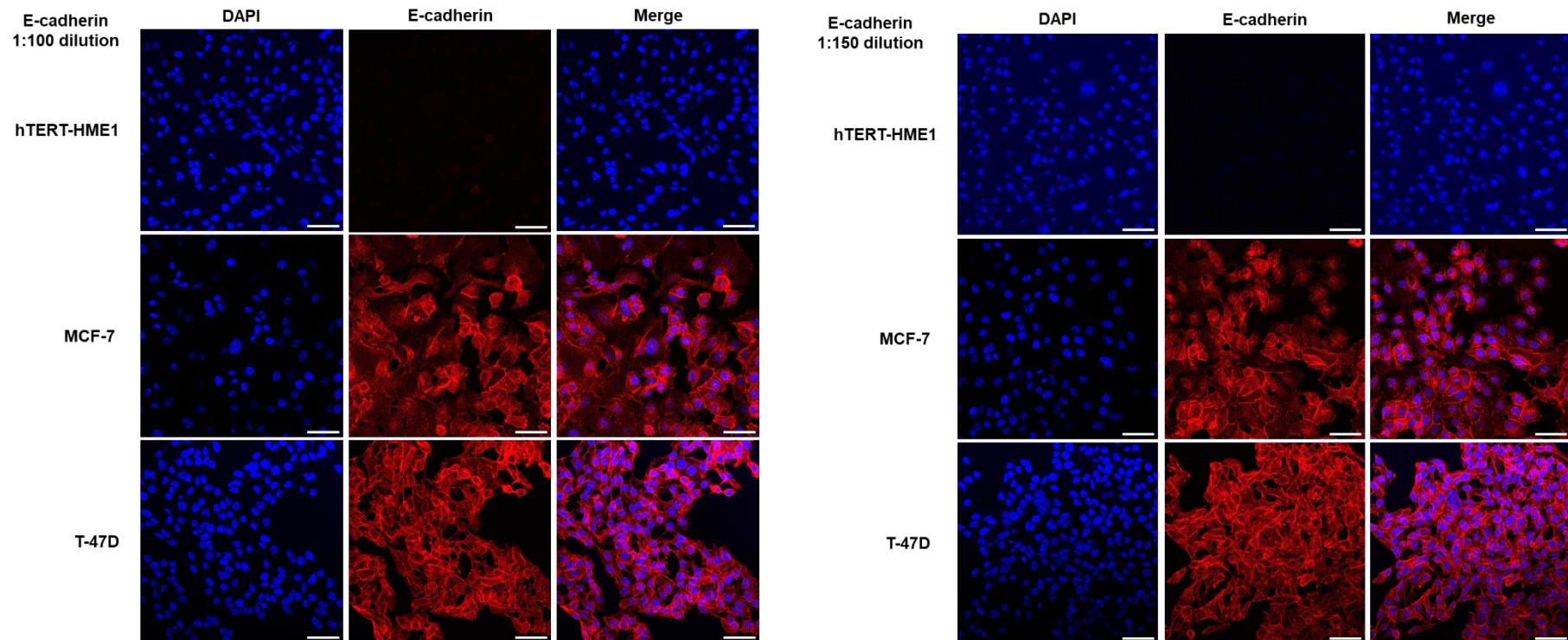
Supplementary Figure 9 Comparison of biotinylated lectin cytochemistry of permeabilized and non-permeabilized cells. A) Comparison of *Helix pomatia* agglutinin (HPA) lectin labelling in non-permeabilized and permeabilized cells. B) Comparison of *Vicia villosa* agglutinin (VVA) lectin labelling in non-permeabilized and permeabilized cells. A Wilcoxon matched-pair signed rank test was used to test for significance, no significant difference was detected.



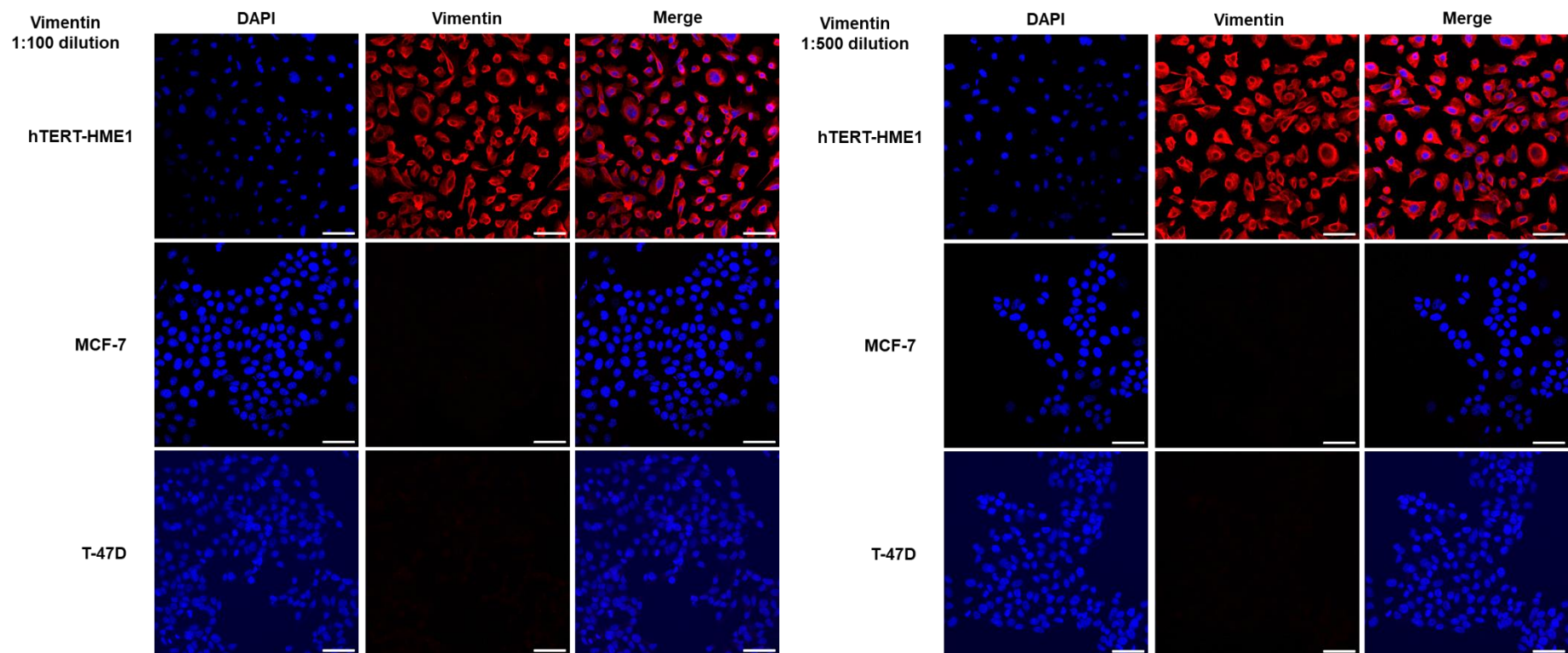
Supplementary Figure 10 *Helix pomatia* agglutinin (HPA) flow cytometry gating strategy. A) Representative singlet gating and excluding cell debris used in flow cytometry analysis. B) Gated histogram profiles of breast epithelial cells and their respective controls. HPA negative control (omitting HPA lectin, pink histogram), and N-acetylgalactosamine (GalNAc) competitive inhibition control (addition of GalNAc sugar with HPA lectin, green histogram) and test condition (cells with HPA, red histogram).



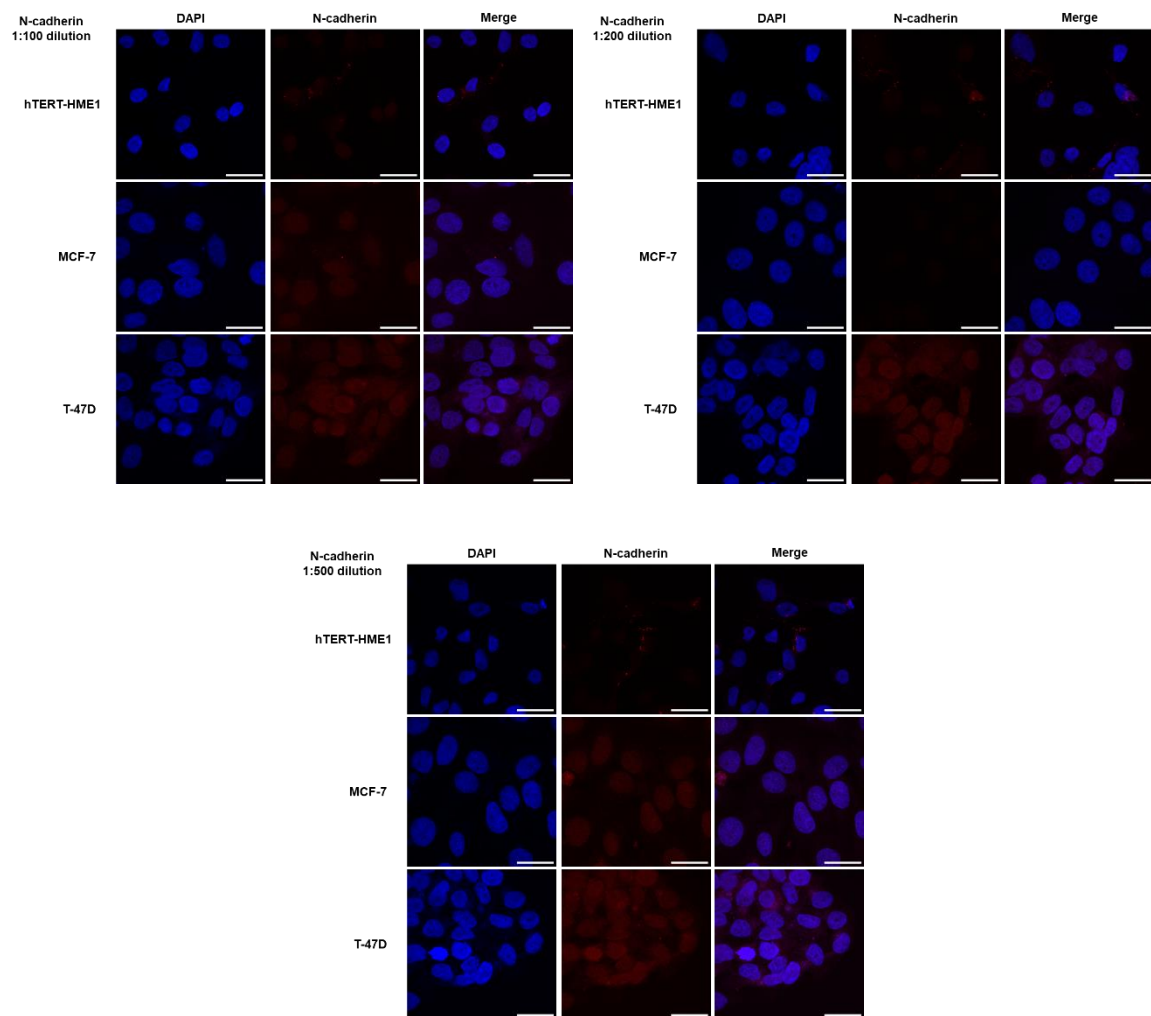
Supplementary Figure 11 Average Allophycocyanin-Area (APC-A) *Helix pomatia* agglutinin (HPA) values ($\pm$ SD) assessed by flow cytometry. Representative of three biological replicates. Unpaired Students t-test was used to test for significance, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.



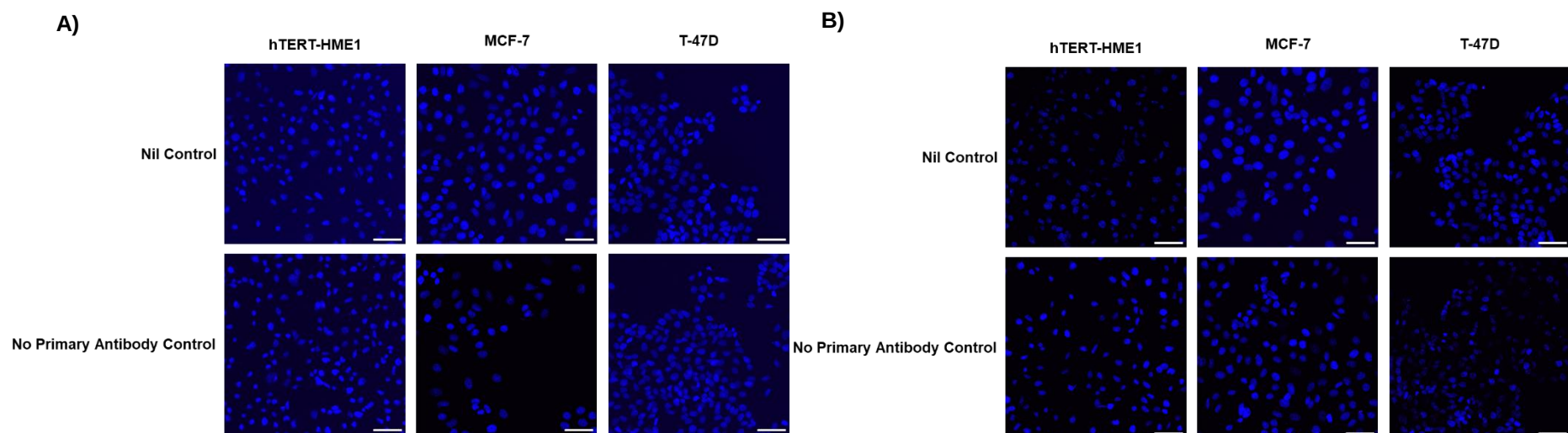
Supplementary Figure 12 E-cadherin primary antibody optimisation for epithelial-mesenchymal transition (EMT) marker immunofluorescence. Representative Max Intensity Z-stack projections of 1:100 and 1:150 dilutions of the E-cadherin primary antibody in hTERT-HME1, MCF-7 and T-47D cell lines. E-cadherin (red) and DAPI (blue). Scale bar = 50µm.



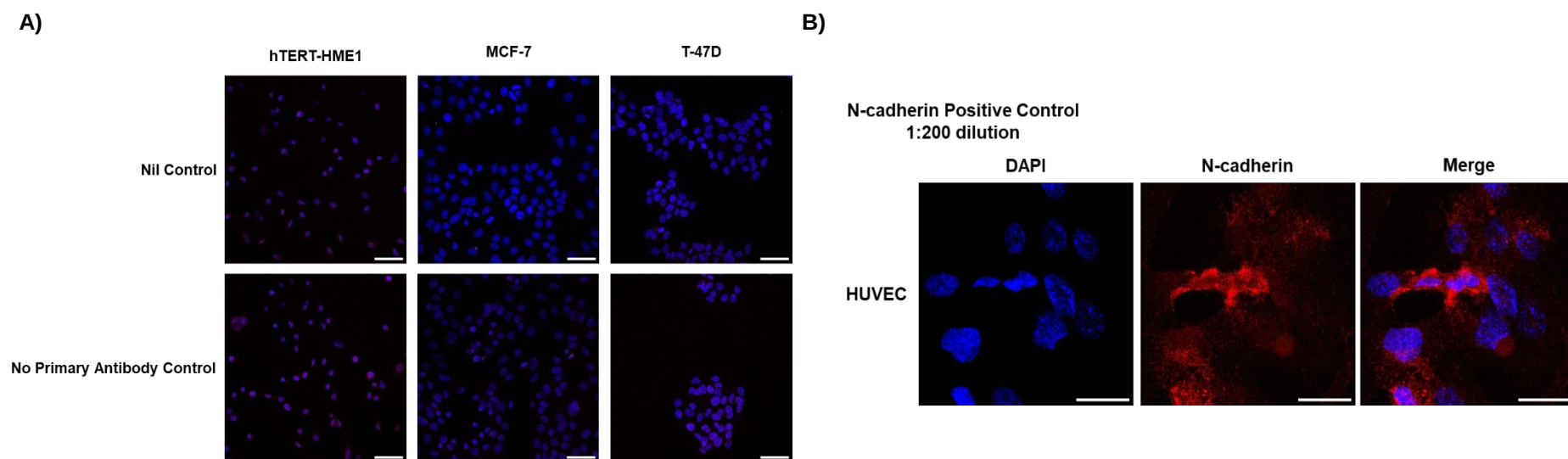
Supplementary Figure 13 Vimentin primary antibody optimisation for epithelial-mesenchymal transition (EMT) marker immunofluorescence. Representative Max Intensity Z-stack projections of 1:100 and 1:500 dilutions of the vimentin primary antibody in hTERT-HME1, MCF-7 and T-47D cell lines. Vimentin (red) and DAPI (blue). Scale bar = 50µm.



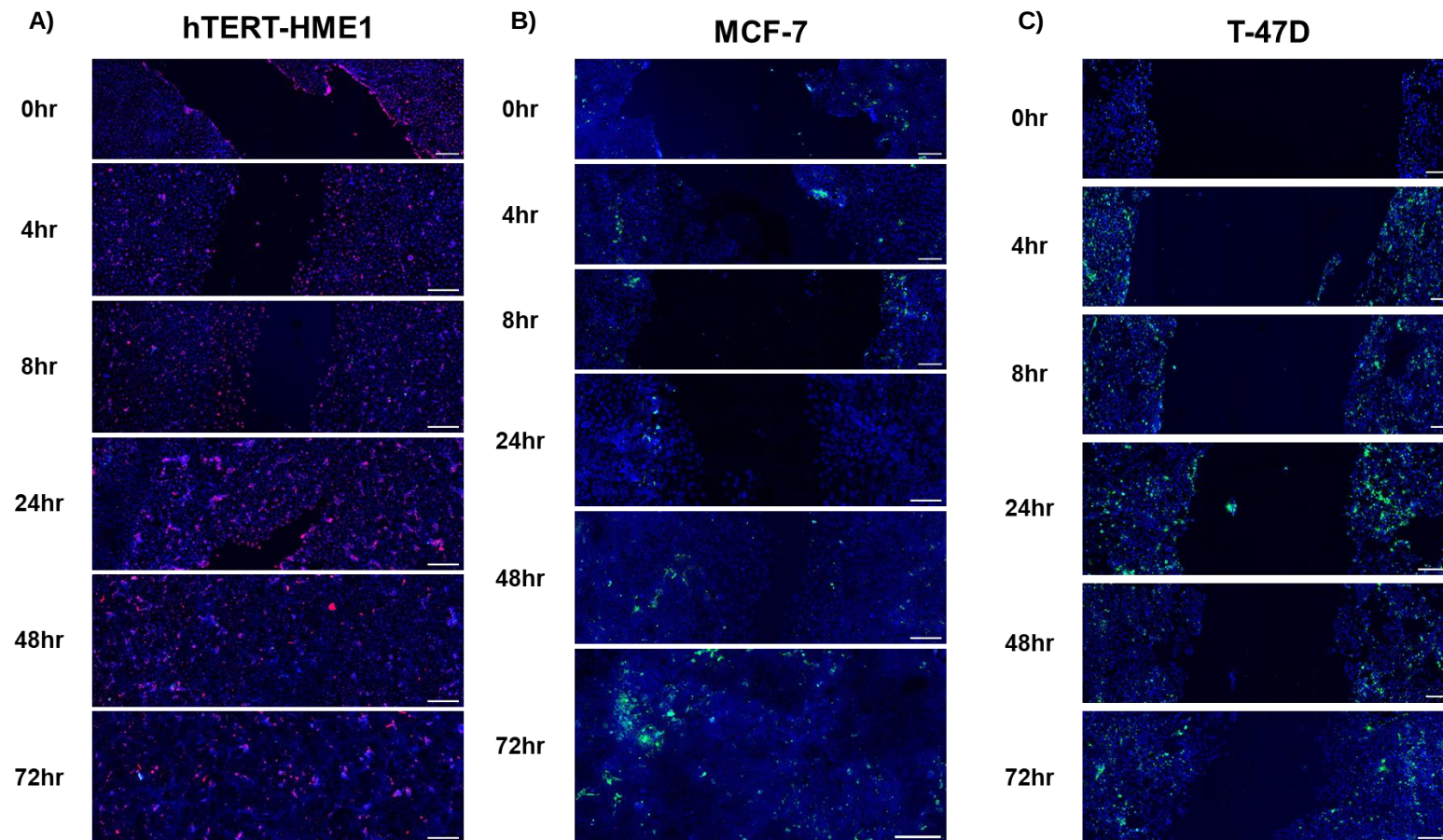
Supplementary Figure 14 N-cadherin primary antibody optimisation for epithelial-mesenchymal transition (EMT) marker immunofluorescence. Representative Max Intensity Z-stack projections of 1:100, 1:200 and 1:500 dilutions of the N-cadherin primary antibody in hTERT-HME1, MCF-7 and T-47D cell lines. N-cadherin (red) and DAPI (blue). Scale bar = 25μm.



Supplementary Figure 15 E-cadherin and vimentin immunofluorescence controls. Representative Max Intensity Z-stack projections of merged channel E-cadherin and vimentin controls in hTERT-HME1, MCF-7 and T-47D cell lines. A) E-cadherin immunofluorescence nil control and no primary antibody control. E-cadherin (red) and DAPI (blue). B) Vimentin immunofluorescence nil control and no primary antibody control. Vimentin (red) and DAPI (blue). Scale bar = 50 μ m.

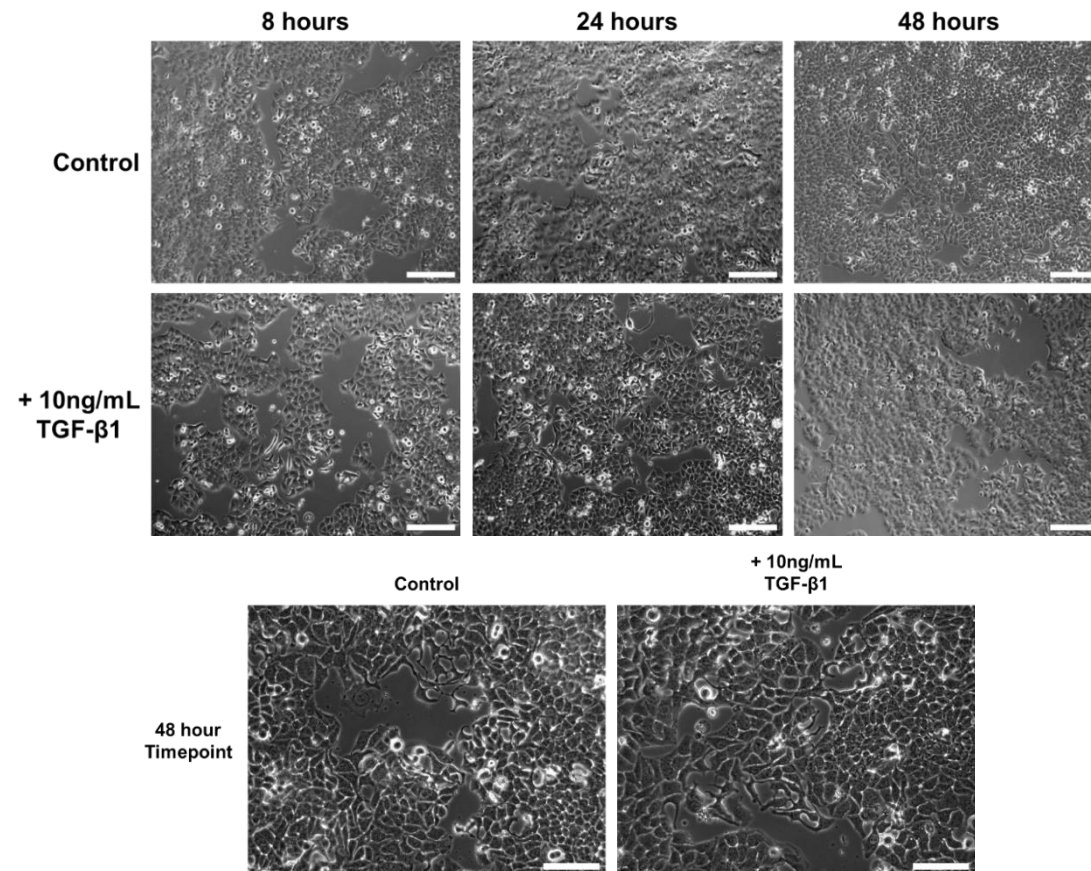


Supplementary Figure 16 N-cadherin immunofluorescence controls. A) Representative images of merged channel N-cadherin nil control and no primary antibody control in the hTERT-HME1, MCF-7 and T-47D cell lines. N-cadherin (red) and DAPI (blue). Scale bar = 50µm. B) Representative Max Intensity Z-stack projection of HUVEC cell line as a positive control at optimised 1:200 dilution of primary N-cadherin antibody. N-cadherin (red) and DAPI (blue). Scale bar = 25µm.

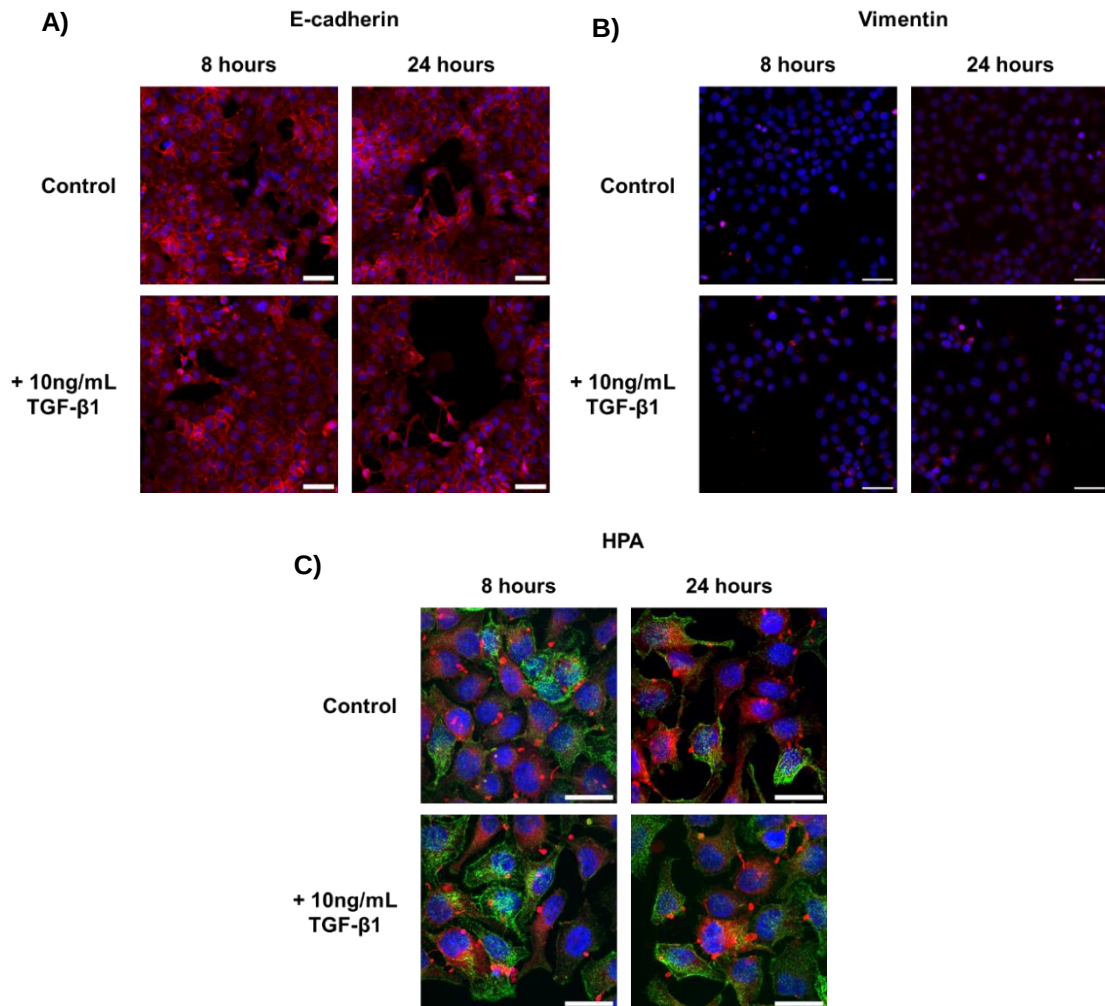


Supplementary Figure 17 Vimentin immunofluorescence and fluorescent *Helix pomatia* agglutinin (HPA) labelling of the wound healing assay. Representative Max Intensity Z-stack projections of wound healing assays at 0, 4, 8, 24, 48 and 72 hours for A) hTERT-HME1, B) MCF-7 and C) T-47D cell lines. Labelled with HPA (green), vimentin (red), and DAPI nucleus staining (blue). Scale bar = 150µm.

T-47D Optimisation

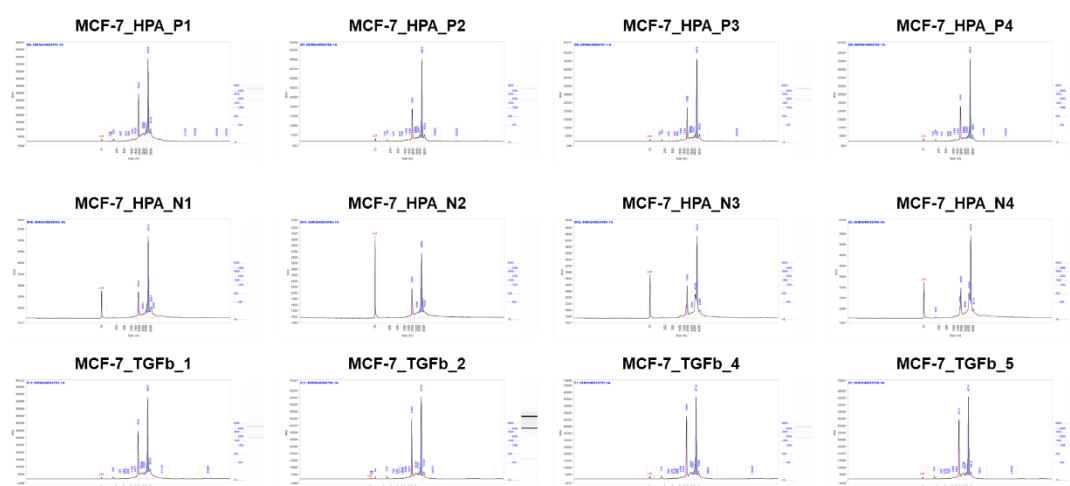


Supplementary Figure 18 T-47D cell line transforming growth factor-β1 (TGF-β1) treatment optimisation – morphology assessment. Visual morphology inspections showed little to no changes at the 48-hour treatment timepoint. Scale bar = 50 μm.

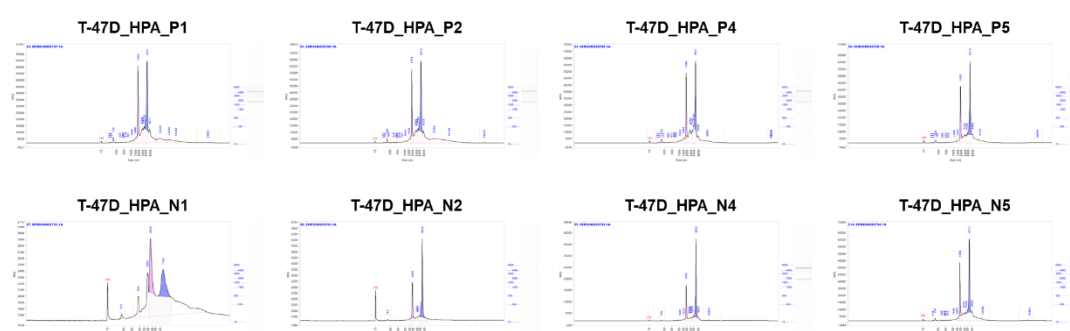


Supplementary Figure 19 T-47D cell line transforming growth factor- β 1 (TGF- β 1) treatment optimisation – E-cadherin, vimentin and *Helix pomatia* agglutinin (HPA) assessment. Visual inspections of E-cadherin, vimentin and HPA labelling showed little to no changes at the 8-hour and 24-hour treatment timepoints. Scale bar = 50 μ m for A) and B). Scale bar = 25 μ m for C).

**Total RNA Sample QC
MCF-7 Integrity Results**



**Total RNA Sample QC
T-47D Integrity Results**



Group Key

MCF_HPA_P = MCF-7 HPA-positive	T47D_HPA_N = T-47D HPA-negative
MCF_HPA_N = MCF-7 HPA-negative	MCF7_TGfb = MCF-7 TGF- β 1 treated
T47D_HPA_P = T-47D HPA-positive	

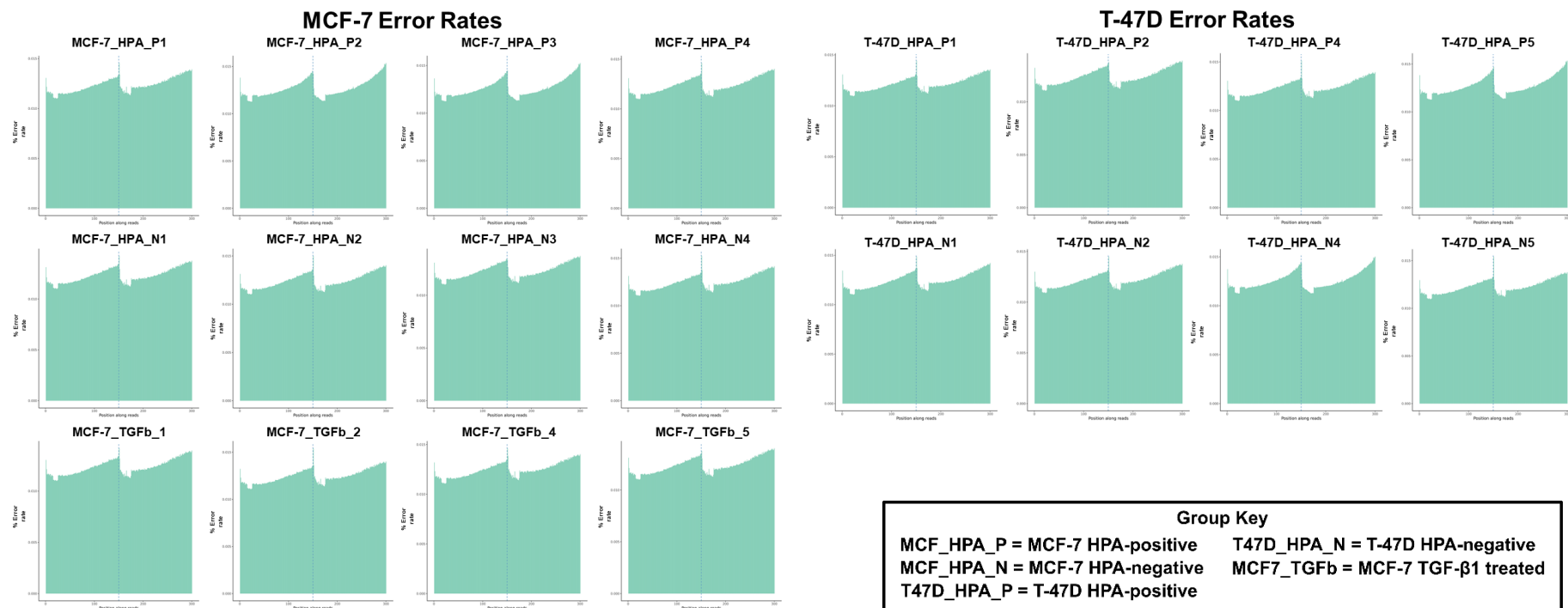
Supplementary Figure 20 Novogene Agilent 5400 RNA sample quality and integrity control – RNA integrity number (RIN) results. RIN estimates sample integrity using gel electrophoresis and analysis of 28S and 18S ribosomal band ratio.

Supplementary Table 2 Novogene Agilent 5400 RNA sample quality control result summary information. All samples passed the sample quality control and had high scores for RNA integrity numbers (RIN) (RIN = 10 is the highest sample quality score and RIN = 1 is the lowest sample quality score).

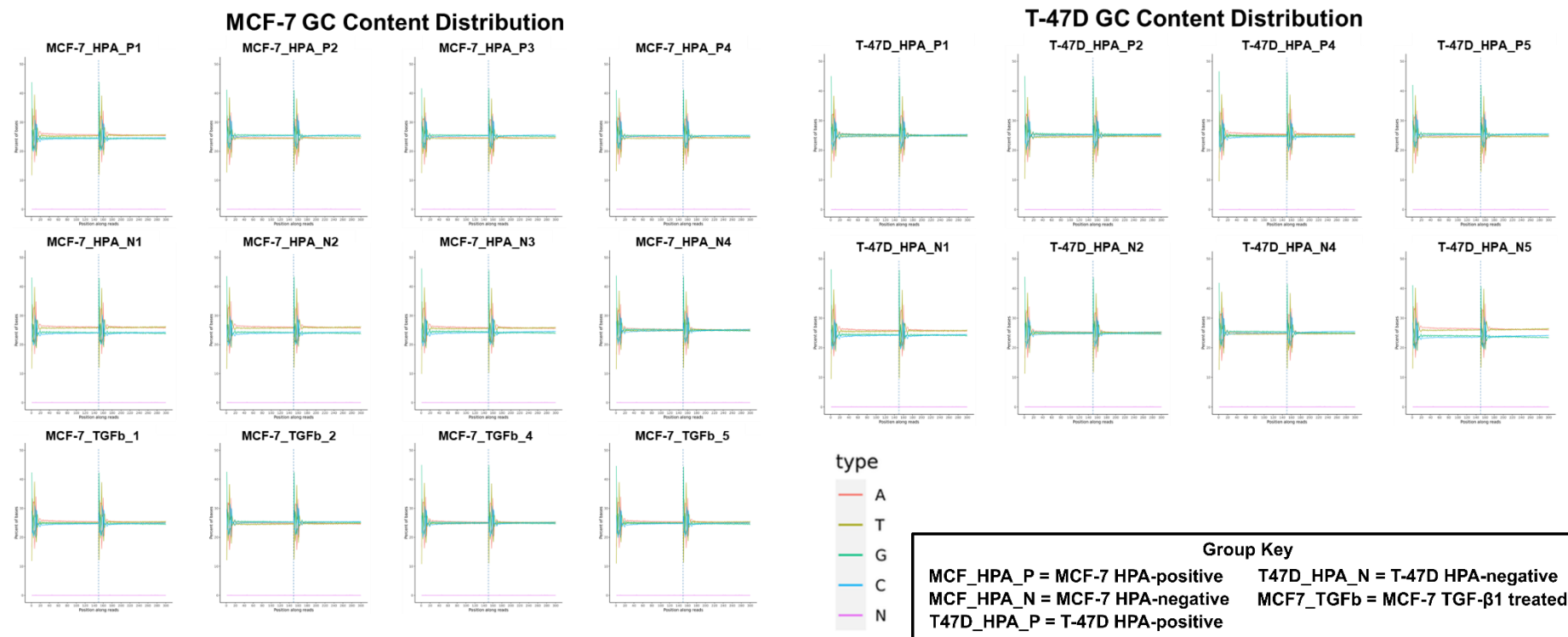
No	Sample	Nucleic Acid ID	Concentration (ng/μl)	Volume (μl)	Total amount (μg)	RIN	Sample QC Results
1	MCF7_HPA_P1	EKRN240028779-1A	763.2	9	6.9	9.8	Pass
2	MCF7_HPA_P2	EKRN240028780-1A	442.3	9	4.0	9.9	Pass
3	MCF7_HPA_P3	EKRN240028781-1A	590.6	9	5.3	9.9	Pass
4	MCF7_HPA_P4	EKRN240028782-1A	447.3	9	4.3	9.9	Pass
5	MCF7_HPA_N1	EKRN240028783-1A	55.9	9	0.5	9.6	Pass
6	MCF7_HPA_N2	EKRN240028784-1A	21.3	9	0.2	9.6	Pass
7	MCF7_HPA_N3	EKRN240028785-1A	44.5	10	0.4	9.7	Pass
8	MCF7_HPA_N4	EKRN240028786-1A	53.2	11	0.6	9.5	Pass
9	T47D_HPA_P1	EKRN240028787-1A	1692.5	9	15.2	9.6	Pass
10	T47D_HPA_P2	EKRN240028788-1A	1460.2	10	14.6	9.5	Pass
11	T47D_HPA_P4	EKRN240028789-1A	1171.9	10	11.7	9.7	Pass
12	T47D_HPA_P5	EKRN240028790-1A	864.8	10	8.6	9.8	Pass
13	T47D_HPA_N1	EKRN240028791-1A	164.7	12	2.0	8.2	Pass
14	T47D_HPA_N2	EKRN240028792-1A	46.6	12	0.6	9.8	Pass
15	T47D_HPA_N4	EKRN240028793-1A	1051.9	12	12.6	9.9	Pass
16	T47D_HPA_N5	EKRN240028794-1A	816.8	12	9.8	9.9	Pass
17	MCF7_TGFb_1	EKRN240028795-1A	978.8	9	8.8	9.9	Pass
18	MCF7_TGFb_2	EKRN240028796-1A	3705.0	9	33.3	9.6	Pass
19	MCF7_TGFb_4	EKRN240028797-1A	703.0	9	6.3	9.9	Pass
20	MCF7_TGFb_5	EKRN240028798-1A	597.4	9	5.4	9.9	Pass

Supplementary Table 3 Overview of sample sampling data quality. Average sequencing error rate, the Q20 and Q30 scores indicating the percentage of bases whose correct base recognition rates were greater than 99% and 99.9% respectively and the percentage of guanine and cytosine bases in the samples.

Sample	Library	error_rate	Q20	Q30	GC_pct
MCF7_HPA_P1	ERAS240028779-1r	0.01	98.43	95.5	48.94
MCF7_HPA_P2	ERAS240028780-2r	0.01	98.22	94.97	50.89
MCF7_HPA_P3	ERAS240028781-2r	0.01	98.27	95.08	50.8
MCF7_HPA_P4	ERAS240028782-1r	0.01	98.37	95.51	50.7
MCF7_HPA_N1	ERAS240028783-1r	0.01	98.44	95.59	48.13
MCF7_HPA_N2	ERAS240028784-1r	0.01	98.35	95.43	48.17
MCF7_HPA_N3	ERAS240028785-1r	0.01	98.47	95.66	48.48
MCF7_HPA_N4	ERAS240028786-1r	0.01	98.33	95.28	49.91
T47D_HPA_P1	ERAS240028787-1r	0.01	98.57	95.86	50.14
T47D_HPA_P2	ERAS240028788-1r	0.01	98.4	95.49	50.6
T47D_HPA_P4	ERAS240028789-1r	0.01	98.42	95.52	49.35
T47D_HPA_P5	ERAS240028790-2r	0.01	98.22	94.94	50.79
T47D_HPA_N1	ERAS240028791-1r	0.01	98.35	95.37	48.46
T47D_HPA_N2	ERAS240028792-1r	0.01	98.5	95.77	49.86
T47D_HPA_N4	ERAS240028793-2r	0.01	98.32	95.17	50.41
T47D_HPA_N5	ERAS240028794-1r	0.01	98.47	95.66	47.64
MCF7_TGFB_1	ERAS240028795-1r	0.01	98.41	95.51	49.46
MCF7_TGFB_2	ERAS240028796-1r	0.01	98.36	95.3	50.57
MCF7_TGFB_4	ERAS240028797-1r	0.01	98.36	95.3	49.85
MCF7_TGFB_5	ERAS240028798-1r	0.01	98.38	95.45	49.56

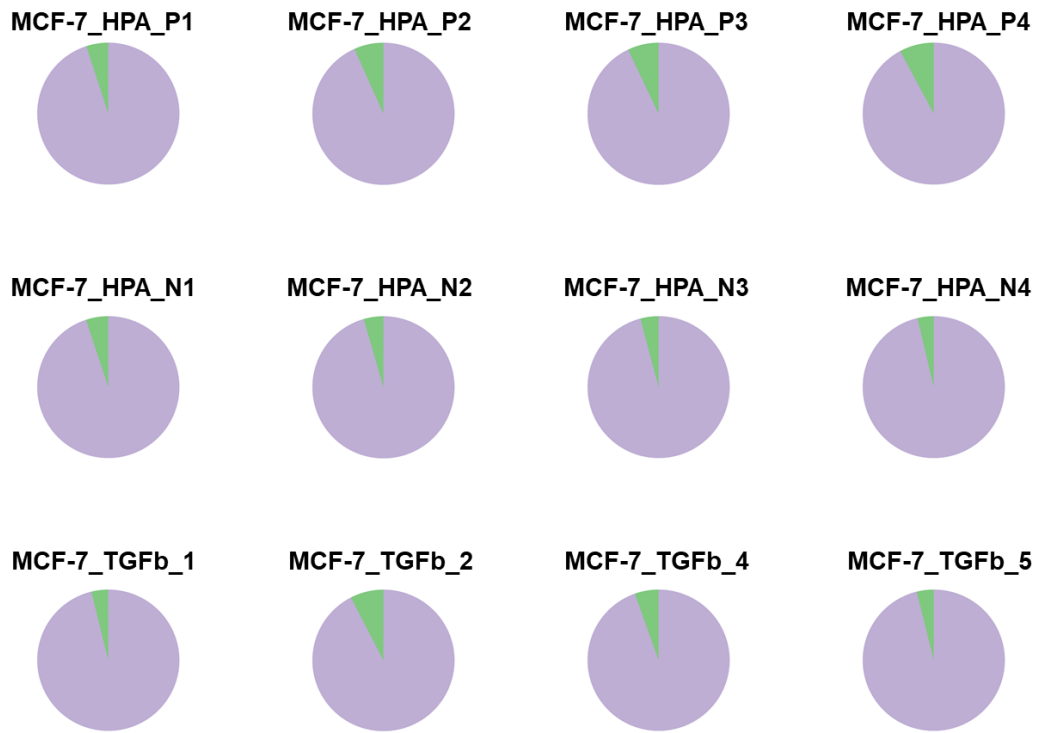


Supplementary Figure 21 Novogene RNA sequencing error rate distribution for MCF-7 *Helix pomatia* agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor-β1 (TGF-β1) treated, T-47D HPA-positive and T-47D HPA-negative samples. The x-axis is the base position along each sequencing read and the y-axis shows the base error rate.

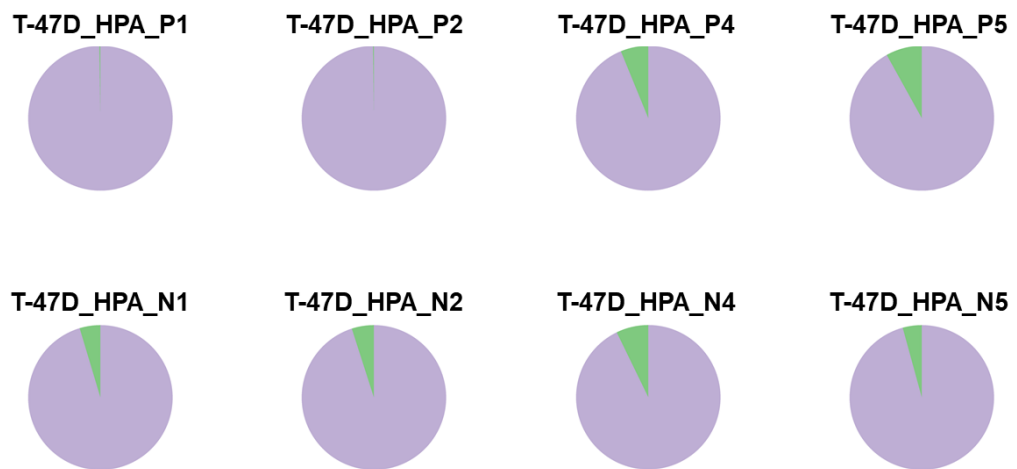


Supplementary Figure 22 Novogene RNA sequencing GC content distribution for MCF-7 *Helix pomatia* agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor- β 1 (TGF- β 1) treated, T-47D HPA-positive and T-47D HPA-negative samples. The x-axis shows each base position within a read; the y-axis shows the percentage of each base with each base represented by a different colour.

MCF-7 Sequencing Data Filtering



T-47D Sequencing Data Filtering



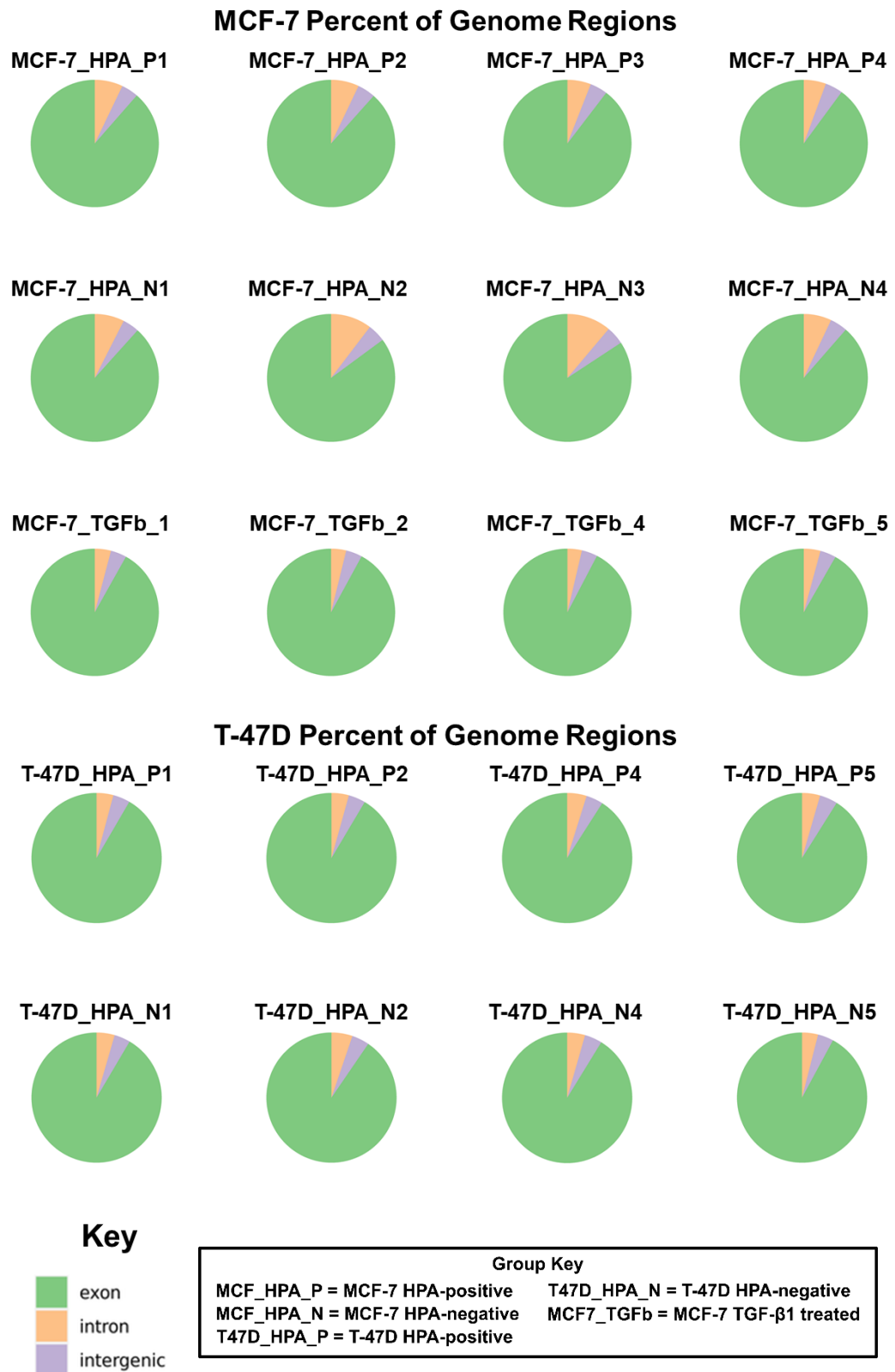
Key



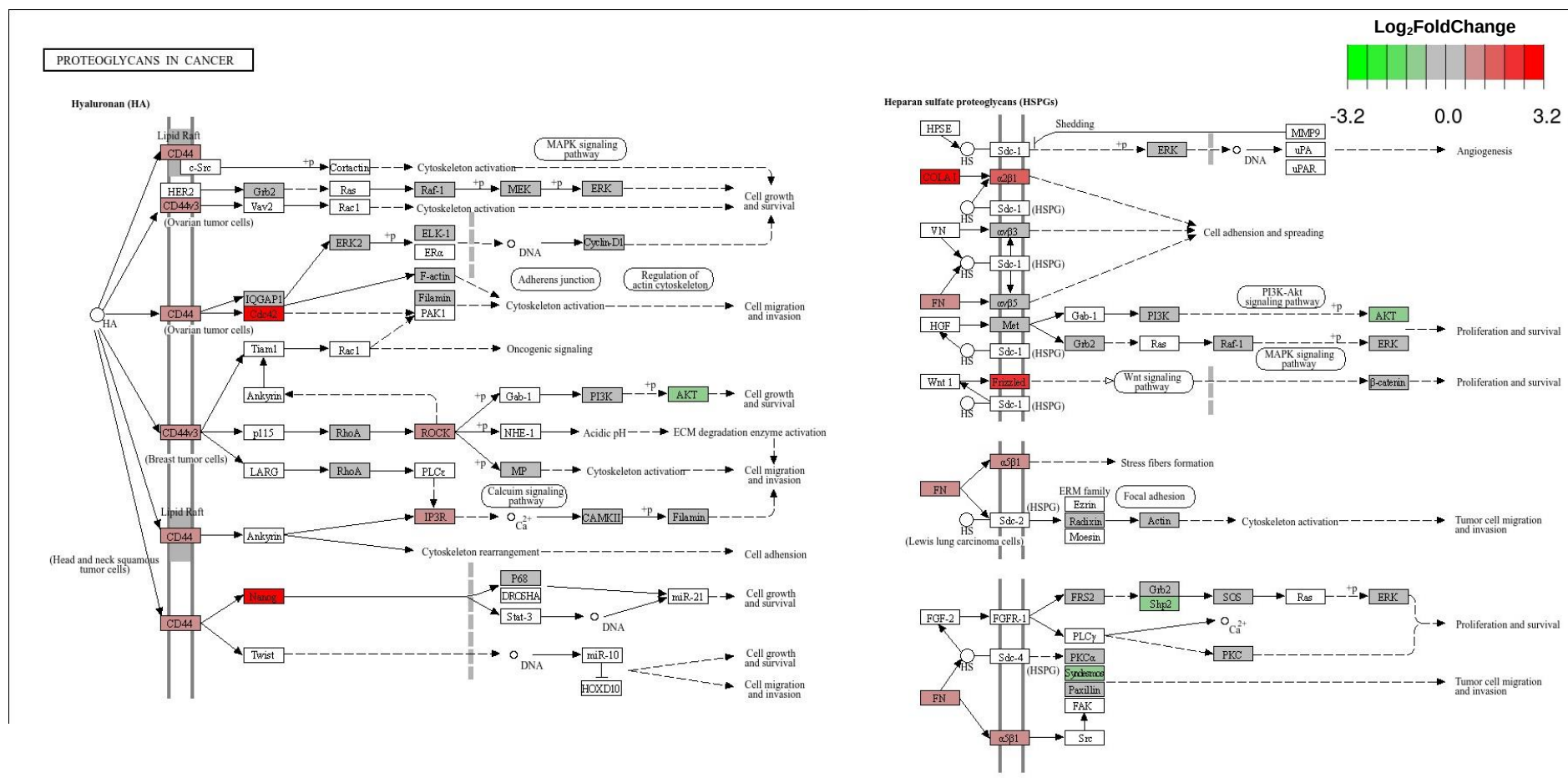
Group Key

MCF_HPA_P = MCF-7 HPA-positive T47D_HPA_N = T-47D HPA-negative
MCF_HPA_N = MCF-7 HPA-negative MCF7_TGFB = MCF-7 TGF-β1 treated
T47D_HPA_P = T-47D HPA-positive

Supplementary Figure 23 Novogene RNA sequencing sample data filtering for MCF-7 *Helix pomatia* agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor-β1 (TGF-β1) treated, T-47D HPA-positive and T-47D HPA-negative samples. Proportions of different colours represent proportions of different components; clean reads, containing N, low quality and adaptor related.



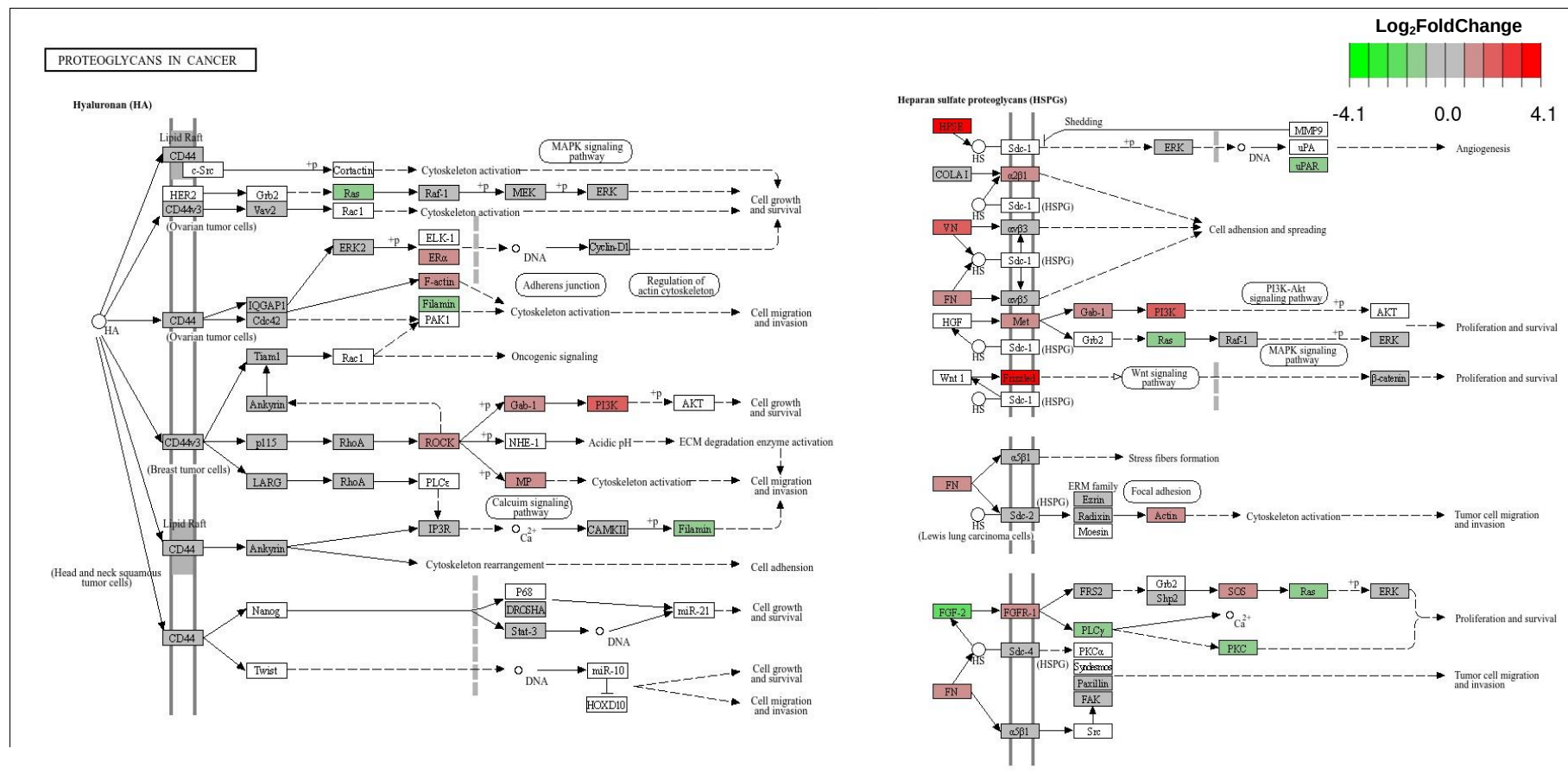
Supplementary Figure 24 Novogene RNA sequencing in the genomic region for MCF-7 *Helix pomatia* agglutinin (HPA)-positive, MCF-7 HPA-negative, MCF-7 transforming growth factor-β1 (TGF-β1) treated, T-47D HPA-positive and T-47D HPA-negative samples. Proportions of different colours represent proportions of different regions; exon, intron and intergenic.



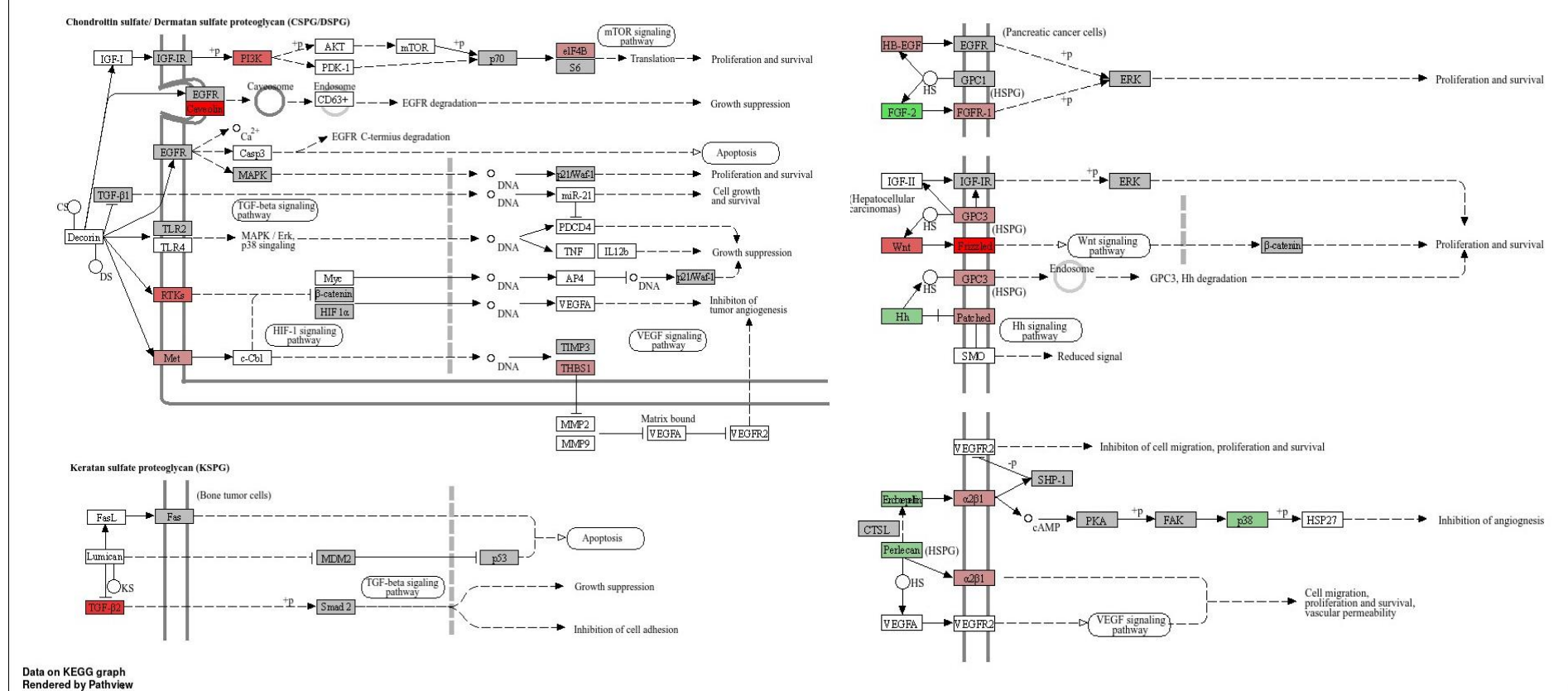
Supplementary Figure 25 Proteoglycans in cancer KEGG (hsa05205) identified 64/2,043 differentially expressed genes (DEGs) ($p < 0.025$) in MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells. Four biological replicates were conducted (1/2).

Supplementary Table 4 MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells proteoglycans in cancer KEGG (hsa05205) differentially expressed genes (DEGs). Regulation direction is HPA-negative compared to HPA-positive; only fold-changes ≥ 1.5 are displayed (64/2,043 genes, $p < 0.025$). Four biological replicates were conducted.

MCF-7 HPA-negative proteoglycans in cancer KEGG DEGs			
Fold-change	Regulation direction	pvalue	Gene name
1.60	Up	$p < 0.025$	CD44
2.83	Up	$p < 0.025$	CDC42
5.94	Up	$p < 0.025$	COL1A2
2.20	Up	$p < 0.025$	EGFR
8.94	Up	$p < 0.025$	ENSG00000229601 (CDC42 Homolog)
2.17	Down	$p < 0.025$	ENSG00000273443 (PTN11)
1.62	Up	$p < 0.025$	ERBB4
1.87	Up	$p < 0.025$	FN1
1.50	Up	$p < 0.025$	HSPG2
9.07	Up	$p < 0.05$	IHH
1.60	Up	$p < 0.025$	ITGA2
1.53	Up	$p < 0.025$	ITGAV
1.69	Up	$p < 0.025$	ITGB1P1
1.61	Up	$p < 0.025$	ITPR1
7.49	Up	$p < 0.025$	NANOGP5
1.63	Down	$p < 0.025$	NUDT16L1
1.52	Down	$p < 0.025$	PPP1CA
1.99	Up	$p < 0.025$	ROCK1P1
1.64	Up	$p < 0.025$	SHH
1.68	Up	$p < 0.025$	THBS1



Supplementary Figure 26 Proteoglycans in cancer KEGG (hsa05205) identified 119/3,589 differentially expressed genes (DEGs) ($p < 0.025$) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 *Helix pomatia* agglutinin (HPA)-positive cells. Four biological replicates were conducted (1/2).



Supplementary Figure 26 (Continued 2/2).

Supplementary Table 5 MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells proteoglycans in cancer KEGG (hsa05205) differentially expressed genes (DEGs). Regulation direction is TGF- β 1 treated compared to *Helix pomatia* agglutinin (HPA)-positive; only fold-changes ≥ 1.5 are displayed (119/3,589 genes, $p < 0.025$). Four biological replicates were conducted.

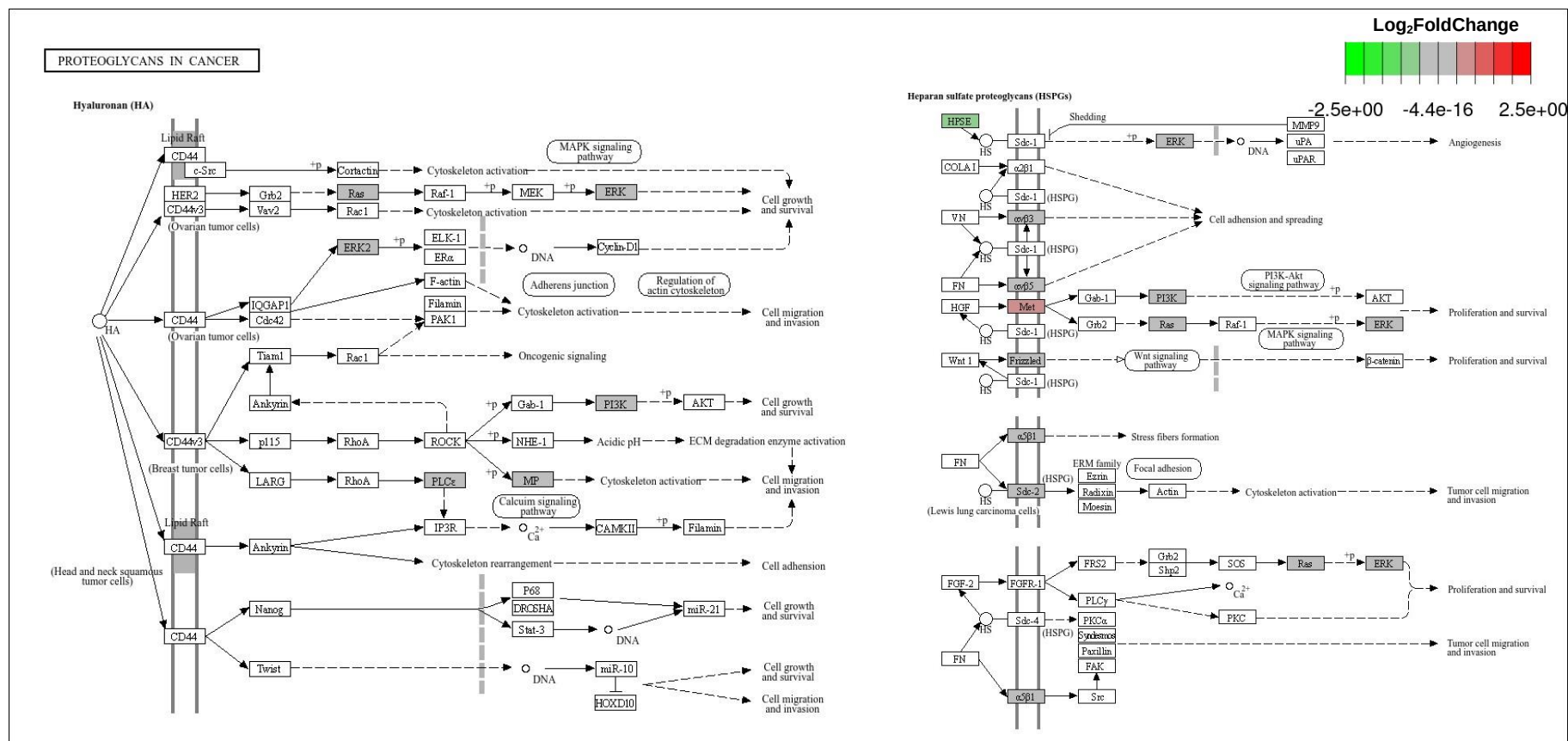
MCF-7 TGF- β 1 treated proteoglycans in cancer KEGG DEGs			
Fold-change	Regulation direction	pvalue	Gene name
6.42	Down	$p < 0.025$	ACTG1P17
1.62	Up	$p < 0.025$	ANK3
1.50	Up	$p < 0.025$	ARHGEF12
6.45	Up	$p < 0.025$	CAV1
2.42	Up	$p < 0.025$	CAV2
1.70	Down	$p < 0.025$	CDC42
5.99	Down	$p < 0.05$	CDC42P4
1.80	Up	$p < 0.025$	EIF4BP7
1.60	Down	$p < 0.025$	ENSG00000217801
1.59	Up	$p < 0.025$	ERBB3
2.48	Up	$p < 0.025$	ERBB4
2.19	Up	$p < 0.025$	ESR1
5.11	Down	$p < 0.025$	FGF2
1.81	Up	$p < 0.025$	FGFR1
2.43	Down	$p < 0.025$	FLNB-AS1
2.44	Up	$p < 0.025$	FN1
1.64	Up	$p < 0.025$	FRS2
1.74	Up	$p < 0.025$	FZD1
2.14	Up	$p < 0.025$	FZD2
1.59	Up	$p < 0.025$	FZD4
1.56	Up	$p < 0.025$	FZD5
1.82	Up	$p < 0.025$	FZD7
2.25	Down	$p < 0.025$	FZD9
1.92	Up	$p < 0.025$	GAB1
2.56	Up	$p < 0.025$	GPC3
1.79	Up	$p < 0.025$	HBEGF
3.09	Down	$p < 0.025$	HIF1A-AS3
16.77	Up	$p < 0.05$	HPSE2
6.11	Down	$p < 0.025$	HRAS
2.41	Down	$p < 0.025$	HSPG2

Supplementary Table 5 (Continued).

1.53	Up	p<0.025	IQGAP1
2.06	Up	p<0.025	ITGA2
2.43	Down	p<0.025	MAPK13
1.89	Up	p<0.025	MET
1.67	Up	p<0.05	novel.541 (ACTG1)
1.61	Up	p<0.025	PIK3R1
1.80	Down	p<0.025	PLAUR
1.92	Down	p<0.025	PLCG2
1.80	Down	p<0.025	PRKCG
2.22	Up	p<0.025	PTCH1
1.66	Up	p<0.025	ROCK2
3.04	Down	p<0.025	SHH
7.53	Up	p<0.025	TGFB2
2.25	Up	p<0.025	THBS1
1.69	Up	p<0.025	VAV2
4.93	Up	p<0.025	VTN
1.79	Up	p<0.025	WNT2B
1.64	Down	p<0.025	WNT4
1.68	Up	p<0.025	WNT6
1.65	Up	p<0.025	WNT9A

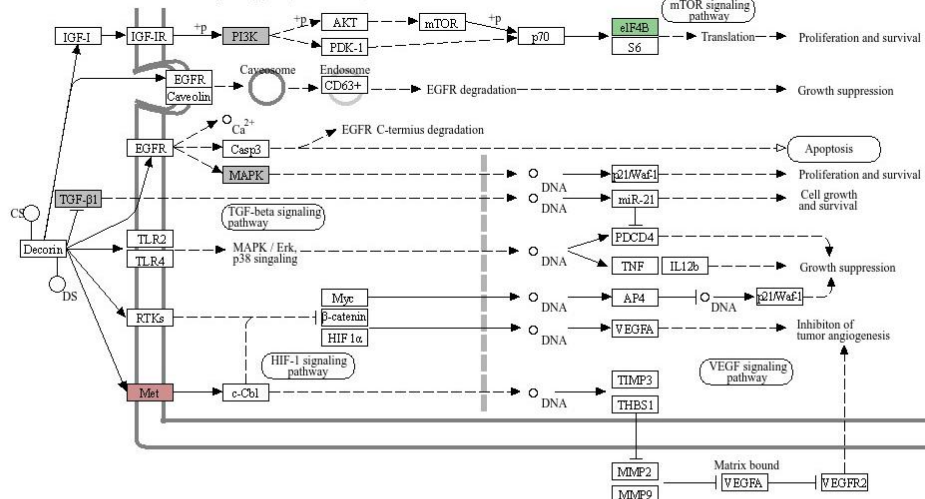
Supplementary Table 6 T-47D *Helix pomatia* agglutinin (HPA)-negative cells proteoglycans in cancer KEGG (hsa05205) differentially expressed genes (DEGs). Regulation direction is HPA-negative compared to HPA-positive; only fold-changes ≥ 1.5 are displayed (19/577 genes, p=0.07). Four biological replicates conducted.

T-47D HPA-negative proteoglycans in cancer KEGG DEGs			
Fold-change	Regulation direction	pvalue	Gene name
1.71	Down	p<0.025	HPSE2
5.00	Up	p<0.025	LUM
5.70	Down	p<0.025	novel.488 (cAMP-dependent protein kinase catalytic subunit alpha)
1.79	Up	p<0.025	MET

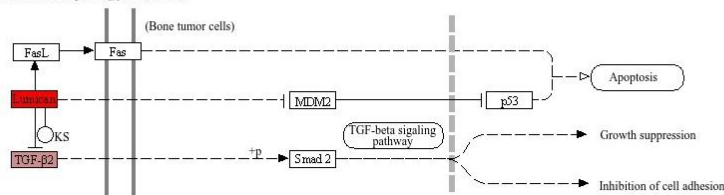


Supplementary Figure 27 Proteoglycans in cancer KEGG (hsa05205) identified 19/577 differentially expressed genes (DEGs) ($p=0.07$) in T-47D *Helix pomatia* agglutinin (HPA)-negative cells compared to T-47D HPA-positive cells. Four biological replicates were conducted (1/2).

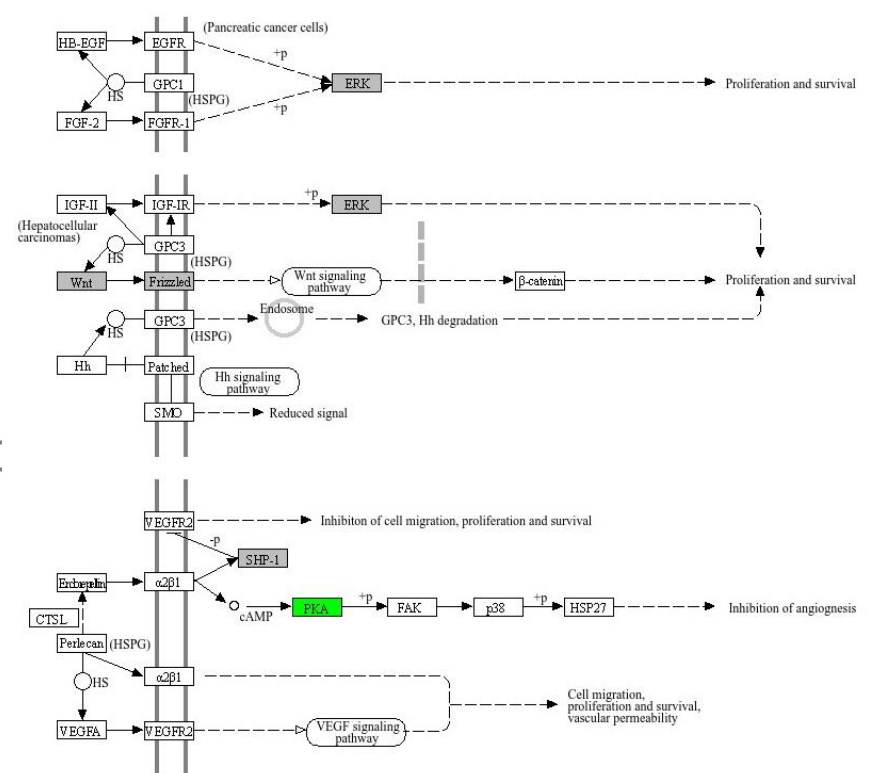
Chondroitin sulfate/ Dermatan sulfate proteoglycan (CSPG/DSPG)



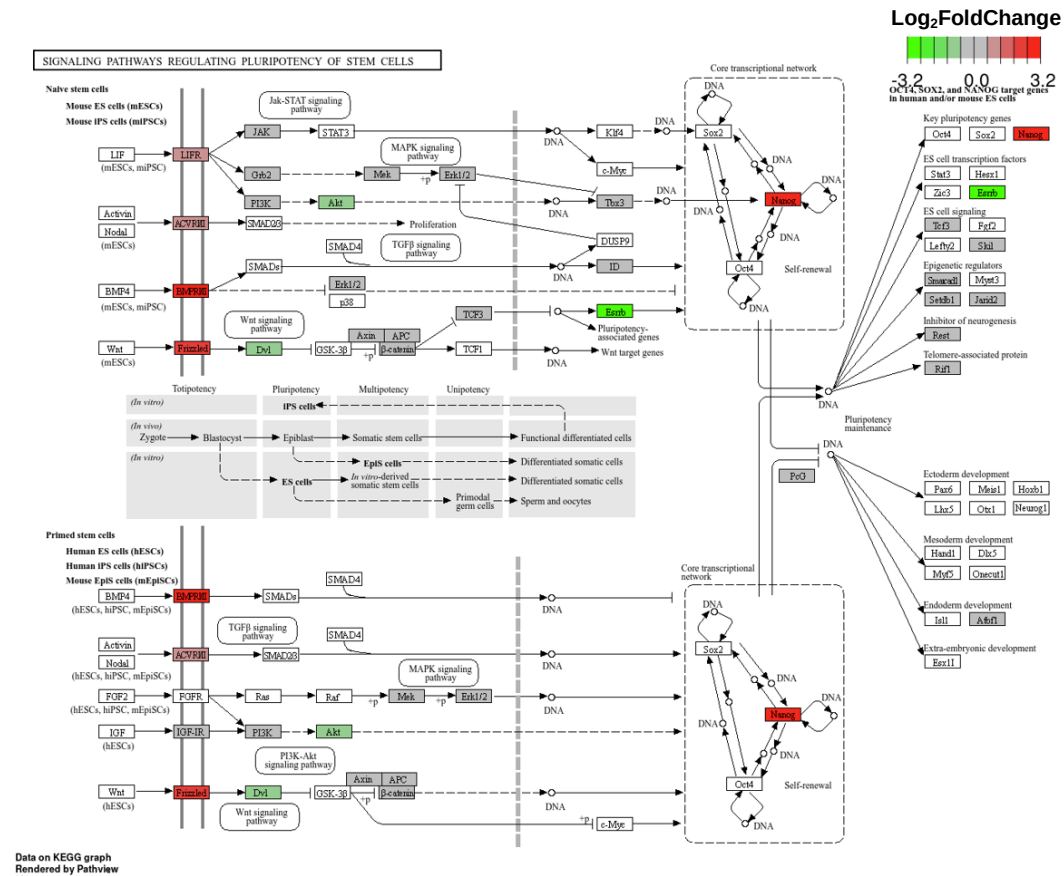
Keratan sulfate proteoglycan (KSPG)



Data on KEGG graph
Rendered by Pathview



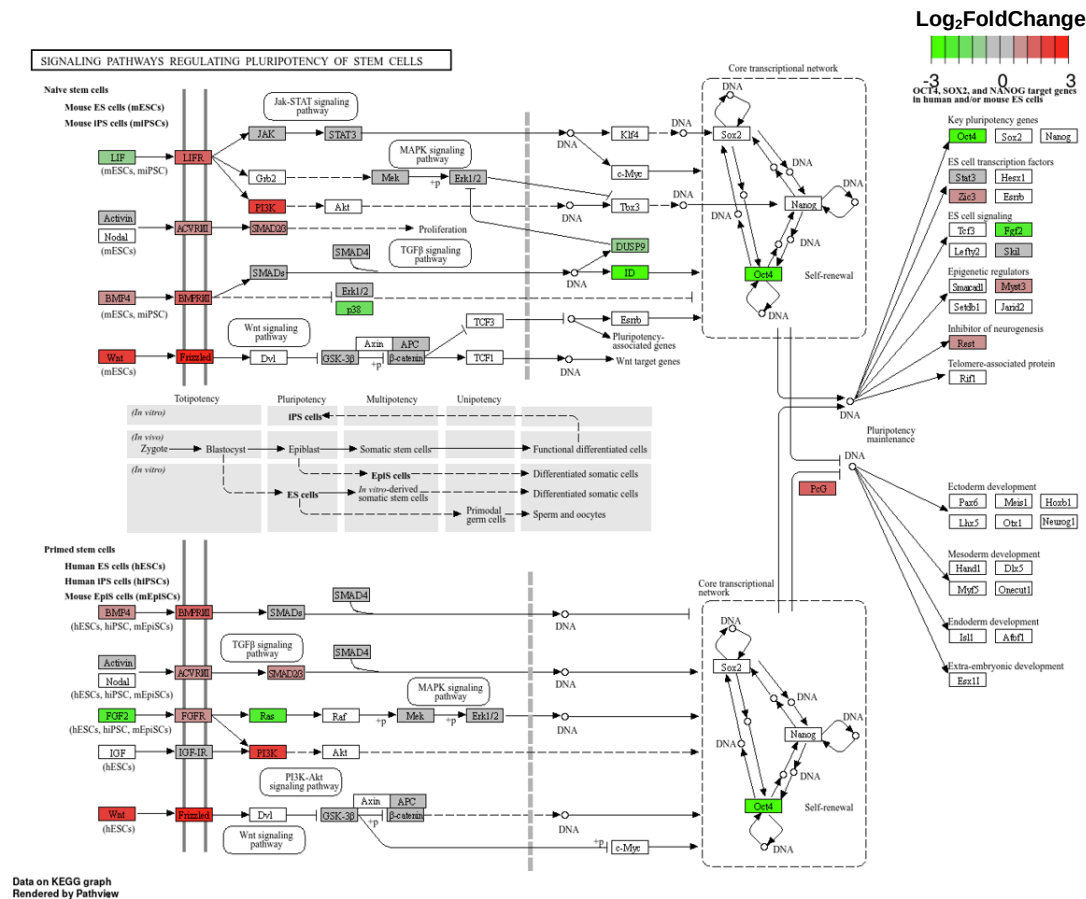
Supplementary Figure 27 (Continued 2/2).



Supplementary Figure 28 Signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) identified 39/2,043 differentially expressed genes (DEGs) ($p < 0.05$) in MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells compared to MCF-7 HPA-positive cells. Four biological replicates were conducted.

Supplementary Table 7 MCF-7 *Helix pomatia* agglutinin (HPA)-negative cells signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) differentially expressed genes (DEGs). Regulation direction is HPA-negative compared to HPA-positive; only fold-changes ≥ 1.5 are displayed (39/2,043 genes, $p < 0.05$). Four biological replicates were conducted.

MCF-7 HPA-negative signalling pathways regulating pluripotency of stem cells			
KEGG DEGs			
Fold-change	Regulation direction	pvalue	Gene name
1.59	Up	$p < 0.05$	AXIN2
9.11	Up	$p < 0.05$	BMPRIAP1
6.55	Down	$p < 0.05$	ESRRB
7.49	Up	$p < 0.05$	NANOGP5



Supplementary Figure 29 Signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) identified 62/3,589 differentially expressed genes (DEGs) ($p < 0.05$) in MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells compared to MCF-7 *Helix pomatia* agglutinin (HPA)-positive cells. Four biological replicates were conducted.

Supplementary Table 8 MCF-7 transforming growth factor- β 1 (TGF- β 1) treated cells signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) differentially expressed genes (DEGs). Regulation direction is TGF- β 1 treated compared to *Helix pomatia* agglutinin (HPA)-positive cells; only fold-changes ≥ 1.5 are displayed (62/3,589 genes, $p < 0.05$). Four biological replicates were conducted.

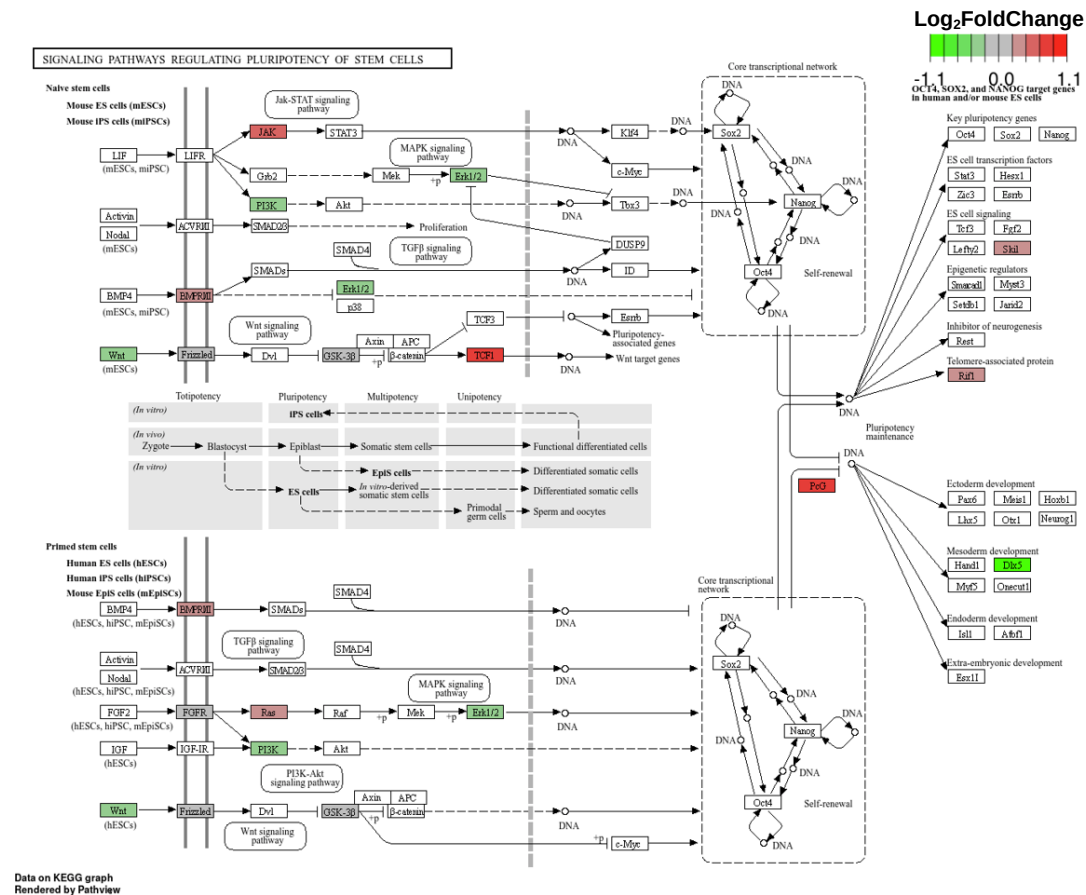
MCF-7 TGF- β 1 treated signalling pathways regulating pluripotency of stem cells			
KEGG DEGs			
Fold-change	Regulation direction	pvalue	Gene name
1.90	Up	$p < 0.025$	BMP4
1.55	Up	$p < 0.025$	BMPR2
2.32	Down	$p < 0.025$	DUSP9
5.11	Down	$p < 0.025$	FGF2
1.81	Up	$p < 0.025$	FGFR1
1.61	Down	$p < 0.025$	FGFR3
1.71	Up	$p < 0.025$	FZD1
2.14	Up	$p < 0.025$	FZD2
1.59	Up	$p < 0.025$	FZD4
1.56	Up	$p < 0.025$	FZD5
1.82	Up	$p < 0.025$	FZD7
2.25	Down	$p < 0.025$	FZD9
6.11	Down	$p < 0.025$	HRAS
3.55	Down	$p < 0.025$	ID1
1.81	Down	$p < 0.025$	ID2
1.50	Up	$p < 0.025$	IL6ST
2.06	Up	$p < 0.025$	INHBA
4.36	Down	$p < 0.025$	INHBE
1.65	Up	$p < 0.025$	KAT6A
1.99	Down	$p < 0.025$	LIF
1.75	Up	$p < 0.025$	LIFR
2.43	Down	$p < 0.025$	MAPK13
1.61	Up	$p < 0.025$	PIK3R1
8.25	Down	$p < 0.025$	POU5F1P3
1.61	Up	$p < 0.025$	SMAD3
1.79	Up	$p < 0.025$	WNT2B
1.64	Down	$p < 0.025$	WNT4
1.68	Up	$p < 0.025$	WNT6
1.65	Up	$p < 0.025$	WNT9A

Supplementary Table 8 (Continued).

1.93	Up	p<0.05	ZIC3
------	----	--------	------

Supplementary Table 9 T-47D *Helix pomatia* agglutinin (HPA)-negative cells signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) differentially expressed genes (DEGs). Regulation direction is HPA-negative compared to HPA-positive; only fold-changes ≥ 1.5 are displayed (17/577 genes, p<0.025). Four biological replicates were conducted.

T-47D HPA-negative signalling pathways regulating pluripotency of stem cells KEGG DEGs			
Fold-change	Regulation direction	pvalue	Gene name
2.14	Down	p<0.05	DLX5
1.76	Up	p<0.05	TCF7



Supplementary Figure 30 Signalling pathways regulating pluripotency of stem cells KEGG (hsa04550) identified 17/577 differentially expressed genes (DEGs) ($p < 0.025$) in T-47D *Helix pomatia* agglutinin (HPA)-negative cells compared to T-47D HPA-positive cells. Four biological replicates were conducted.