

# **Transcriptomic Profiling and Molecular Characterisation of Dormant and Proliferating Breast Cancer Cells in the Brain**

**Ashley James Sunderland**

Submitted in accordance with the requirements for the degree of  
**Doctor of Philosophy**

The University of Leeds  
School of Medicine  
January 2020

## **Intellectual Property and Publication**

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Ashley James Sunderland to be identified as Author of this work has been asserted by Ashley James Sunderland in accordance with the Copyright, Designs and Patents Act 1988.

## Acknowledgments

First and foremost, I would like to thank Dr Mihaela Lorgler for being an incredible main supervisor. She has pushed me beyond what I ever thought was possible and sustained morale even during the most difficult periods. As a result of her mentorship, the highlight of my PhD was being selected for an oral presentation to demonstrate my work in 2019, at an international brain metastases conference in Vienna, Austria. Incredibly, I won the prize for Best Selected Talk. Having reached the end of my PhD, I have since started my academic career as a Research Fellow in Molecular Biology. Thank you, Mihaela.

My appreciation also extends to my secondary supervisors: Prof Valerie Speirs, Dr Tereza Andreou and Dr Alastair Droop. I am very grateful for the close mentoring and support they have offered me. My PhD and my time in Leeds have been significantly enriched by their efforts. In addition to the amazing people I have had the privilege to work with at St James', Wellcome Trust Brenner Building, I wish to acknowledge Jenny Williams for everything she has helped with over the years (and putting up with me!), Christopher Fife for being a top-notch Australian and friend, and Yolanda Kartika for bringing a unique taste of Indonesia and an upbeat attitude to lab life. I would also like to highlight that this research and the opportunities afforded to me during my PhD journey would not have been possible without the Leeds Anniversary Research Scholarship. Importantly, completing my PhD in God's Own Country, Yorkshire, has made this Yorkshireman very proud.

During my PhD, however, we sadly lost my auntie, Wendy, to pancreatic cancer. This only makes me prouder to have been involved in cutting edge cancer research during my time in Leeds. I hope to continue to be a part of the cancer research community in my future endeavours. Finally, I would like to dedicate this thesis to my Mum (Diane), my sister (Courtney), my Dad (Paul), my Nanna (Shirley), and of course my friends. These past few years have been quite a rollercoaster! Their unrelenting support has kept me going throughout.

## Abstract

In metastatic breast cancer, brain metastases (BrM) incidence is 15-30% and effective therapies for BrM represent an unmet clinical need. Cancer dormancy describes the extended period of time in which disseminated tumour cells stop dividing and enter a clinically undetectable, reversible state of quiescence. Under more favourable conditions, the cancer cells may be able to exit dormancy, leading to metastatic progression. It was reasoned that a deeper understanding of dormancy mechanisms may reveal novel therapeutic targets for metastatic breast cancer.

I utilised a xenograft model of BrM and isolated dormant and proliferating breast adenocarcinoma (MDA-MB-231) cells from mouse brains. We performed mRNAseq and differential gene expression analysis to identify differences in gene expression profiles.

GO and KEGG enrichment analysis comparing dormant and proliferating cancer cells revealed an enrichment of extracellular matrix constituents and reduced metabolic processes. This was accompanied by an enrichment of terms pertaining to HIPPO, Wnt and TGF- $\beta$  signalling pathways, amongst other pathways. Inhibition of aerobic glycolysis *in vitro* induced reversible G2/M cell cycle arrest in MDA-MB-231 cancer cells, concurrent with phosphorylation and cytoplasmic retention of a core HIPPO transcriptional regulator, YAP. The most significantly upregulated gene identified in dormancy was the small leucine-rich proteoglycan, biglycan. In primary breast tumours, high biglycan expression was associated with increased occurrence of metastasis and relapse. In a meta-analysis of breast cancer patient gene expression data, biglycan expression was reduced in metastases compared to that of the primary tumour. Furthermore, overexpression of biglycan *in vitro* induced growth arrest of MDA-MB-231 cells. Immunofluorescence imaging of brain tumour xenografts revealed that biglycan-expressing dormant cells as well as proliferating cancer cells *in vivo* localise along the brain parenchymal vasculature, implicating the perivascular niche in the induction of dormancy.

The results are expected to offer a novel insight into the induction of dormancy and suggest targetable pathways for the inhibition of cancer relapse.

Keywords: *Breast Cancer, Dormancy, Brain Metastases, mRNAseq, HIPPO, Wnt, TGF- $\beta$ , Aerobic Glycolysis, YAP, Biglycan, Perivascular Niche*

## Table of Contents

	Page
<b>Intellectual Property and Publication</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Table of Contents</b>	<b>vi</b>
<b>List of Tables</b>	<b>xii</b>
<b>List of Figures</b>	<b>xiv</b>
<b>List of Abbreviations</b>	<b>xix</b>
<b>Chapter 1    Introduction</b>	<b>1</b>
1.1    Breast Cancer	1
1.2    Breast Cancer Brain Metastases	2
1.3    Introduction to Cancer Dormancy	6
1.3.1    Tumour Mass Dormancy	6
1.3.2    Cellular Dormancy	7
1.3.3    Cancer Cell Fate Following Dissemination to the Brain	9
1.4    The Cell Cycle, Cancer and Dormancy	10
1.5    Major Factors and Pathways Involved in Cancer Cell Dormancy	14
1.5.1    The Extracellular Matrix	14
1.5.1.1    Extracellular Matrix Stiffness	16
1.5.1.2    Biglycan	17
1.5.2    The Perivascular Niche	19
1.5.3    Hypoxia	22
1.5.4    Wnt Signalling Cascades	24
1.5.5    PI3K-Akt-mTOR Pathway	26
1.5.6    The Transforming Growth Factor- $\beta$ (TGF- $\beta$ ) Superfamily	28
1.5.7    Mitogen-Activated Protein Kinase Pathways	29
1.5.8    HIPPO Signalling Pathway	31
1.5.9    Cancer Cell Respiration	33

1.6	Exploitation of Dormancy Mechanisms – The Future of Cancer Therapy?	34
1.7	Hypothesis	37
1.7.1	Aims and Objectives	37
<b>Chapter 2</b>	<b>Materials and Methods</b>	<b>38</b>
2.1	Mammalian Cell Culture	38
2.1.1	2-Deoxy-D-Glucose	40
2.1.2	Recombinant Human Biglycan	41
2.2	Cell Membrane Staining	41
2.2.1	Staining of MDA-MB-231 cells with CellVue® Claret Membrane Label	41
2.2.2	Staining of MDA-MB-231 cells with Dil Long-Term Tracer	42
2.3	<i>In Vivo</i> Experiments	42
2.3.1	Preparation of Cancer Cells for Surgery	42
2.3.2	Generation of Mouse Brain Tumour Xenografts	43
2.3.3	Terminal Perfusion and Mouse Brain Isolation	43
2.3.4	Isolation of Cancer Cells from Mouse Brain	44
2.4	Flow Cytometry	45
2.4.1	Flow Cytometry Analysis	45
2.4.2	Fluorescence Activated Cell Sorting (FACS) – Isolation of Dormant and Proliferating MDA-MB-231 Cells from Mouse Brain Tumour Xenografts	46
2.4.3	Cell Cycle Analysis	47
2.5	RNA Isolation and Purification	48
2.5.1	RNA Isolation from Cells Isolated by FACS	48
2.5.2	RNA Isolation from Adherent Cells	49
2.6	cDNA Synthesis, Library Construction and mRNAseq	49
2.6.1	mRNAseq Data Processing and Analysis	49
2.6.2	Alignment to Reference Genomes	50
2.6.3	Read Counting	50
2.6.4	Differential Gene Expression Analysis	50

2.6.5	Calculation of Allele Frequencies	52
2.6.6	GO and KEGG Enrichment Analysis	53
2.6.7	Transcription Factor, Protein Network and Cell Cycle Analyses	53
2.6.8	Analysis of Publicly Available Gene Expression Datasets	53
2.7	Immunofluorescence	54
2.8	Quantitative Polymerase Chain Reaction (qPCR)	57
2.9	Protein Extraction and Quantification	59
2.10	Patient-Derived Tumour Xenografts	61
2.11	Western Blot	61
2.12	Plasmid Propagation	64
2.13	Generation of Lentivirus Stocks	65
2.14	Lentiviral Transduction	67
2.15	Statistical Analysis	67
<b>Chapter 3</b>	<b>Isolation and Transcriptomic Profiling of Dormant and Proliferating Breast Adenocarcinoma Cells</b>	<b>68</b>
3.1	Introduction	68
3.1.1	Hypothesis	69
3.1.1.1	Aims and Objectives	69
3.2	Tracking of Dormant and Proliferating MDA-MB-231 Cell Proliferation using a Lipophilic Cell Membrane Dye	70
3.3	Generation of Mouse Brain Tumour Xenografts	72
3.3.1	Sorting of Dormant and Proliferating MDA-MB-231 Cells	75
3.4	Generation of mRNAseq Sequencing Library from a Small Starting Input of RNA	76
3.4.1	Processing Stream I: Pre-Processing of mRNAseq Data and Quality Control (Individual Genome Alignment)	80
3.4.2	Processing Stream II: Pre-Processing of mRNAseq	82

	Data and Quality Control (Concatenated Genome Alignment)	
3.5	Analysis of Dormant and Proliferating MDA-MB-231 Cell Gene Sets	88
3.6	Analysis of Differentially Expressed Genes	91
3.6.1	Transcription Factor Analysis	94
3.6.2	Gene Ontology and Kyoto Encyclopedia of Genes And Genomes Enrichment Analysis	97
3.7	Validation of mRNAseq Data	105
<b>3.8</b>	<b>Discussion</b>	<b>108</b>
3.8.1	Identification and Isolation of Dormant and Proliferating Breast Adenocarcinoma Cells	108
3.8.2	mRNAseq and Differential Gene Expression Analysis	109
3.8.3	Comparison to Known Dormancy Mechanisms	110
3.8.4	Enriched Molecular Pathways in Dormant Breast Cancer Cells <i>in vivo</i>	112
3.8.4	Limitations	114
<b>3.8.6</b>	<b>Summary</b>	<b>114</b>
<b>Chapter 4</b>	<b>Characterisation of Biglycan as a Potential Mediator and/or Marker of Cellular Dormancy</b>	<b>116</b>
4.1	Introduction	116
4.1.1	Hypothesis	117
4.1.1.1	Aims and Objectives	117
4.2	Determining the Localisation of Dormant MDA-MB-231 Cells In Mouse Brain Tumour Xenografts	119
4.3	Analysis of Biglycan Overexpression in MDA-MB-231 Cell Line	123
4.4	Doxycycline-Inducible Model of Biglycan Overexpression	130
4.5	Determining the Anti-Proliferative Properties of Soluble Biglycan in MDA-MB-231 Cells	132
4.6	Analysis of Biglycan Expression in Human Primary Breast	133

	Cancer and Brain Metastases	
4.7	Analysis of Biglycan Protein Interaction Network	139
<b>4.8</b>	<b>Discussion</b>	<b>140</b>
4.8.1	Localisation of Dormant Breast Cancer Cells to Brain Parenchymal Vascular Endothelium	140
4.8.2	Regulation of Biglycan Expression in Dormancy	141
4.8.3	A Dual Role of Biglycan?	142
4.8.4	Biglycan Mechanisms of Action	142
<b>4.8.7</b>	<b>Summary</b>	<b>144</b>
<b>Chapter 5</b>	<b>Characterisation of Aerobic Glycolysis in the Induction and/or Maintenance of Cellular Dormancy</b>	<b>145</b>
5.1	Introduction	145
5.1.1	Hypothesis	146
5.1.2	Aims and Objectives	146
5.2	Inhibition of Aerobic Glycolysis Induces Reversible Cancer Cell Growth Arrest <i>In Vitro</i> , Mimicking Cancer Cell Dormancy	147
5.3	2DG-Induced Growth Inhibition does not Influence Glycolysis Enzymes or Biglycan Expression <i>In Vitro</i>	155
5.4	Inhibition of Aerobic Glycolysis Induces Reversible G2/M Phase Cell Cycle Arrest	157
5.5	Inhibition of Aerobic Glycolysis may Induce Dormancy in Disseminated Breast Cancer Cells by Modulating YAP Activity	159
5.6	Inhibition of Aerobic Glycolysis does not Inhibit nor Promote Canonical Wnt Signalling in MDA-MB-231 Cells <i>In Vitro</i>	166
5.7	Proposed Schematic of Dormancy Induction Facilitated by a Downregulation of Aerobic Glycolysis	168
<b>5.8</b>	<b>Discussion</b>	<b>171</b>
5.8.1	Relevance of 2DG in the Treatment of Triple Negative Breast Cancer Brain Metastases	171
5.8.2	Inhibited Glycolysis and Cell Survival, Growth Arrest and YAP	172

5.8.3	Biglycan, TLRs, Glycolysis Inhibition and YAP – A Potential Crosstalk	193
5.8.3	Decreased Acidity of the Extracellular Milieu and Growth Arrest	175
<b>5.8.4</b>	<b>Summary</b>	<b>176</b>
<b>Chapter 6</b>	<b>Conclusions and Future Work</b>	<b>177</b>
6.1	Chapter 3	177
6.1.1	Conclusions	177
6.1.2	Critical Analysis	178
6.1.3	Future Work	178
6.2	Chapter 4	179
6.2.1	Conclusions	179
6.2.2	Critical Analysis	180
6.2.3	Future Work	180
6.3	Chapter 5	181
6.3.1	Conclusions	181
6.3.2	Critical Analysis	181
6.3.3	Future Work	182
6.4	Final Conclusion	183
<b>Appendix</b>		<b>184</b>
7.1	List of Suppliers	184
7.2	Supplementary Figures	187
7.2.1	Custom R Scripts	187
7.2.2	Human Biglycan mRNA Sequence in Dormant MDA-MB-231 cells	191
7.3.3	Plasmid Maps	192
<b>Bibliography</b>		<b>199</b>

## List of Tables

### Chapter 1

Table 1.1	A comparison between primary breast cancer and brain metastases
Table 1.2	Reported first site of breast cancer metastases as a percentage of cases Dormancy signature by gene expression
Table 1.3	Association of cyclin-CDK complexes within each point of the eukaryotic cell cycle
Table 1.4	The influence of ECM components on cancer cells
Table 1.5	The effect of ECM stiffness on cancer proliferation
Table 1.6	The effect of ECM stiffness on growth inhibition and dormancy in cancer
Table 1.7	Clinical trials for the targeting of residual disseminated tumour cells in breast cancer

### Chapter 2

Table 2.1	Culture medium constituents for MDA-MB-231 and HEK293 cells
Table 2.2	Constituents of phosphate buffered sucrose solution
Table 2.3	Components of collagenase/hyaluronidase solution for mouse brain dissociation
Table 2.4	Laser and fluorescent channel configurations for flow cytometry (CytoFlex)
Table 2.5	Solution constituents for cell cycle FACS analysis
Table 2.6	Download parameters for human (GRCh38/hg38) and mouse (GRCm38/mm10) reference genomes (NCBI RefSeq), processing stream I
Table 2.7	Components of Walter's antifreeze for storage of PFA-fixed floating frozen sections
Table 2.8	Components of antibody blocking buffer for immunofluorescence
Table 2.9	Summary table of antibodies
Table 2.10	Mastermix I components for cDNA synthesis
Table 2.11	Mastermix II components for cDNA synthesis
Table 2.12	Summary table of Taqman® probes

Table 2.13	Constituents of RIPA buffer
Table 2.14	Constituents of hypotonic buffer for cellular lysis
Table 2.15	Constituents of 6X SDS loading buffer
Table 2.16	Polyacrylamide gel constituents (1 gel, 0.75mm thickness)
Table 2.17	SDS running buffer constituents
Table 2.18	Transfer buffer constituents for Western blot
Table 2.19	Constituents of solution containing vectors for generation of lentiviral stocks via transfection of HEK293T cells
Table 2.20	Buffer constituents for generation of lentiviral stocks.

### Chapter 3

Table 3.1	Dormant cells account for a small fraction of the total cancer cell population within mouse brain tumour xenografts
Table 3.2	Yields of total RNA derived from dormant and proliferating MDA-MB-231 cells <i>in vivo</i>
Table 3.3	Descriptive statistics for mRNAseq alignment data (processing stream I)
Table 3.4	Differentially expressed genes between dormant and proliferating breast cancer cells <i>in vivo</i> in the context of c-Myc and HIF-1A transcription factors
Table 3.5	Comparison to known dormancy signatures by gene expression

### Chapter 5

Table 5.1	Inhibition of aerobic glycolysis using the glycolysis inhibitor, 2DG, does not inhibit the expression of core glycolysis genes
Table 5.2	Inhibition of aerobic glycolysis does not affect the expression of biglycan <i>in vitro</i>
Table 5.3	Inhibition of aerobic glycolysis does not affect the expression of the YAP activator <i>L1CAM</i>
Table 5.4	Inhibition of aerobic glycolysis does not affect the expression of the Wnt obligate genes, <i>AXIN2</i> and <i>NKD1</i> , <i>in vitro</i>
Table 5.5	A comparison of dormant cancer cells <i>in vivo</i> and MDA-MB-231 cells <i>in vitro</i> (after addition of 10mM 2DG)

## List of Figures

### Chapter 1

- Figure 1.1 Potential crosstalk between breast cancer cells disseminated to the brain and local brain cell populations
- Figure 1.2 Cancer progression and dormancy
- Figure 1.3 Tracking of breast cancer cell fate in the brain by magnetic resonance imaging
- Figure 1.4 An overview of some of the critical regulators in cell cycle regulation
- Figure 1.5 Signal transduction pathways activated downstream of extracellular biglycan
- Figure 1.6 Overview of the canonical and non-canonical Wnt signalling pathways
- Figure 1.7 Simple schematic of the PI3K-Akt-mTOR signalling pathway in cancer
- Figure 1.8 The TGF- $\beta$  superfamily: ligands, receptors and SMADs
- Figure 1.9 An overview of the MAPK pathways
- Figure 1.10 HIPPO signalling pathway cascade in humans
- Figure 1.11 Schematic overview of glycolysis

### Chapter 3

- Figure 3.1 Proliferation tracking of MDA-MB-231 cells *in vitro* with the vital membrane dye, CellVue® Claret
- Figure 3.2 *In vivo* model to study breast cancer cell dormancy in the brain
- Figure 3.3 Identification of brain tumour xenografts *in vivo* by immunofluorescence staining
- Figure 3.4 Isolating dormant (231/GFP/CV) and proliferating (231/GFP) cancer cells by fluorescence activated cell sorting
- Figure 3.5 Quality control of RNA isolated from a small number of dormant and proliferating cancer cells
- Figure 3.6 Quality control of cDNA library for mRNAseq
- Figure 3.7 Proportion of human- and mouse-aligned reads in dormant and proliferating MDA-MB-231 cells (processing stream I)

- Figure 3.8** Workflow of mRNAseq data processing – alignment to concatenated human and mouse reference genomes (processing stream I)
- Figure 3.9** Workflow of mRNAseq data processing – alignment to concatenated human and mouse reference genomes (processing stream II)
- Figure 3.10** Per-sequence GC content of mRNAseq reads derived from dormant and proliferating breast cancer cells *in vivo* (processing stream II)
- Figure 3.11** Proportion of human- and mouse-aligned reads in dormant and proliferating MDA-MB-231 cells (processing stream II)
- Figure 3.12** MA-plots of normalized mRNAseq data
- Figure 3.13** Pairwise comparison of known MDA-MB-231 cell SNPs between dormant and proliferating breast adenocarcinoma cells (processing stream I)
- Figure 3.14** Dormant and proliferating MDA-MB-231 cells display distinct gene expression profiles *in vivo*
- Figure 3.15** Top 30 up- and down-regulated protein-coding genes in dormant versus proliferating MDA-MB-231 cells derived from mouse brain tumour xenografts
- Figure 3.16** Prediction of cell cycle phase in dormant breast adenocarcinoma cells *in vivo*
- Figure 3.17** Predicted transcription factors implicated in breast cancer cell dormancy in the brain
- Figure 3.18** KEGG and KEGG module enrichment analysis (processing streams I and II)
- Figure 3.19** KEGG and KEGG module enrichment analysis (processing stream II) with no statistical cutoff on output terms
- Figure 3.20** Potential crosstalk between Wnt, HIPPO and TGF- $\beta$  signalling pathways in dormant MDA-MB-231 cells *in vivo*
- Figure 3.21** Gene Ontology (GO) enrichment analysis (processing stream I)
- Figure 3.22** Gene Ontology (GO) enrichment analysis (processing stream II)
- Figure 3.23** Gene Ontology (GO) enrichment analysis (processing stream II) with no statistical cutoff on output terms

Figure 3.24 Biglycan expression in proliferating and dormant MDA-MB-231 cells *in vivo*, compared to cells cultured *in vitro*

Figure 3.25 Collagen I expression in proliferating and dormant MDA-MB-231 cells *in vivo*, compared to cells cultured *in vitro*

#### Chapter 4

Figure 4.1 The majority of MDA-MB-231 cells which colonise mouse brains following intra-carotid injection are negative for biglycan

Figure 4.2 Biglycan-expressing MDA-MB-231 cells in brain tumour xenografts localize around cerebral vasculature

Figure 4.3 Three-dimensional rendering of biglycan-secreting cancer cells *in vivo*

Figure 4.4 Generation of biglycan-expressing MDA-MB-231 cells

Figure 4.5 Overexpression of human biglycan in MDA-MB-231 and HEK293 cells

Figure 4.6 Overexpression of biglycan in MDA-MB-231 cells at low level does not induce cell cycle arrest under high serum conditions *in vitro*

Figure 4.7 Stronger biglycan expression induces downregulation of *GAPDH* and complete growth inhibition of MDA-MB-231 cells *in vitro*

Figure 4.8 Biglycan is secreted when overexpressed in MDA-MB-232 cells *in vitro*

Figure 4.9 Biglycan secretion is inhibited using a Tet-OFF biglycan expression system in MDA-MB-231 cells *in vitro*

Figure 4.10 Doxycycline-inducible biglycan over-expression induces growth inhibition of MDA-MB-231 cells *in vitro*

Figure 4.11 Recombinant human biglycan induces growth arrest of MDA-MB-231 cells *in vitro*

Figure 4.12 Meta-analysis of biglycan expression in breast cancer metastases compared to primary tumours

Figure 4.13 Analysis of biglycan expression in patient-derived brain tumour xenografts

Figure 4.14 Correlation of biglycan expression in primary breast tumours and relapse-free survival, by subtype

Figure 4.15 Correlation of biglycan expression in primary breast tumours and distant metastasis-free survival, by subtype

Figure 4.16 Human biglycan protein interaction network

**Chapter 5**

- Figure 5.1** Core enzymes in the glycolysis pathway are downregulated in dormant cells *in vivo*
- Figure 5.2** Inhibition of aerobic glycolysis with the glycolysis inhibitor, 2DG, induces reversible growth arrest in MDA-MB-231 cells *in vitro* (I)
- Figure 5.3** Inhibition of aerobic glycolysis with the glycolysis inhibitor, 2DG, induces reversible growth arrest of MDA-MB-231 cells *in vitro* (II)
- Figure 5.4** Inhibition of aerobic glycolysis with 10mM 2DG inhibits growth of breast adenocarcinoma cells whilst retaining normal cell size and morphology
- Figure 5.5** Tracking the growth-inhibitory effects of 2DG on MDA-MB-231 cells by membrane staining
- Figure 5.6** Inhibition of aerobic glycolysis in MDA-MB-231 cells prevents acidification of growth medium *in vitro*
- Figure 5.7** Inhibition of aerobic glycolysis induces reversible G2/M phase arrest in MDA-MB-231 cells *in vitro*
- Figure 5.8** Inhibition of aerobic glycolysis induces phosphorylation of the transcriptional activator, YAP, *in vitro*
- Figure 5.9** Inhibition of aerobic glycolysis induces cytoplasmic retention of the transcriptional activator YAP
- Figure 5.10** Effect of inhibited aerobic glycolysis on the expression of laminin receptor subunits
- Figure 5.11** Inhibition of aerobic glycolysis does not induce nuclear localisation of  $\beta$ -catenin *in vitro*
- Figure 5.12** Schematic of proposed dormancy mechanisms in breast cancer cells

**Appendix**

- Supplementary Figure S1** R script for the concatenation of human and mouse reference genomes prior to generation of human-mouse mRNAseq alignment index, using R package Biostrings
- Supplementary Figure S2** R script for mRNAseq read count generation from BAM files using R package RSubread

<b>Supplementary Figure S3</b>	<b>R script for Cell cycle phase analysis using R package Scran.</b>
<b>Supplementary Figure S4</b>	<b>R script for the normalisation of microarray data (CEL files)</b>
<b>Supplementary Figure S5</b>	<b>Alignment of dormant cell human BGN mRNA sequence to known human BGN sequence</b>
<b>Supplementary Figure S6</b>	<b>Plasmid map for the ubiquitin C promotor-controlled human biglycan lentiviral expression vector, pFUW_hBGN</b>
<b>Supplementary Figure S7</b>	<b>Plasmid map for the ubiquitin C promotor-controlled enhanced green fluorescent protein (eGFP) lentiviral expression vector, pFUGW</b>
<b>Supplementary Figure S8</b>	<b>Plasmid map for the ubiquitin C promotor-controlled lentiviral expression vector, pFUW (empty vector)</b>
<b>Supplementary Figure S9</b>	<b>Plasmid map for the doxycycline-inducible (Tet-ON) enhanced green fluorescent protein (eGFP) lentiviral expression vector pTREAUTO3</b>
<b>Supplementary Figure S10</b>	<b>Plasmid map for the doxycycline-inducible (Tet-ON) human biglycan lentiviral expression vector, pTREAUTO_hBGN_NeoR (pTREA_BGN)</b>
<b>Supplementary Figure S11</b>	<b>Plasmid map for the doxycycline-inducible (Tet-OFF) human biglycan lentiviral expression vector, pLenti_BGN</b>
<b>Supplementary Figure S12</b>	<b>Plasmid map for the doxycycline-inducible (Tet-OFF) lentiviral expression vector, pLenti Puro (empty vector)</b>

## List of Abbreviations

18S	Ribosomal subunit 18S
2,3-DPG	2,3-diphosphoglycerate
28S	Ribosomal subunit 28S
2DG	2-deoxy-D-glucose
2PG	2-phosphoglyceric acid
3PG	3-phosphoglyceric acid
5FU	5-fluorouracil
A35	Berberine of 1,13-cyclization
ActR	Activin receptor
ALDOA	Aldolase
ALK	Anaplastic lymphoma kinase
AMH/MIS	Anti-Müllerian hormone
AMPK	AMP-activated protein kinase
AmpR	Ampicillin resistance
ANGPTL4	Angiopoietin-like 4
ANOVA	Analysis of Variance
APC	Allophycocyanin
APS	Ammonium persulphate
ARC3	Advanced Research Computer 3
ASK	Apoptosis signal-regulating kinase
AT	Ambient temperature
ATF2	Activating transcription factor 2
ATP	Adenosine triphosphate
AURKA	Aurora kinase A
AWERC	Animal Welfare & Ethical Review Committee
AXIN2	Axis inhibition protein 2
BAD	Bcl2-associated death promotor
BAM	Binary Alignment MAP
BBB	Blood-brain barrier
BC	Breast cancer
BCBM	Breast cancer brain metastasis (PDTXs)
BCBrM	Breast cancer brain metastasis
BGN	Biglycan
BHLH	Basic helix-loop-helix
BHLHE41	Basic helix-loop-helix family member E41
BLBC	Basal-like breast cancer
BMDCs	Bone marrow derived immune cells
BMP	Bone morphogenic protein
BRCA1	Breast cancer type 1 susceptibility protein

BRCA2	Breast cancer type 2 susceptibility protein
BrdU	5-bromo-2'-deoxyuridine
BrM	Brain metastases
BSA	Bovine serum albumin
CA	Carbonic anhydrase
CAK	CDK activating kinase
CAMs	Cellular adhesion molecules
CAP	Catabolite activator protein
CCND1	Cyclin D
ccRCC	Clear cell renal cell carcinoma
CD	Cluster of differentiation
CDC42	Cell division protein 42 homolog
CDKs	Cyclin-dependent kinases
cDNA	Complementary DNA
Chk1	Checkpoint kinase 1
chMDA-MB-231	Hypoxia resistant MDA-MB-231
CK1 $\gamma$	Casein kinase 1 $\gamma$
CMV	Cytomegalovirus
COL1A1	Collagen I
COSMIC	Catalogue of Somatic Mutations In Cancer
cPPT	Central polypurine tract
CREB1	CAMP responsive element binding protein 1
CTNNB1	$\beta$ -catenin
CuB	Curcubitacin B
CuE	Cucurbitacin E
CV	CellVue Claret <sup>®</sup>
CXCL	C-X-C motif chemokine ligand
CXCR	C-X-C chemokine receptor
DAAM1	Dvl associated activator of morphogenesis 1
DAPI	4',6-diamidino-2-phenylindole
DCA	Dichloroacetate
DCIS	Ductal carcinoma <i>in situ</i>
ddH <sub>2</sub> O	Double distilled water
DEGs	Differentially expressed genes
Dil	Vybrant Dil
DKK1	Dickkopf-related protein 1
DLG	Discs large
DLK	Dual leucine zipper kinase
DMEM	Dulbecco's Modified Eagle Medium
DMFS	Distant metastasis-free survival

DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT1	DNA methyltransferase 1
DOC	Sodium deoxycholate
DTCs	Disseminated tumour cells
DTT	Dithioereitol
E-CAD	E-cadherin
<i>E.coli</i>	<i>Escherichia coli</i>
EACR	Extracellular acidification rate
ECL	Enhanced chemiluminescence
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic
EdU	5-ethynyl-2'-deoxyuridine
EGA	European Genome-phenome Archive
eGFP	enhanced green fluorescent protein
EGFR	Epidermal growth factor receptor
EMEM	Eagle's Minimum Essential Medium
EMT	Epithelial-mesenchymal transition
ENO	Enolase
EPCAM	Epithelial cell adhesion molecule
ER	Oestrogen receptor
ERK	Extracellular regulated kinase
ESA	Epithelial specific antigen
F-12	Nutrient Mixture F-12
FA	Focal adhesion
FACS	Fluorescence activated cell sorting
FAK	Focal adhesion kinase
FBC	Female breast cancer
FBS	Foetal Bovine Serum
FDR	False Discovery Rate
FGF	Fibroblast growth factor
FGFR1	FGF receptor
FITC	Fluorescein isothiocyanate
FOXM1	Forkhead box M1
Fru-1,6-dip	Fructose-1,6-diphosphate
Fru-6P	Fructose-6-phosphate
FZD	Frizzled
G1	Gap1
G2	Gap2
G3P	Glyceraldehyde-3-phosphate
G6Pase	Glucose-6-phosphatase

GABA	Gamma-aminobutyric acid
Gagg45 $\alpha$	Growth arrest and DNA damage inducible alpha
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GBM	Glioblastome multiforme
GDF1	Growth differentiation factor 1
GEO	Gene Expression Omnibus
GFP	Green fluorescent protein
GFP+	Proliferating cells
GFP+CV+	Dormant cells
Glt <sub>n</sub>	Gallotannin
Glu-6P	Glucose-6-phosphate
GLUT1	Glucose transporter 1
GO	Gene ontology
GPCR	G protein-coupled receptor
GRB2	Growth factor receptor-bound protein 2
GSK3B/GSK3 $\beta$	Glycogen synthase kinase 3 $\beta$
GSK $\beta$	Glycogen synthase kinase $\beta$
GTP	Guanosine triphosphate
hBGN	Human biglycan
HER	Human epidermal growth factor receptor
HIF-1A	Hypoxia-inducible factor 1 alpha
HIF-1B	Hypoxia-inducible factor 1 beta
HIV	Human immunodeficiency virus
HK	Hexokinase
HMB	High mobility group
hnRNPA0	Heterogeneous ribonucleoprotein A0
HR	Hazard ratio
HRP	Horseradish peroxidase
HSPE	Heparanase
IC <sub>50</sub>	Half maximal inhibitory concentration
ID1	Inhibitor of differentiation 1
IF	Immunofluorescence
IGF-1	Insulin-like growth factor 1
IGV	Integrative Genome Viewer
IKK	Inhibitor of nuclear factor kappa-B
IL	Interleukin
IL-23R	Interleukin-23 receptor
ILK	Integrin linked kinase
IND	Investigational new drug
INHBA	Inhibin beta A chain precursor

IRS1	Insulin receptor substrate 1
JAG	Jagged
JAK-STAT	Janus kinase signal transduction and activator of transcription
JNK	c-Jun N-terminal kinase
KEGG	Kyoto encyclopeadia of genes and genomes
KM	Kaplan Meier
L1CAM	Cell adhesion molecule 1
LABC	Luminal A breast cancer
Lac	Lactose
LAP	Leucyl aminopeptidase
LATS	Large tumour suppressor
LBBC	Luminal B breast cancer
LCC	Latency competent cancer
LDHA	Lactate dehydrogenase
LEF	Lymphoid enhancer-binding factor
LGR5	Leucine-rich-repeat-containing G-protein coupled receptor 5
lncRNA	Long noncoding RNA
LOX	Lysyl oxidase
LPA1	Lysophosphatidic acid receptor
LRP	Low-density lipoprotein receptor-related protein 1
LRR2-3	Leucine-rich repeats 2-3 domain
LSCC	Laryngeal squamous cell carcinoma
LTR	Long terminal repeat
M	Mitosis
mA	Milliamps
MAPK	Mitogen-activated protein kinase
MAPKK	MAPK kinase
MAPKKK	MAPK kinase kinase
MARC3	Medical Advanced Research Computer 3
MASK	Ankyrin repeat and HK domain containing
MBC	Male breast cancer
MCF7	Michigan cancer foundation 7
MCM	Minochromosome maintenance
MDA-MB-231/brain	Brain-homing MDA-MB-231
MDM2	Murine double minute
MDSCs	Myeloid derived suppressor cells
MEK	Mitogen-activated protein kinase kinase
Met.	Metastasis
MFs	Molecular functions

MK2	MAP kinase-activated protein kinase 2
MLK	Mixed-lineage protein kinase
MM	Multiple myeloma
MMP	Matrix metalloproteinase
Mos	Moloney sarcoma
MPIO	Micron-sized superparamagnetic iron oxide
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
mRNAseq	Messenger RNA sequencing
MSCs	Mesenchymal stem cells
MST	Macrophage stimulating protein
mTOR	Mammalian target of rapamycin
MYC	c-Myc
Na <sub>3</sub> VO <sub>4</sub>	Sodium orthovanadate
NaCl	Sodium chloride
NCBn1/SLC4A7	Sodium bicarbonate transporter 3
ND	Not determined
NDRG1	N-myc downstream-regulated 1
NEMO	Nuclear factor-kappa B essential modulator gene
NeoR	Neomycin resistance
NF-kB	Nuclear factor kappa-light-chain-enhancer of B cells
NGS	Next Generation Sequencing
NHE	Na <sup>+</sup> -H <sup>+</sup> exchanger
NHE1/SLC9A1	Sodium-hydrogen exchanger 1
NKD1	Naked cuticle homolog 1
NLBC	Normal-like breast cancer
NLK	Nemo-like kinase
NOG	Noggin
NR2F1	Nuclear receptor subfamily 2 group F member 1
NTPs	Nucleoside triphosphate
OCR	Oxygen consumption rate
p-	phosphorylated-
Pa	Elastic modulus ( <i>E</i> )
PA	Plasminogen activator
PBS	Phosphate buffered saline
PCA	Principle Component Analysis
PCM	Pericellular matrix
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDE1A	Phosphodiesterase 1A

PDK1	3-phosphoinositide-dependent protein kinase 1
PDTX	Patient-derived tumour xenograft
PEP	Phosphoenolpyruvate carboxylase
PFA	Paraformaldehyde
PFK	Phosphofructokinase
PGC-1 $\beta$	nuclear peroxisome proliferator-activated receptor gamma cofactor-1- $\beta$
PGI	Phosphoglycerate isomerase
PGK	Phosphoglycerate mutase
PGK1	Phosphoglycerate kinase 1
PHDs	Prolyl-4-hydroxylases
PI	Propidium iodide
PI3K	Phosphoinositide 3-kinase
PIP <sub>2</sub>	Phosphatidylinositol-(4, 5)-bisphosphate
PIP <sub>3</sub>	Phosphatidylinositol-(3, 4, 5)-triphosphate
PKB/Akt	Protein kinase B
PKM2	Pyruvate kinase M2
POLR2A	RNA polymerase II subunit A
polyA	Polyadenylated
PP1	Protein phosphatase 1
PTEN	Phosphatase and tensin homolog
PuroR	Puromycin resistance
PVN	Perivascular niche
Pyr	Pyruvate
qPCR	Quantitative polymerase chain reaction
Rac1	Ras-related C3 botulinum toxin substrate 1
Raf	Rapidly accelerated fibrosarcoma
RANKL	Receptor Activator of Nuclear Factor Kappa-B Ligand
RAPTOR	Regulatory associated protein of TOR
RELA	Transcription factor p65
RFS	Relapse-free survival
RRE	Rev response element
rhBGN	Recombinant human biglycan
Rheb	Ras homolog enriched in brain
RICTOR	Rapamycin-insensitive companion of mammalian target of rapamycin
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
ROCK	Rho-associated protein kinase
ROR	Tyrosine-protein kinase transmembrane receptor
ROR2	Receptor tyrosine kinase-like orphan receptor 2

RSV	Rous sarcoma virus
S	Synthesis
SAPK	Stress activated protein kinase
SBS	St James' Biological Services
SCID	Severe combined immunodeficiency
SCRIB	Scribble
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
Ser	Serine
SFK	Src family kinase
shRNA	Short hairpin RNA
SIAH2	Siah E3 ubiquitin protein ligase 2
siRNA	Small interfering RNA
SLRP	Small leucine-rich proteoglycan
SNP	Single nucleotide polymorphism
SOG	Short gastrulation
SOX	Sry-type HMG box
SP1	Specificity protein 1
STAR	Spliced Transcripts Alignment to a Reference
TAMs	Tumour associated macrophages
TAO	Thousand-and-one amino acid
TAZ	Tafazzin
TCF	T-cell factor
TCGA	The Cancer Genome Atlas
TE	Tris-EDTA
TEAD	TEA domain
TEMED	Tetramethylethylenediamine
Tet	Tetracycline
Tet2	Tet methylcytosine dioxygenase 2
TetO	Tetracycline-ON
TF	Transcription factor
TGF- $\beta$	Transforming growth factor $\beta$
	TIA1 cytotoxic granule associated DNA binding protein like 1
TIAR	
TIE2	Angiopoietin-1 receptor
TLR	Toll-like receptor
TPI	Triphosphate isomerase
Tpl2	Tumour progression locus 2
TRCs	Tumour repopulating cells
TRITC	Tetramethylrhodamine
TSC	Tuberous sclerosis

TSP-1	Thrombospondin-1
TUNEL	Terminal deoxynucleotidyl transferase deoxyuridine trisphosphate
TWIST1	Twist family BHLH transcription factor 1
T $\beta$ R	Transforming growth factor $\beta$ receptor
T $\beta$ RI	TGF- $\beta$ type I receptor
UCSC	University of California Santa Cruz
ULBP	UL16-binding protein
uPAR	Urokinase plasminogen receptor
USF	Upstream stimulatory factor
V-ATPase	Vacuolar-type H <sup>+</sup> -ATPase
Vangl	Van Gogh-like
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
Veh.	Vehicle
VHL	von Hippel-Lindau
w/v	Weight to volume ratio
WB	Western blot
WPRE	Woodchuck hepatitis virus posttranscriptional regulatory element
YAP	Yes-associated protein
ZNF280A	Zinc finger 280A
$\alpha$	Alpha
$\beta$	Beta
$\beta$ -TrCP	$\beta$ -transducin repeats-containing protein
$\mu$ l	Microlitre
$\mu$ m	Micrometre

## Chapter 1

### Introduction

#### 1.1 Breast Cancer

Breast cancer (BC) is currently the joint first most common cancer across the sexes, worldwide. In 2018, this accounted for a total of approximately 2.1 million new breast cancer cases and over 600,000 deaths. Female breast cancer (FBC) represents 99% of all diagnosed breast cancers and poses a significant clinical burden for women globally [1, 2]. FBC is a malignant disease resulting from the chronic proliferation of breast epithelial cells [3]. The breast tissue lies on the anterior chest wall over the pectoralis, extending from the second to sixth rib in the midclavicular line [4, 5]. Breast tissue is composed of glandular and adipose tissue within a fibrous framework, known as Cooper's ligaments [6]. Histologically, secretory epithelial cells form clusters of alveoli, which form lobules and a series of inter-connecting ducts; these form the lobes of which there are 15-20 radiating out from the nipple [5, 6]. In BC, the majority of neoplasms originate from lobules and ducts [5].

FBC is a heterogeneous disease encompassing a range of biologically distinct pathological features and prognoses. Molecular profiling has identified five intrinsic subtypes of BC, for each of which there are further molecular subclassifications: luminal A (LABC), luminal B (LBBC), human epidermal growth factor receptor 2 (HER2)-enriched (HER2+), basal-like (BLBC) and normal-like (NLBC) [7]. Male breast cancer (MBC) is distinct to that of FBC, on account of risk factors, genetics, epigenetics, molecular profiles and response to therapies [8]. Whilst MBC and FBC are reported to share some pathological characteristics, most available treatment strategies on MBC is derived from studies involving females [9]. In women, age is the most common risk factor for development of BC [10]. However, BC type 1 and 2 susceptibility proteins (*BRCA1* and *BRCA2*) germline mutations have been identified as risk factors, as well as positive family history of BC, early menarche, late first full-term pregnancy and Caucasian race [10].

The prognosis of BC patients depends on potential lymph node involvement, size of the primary tumour, grade and subtype [14]. In England, women diagnosed with BC at stage 1 or 2 have a similar 1-year survival rate to the general population, likely as a result of advances and education in symptom recognition, and access to screening programs and improved therapies [15]. For patients with stage 4 BC at the time of diagnosis, the 1-year survival rate is significantly reduced, at 63% [15] and therefore the prognosis for metastatic disease is much less favourable. Following the removal of a primary breast tumour, local irradiation, chemotherapy or endocrine therapy may be administered, as adjuvant therapies, to treat potential remaining micrometastases [11]. When first introduced, such therapies were correlated with a 25% reduction in BC-associated mortality at ages 20 – 69 years [13]. However, 30% of women with breast-confined cancers and 75% with lymph node involvement will eventually relapse [11, 14]. Relapse usually occurs within the first 2 – 3 years following removal of the primary tumour, depending on initial tumour stage at treatment, therapy given and tumour biology [18]. Recurrent BC occurs in approximately 40% of BC patients and is almost always fatal [18]. In stage I-III breast cancer patients who are treated with surgery, radiotherapy and adjuvant chemotherapy, the median survival time from first relapse is 17 months [16]. For BC-related deaths, 90% of deaths occurred within 40 months of diagnosis of recurrence, with TNBC being a significant risk factor for patients who died within 1 year from diagnosis [16]. The source of distant recurrences is likely to be from cells which disseminated from the primary tumour that persisted asymptotically, despite systemic therapies [17]. Common sites of metastasis in BC include the bone, lung, liver, lymph nodes and the brain [18]. BC, after lung cancer, is the second most common cause of brain metastases (BrMs) [19].

## **1.2 Breast Cancer Brain Metastases**

The development of BrMs in BC patients is associated with an extremely poor prognosis. The percentage of metastatic BC patients who develop BrMs is 15-30% [20]. Survival rate is dependent on BC subtype, with median survival times from diagnosis of BrMs of triple negative, HER2+ and luminal breast cancers being 3.7, 9 and 15 months, respectively [21]. There is clear unmet medical need in the identification of exploitable

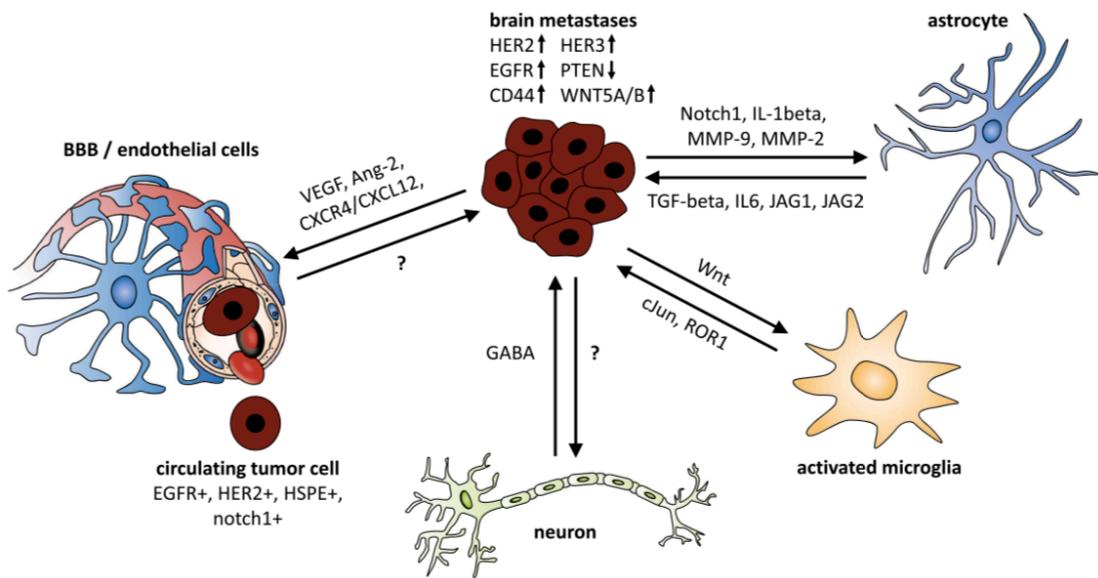
cellular pathways for the treatment of BCBrMs and the development of novel therapeutics. Shown in **Table 1.1** is a comparison of primary breast cancer, compared to brain metastases in terms of their composition and response to therapy.

Primary Breast Cancer	Brain Metastases
<ul style="list-style-type: none"> <li>- The tumour microenvironment of primary breast cancer typically contains vasculature, lymphatics, fibroblasts, pericytes, sometimes adipocytes, myeloid-derived suppressor cells (MDSCs), mesenchymal stem cells (MSCs), monocytes and endothelial cells of surrounding vasculature [22, 23]</li> <li>- Present with a high degree of inflammation [24]</li> <li>- Typically, primary breast cancer is responsive to adjuvant therapies [13]</li> <li>- Leukocyte infiltration can be observed at the invasive edge of the tumour [23]</li> </ul>	<ul style="list-style-type: none"> <li>- Blood-brain barrier is difficult to penetrate with therapies [25].</li> <li>- Brain metastases tend to be faster growing and resistant to chemotherapy, for example by upregulation of HER3 [26].</li> <li>- Tumour microenvironment typically contains endothelial cells of the blood-brain barrier, neurons, activated microglia and astrocytes [27]</li> <li>- Contains few bone-marrow-derived immune cells, known to be “immune-privileged” [28]</li> <li>- Microglia are abundant within the brain and may possess a tumour-supportive phenotype [29].</li> </ul>

**Table 1.1. A comparison between primary breast cancer and brain metastases.**

Where there is intracranial involvement in BC, there is no high-level randomised data to guide treatment [30]. Surgery and systemic therapy may be offered to improve quality of life for such patients, however there is little to no benefit to overall survival [30]. There is also infiltrative growth of cancer cells beyond the border of brain metastases in the majority of cases, making treatment very difficult [31]. Furthermore, due to increasingly advanced imaging techniques available and improved treatment of primary breast cancers, BCBrMs are also becoming more frequent [32, 33]. The discovery and validation of novel drug targets for treating BCBrMs is often hindered as a result of patients with BCBrMs not being allowed to enter into clinical trials [26]. This is due to the concern that most therapies do not penetrate the blood-tumour barrier with much efficacy [26]. For example, patients given the monoclonal antibody trastuzumab (Herceptin) for primary breast cancer experience a reduction in overall breast cancer mortality by one fifth after two years of follow-up [34]. Until recently, it was thought

that trastuzumab is unable to cross the blood-brain-barrier (BBB), like many other monoclonal antibodies [25]. However, there is evidence from animal models and patient data which suggests that the BBB is sufficiently disrupted in brain metastatic patients to allow permeation of monoclonal antibody-based therapies [26]. As shown in **Figure 1.1**, there are also a wide number of ligands exchanged between disseminated breast cancer cells to the brain and local brain populations. This is highly influential on the characteristics of brain metastases, determining treatment response and growth [35].



**Figure 1.1. Potential crosstalk between breast cancer cells disseminated to the brain and local brain cell populations.** Breast cancer cell dissemination to the brain is reliant on direct and indirect cell interactions to enable blood brain barrier penetration and establishment within brain parenchyma. Interactions between host and cancer cells induce activation of a number of signal transduction pathways in both cell types. Epidermal growth factor receptor (EGFR); gamma-aminobutyric acid (GABA); human epidermal growth factor receptor (HER); interleukin (IL); jagged (JAG); matrix metalloproteinase (MMP); transforming growth factor (TGF); vascular endothelial growth factor (VEGF), phosphatase and tensin homolog (PTEN); cluster of differentiation (CD); tyrosine-protein kinase transmembrane receptor (ROR); C-X-C chemokine receptor (CXCR); C-X-C motif chemokine ligand (CXCL); heparanase (HSPE) [35]. *Breast Cancer Research*. 2016; 18: 8, Copyright 2018. DOI: 10.1016/j.ijrobp.2006.03.050

Statistically, there is a significant bias for cerebellar metastasis in BC [36]. However, the spatial distribution of BCBrMs also varies according to BC subtype [36]. TNBCs tend to metastasise more often to the frontal lobe, limbic region and parietal lobe. BCBrMs from HER2+ cancers occur less frequently in the frontal lobe and subcortical region, whilst

those from luminal type occur less frequently in the occipital lobe, subcortical region and cerebellum [36]. Overall, there are three types of central nervous system complications seen in BC: brain parenchymal metastases (most common complication), leptomeningeal carcinomatosis (5% of BCBrMs) and intracranial dural metastases (infrequent but mostly seen in breast and prostate cancers) [37]. Compared to other common sites of metastasis for BC, BCBrMs develop late in disease progression [38]. The brain is seldom the first site of metastasis among BC patients for any subtype (**Table 1.2**) [27, 39]. The “Seed and Soil” hypothesis dictates that some organs present a more fertile environment than others for certain cancer metastases [40]. Circulating tumour cells must penetrate the blood brain barrier, establish a metastatic niche and successfully colonise the brain, before metastatic outgrowth can occur [41].

Site of metastasis	Brain (%)	Bone (%)	Lung (%)	Liver (%)	Pleura (%)
All subtypes	7-16	40-51	13-22	6-18	ND
Luminal A	2	47	8	18	7
Luminal B	0	35	16	12	12
TNBC	10	29	21	10	7
HER2+	2	29	23	27	8

**Table 1.2. Reported first site of breast cancer metastases as a percentage of cases.** Human epidermal growth factor receptor (HER); triple negative breast cancer (TNBC). Not determined (ND) [27]. *Chinese Clinical Oncology*. 2018; 7(3). Copyright 2018. DOI: 10.21037/cco.2018.05.06

Consequently, a major question remains as to whether there is anything unique about the interaction of BC cells and the brain. Is there a mechanism to explain relapse of breast cancer years after successful removal of the primary tumour, or why brain metastases develop late in disease progression? This may be explained by the phenomenon of cancer dormancy. Our understanding of the mechanisms surrounding dormancy are currently limited. Further investigations of the biology of dormancy in this emerging field may open new avenues for the treatment of metastatic disease.

### **1.3 Introduction to Cancer Dormancy**

Cancer dormancy refers to the extended period of time in which disseminated tumour cells (DTCs) persist asymptotically and undetectable, in a reversible state of quiescence [42, 43, 44]. Resurgence from dormancy under more favourable conditions may be associated with potentially lethal metastatic progression [42, 43, 44]. Despite unmet medical need, our understanding of the mechanisms surrounding the induction of, maintenance and escape from a dormant state remains relatively poor. This is despite the concept of a dormancy phenomenon first being postulated in 1934 [45]. Three types of dormancy have been described: cellular (unique to single cell populations), angiogenic and immunologic dormancy [46, 47]. The latter two are referred to as tumour mass dormancy.

#### **1.3.1 Tumour Mass Dormancy**

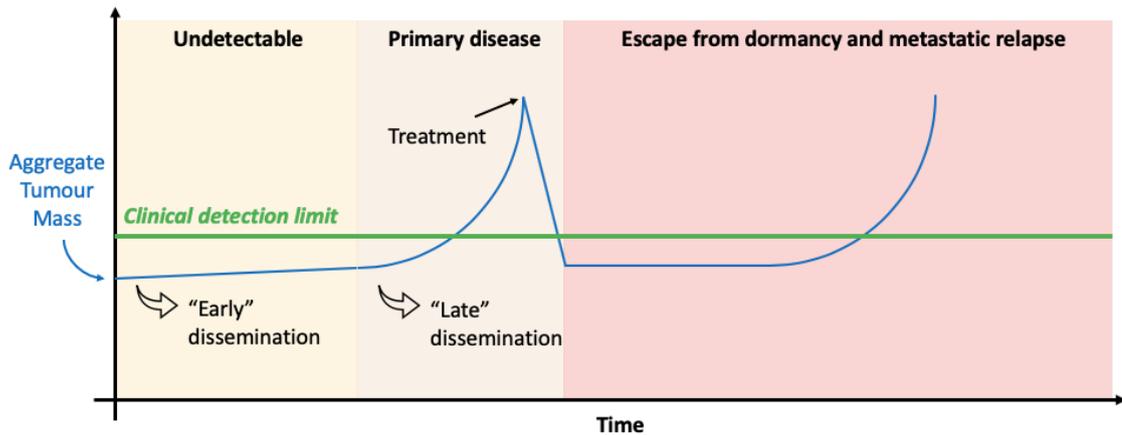
Normal cellular growth and survival, in both neoplastic and non-neoplastic cells, depends on the regular supply of oxygen and nutrients through a network of vasculature [48, 49]. Oxygen molecules have a diffusion limitation of 150-200 $\mu$ m away from blood vessels, therefore the growth of tumours beyond 1-2mm in diameter is an angiogenesis-dependent process [50, 51]. In this context, angiogenic dormancy, a subtype of tumour mass dormancy, refers to inhibition of the formation of macroscopic lesions as consequence of non-functioning angiogenic mechanisms [52]. For example, a disruption in the balance between pro-angiogenic factors, such as vascular endothelial growth factor (VEGF), and anti-angiogenic factors, such as thrombospondin-1 (TSP-1) [52]. The second mediator of tumour mass dormancy are host immune populations. Cells from both innate and adaptive immune system are involved in the detection and destruction of cancer cells [43]. When the growth of micrometastases is in equilibrium with the rate at which the immune system can destroy them, this is immunologic dormancy. This can occur through the induction of a host-protective immune response or an adaptation of cancer cell-surface receptors to match that of their microenvironment [43].

### 1.3.2 Cellular Dormancy

Cellular dormancy refers to solitary cancer cells, disseminated to distant tissues and within a reversible state of quiescence (**Figure 1.3**) [44]. Clinically, dormant cells are problematic as they present with a resistance to chemotherapeutic agents, which typically target actively proliferating cancer cells [14, 53]. Biologically, each of the breast cancer subtypes display phenotypic variations, however the biology of DTCs may also be divergent of that of the primary tumour, which may explain why adjuvant therapies do not always result in a favourable clinical outcome [54, 55]. Transcriptomic profiling of primary breast tumours has made it possible to predict, to some extent, the prognosis of cancer patients in terms of whether a cancer will remain localised, or if it will become metastatic with the potential to relapse [56]. Despite this, solitary cancer cells cannot be detected in living patients, therefore it is not currently possible to predict where, when and if breast cancer will relapse with any accuracy. This makes treatment difficult.

A number of sources provided evidence for the existence of dormant cells in patients. For example, disseminated solitary cells have been observed in the brains of cancer patients at autopsy [57]. Dormant cells are also found in the brains of cancer patients dying of non-cancerous causes [58]. Chronologically, evidence suggests that the generation of such cells may occur very early in cancer pathogenesis. Cancer progression can be explained by one of two models. The first is the parallel progression model, in that metastasis occurs early in disease progression and disseminated cells evolve independently, in parallel, to the primary tumour. Alternatively, the linear model describes that metastases are derived from late-disseminating clones that resemble the primary tumour [59]. Current evidence suggests that the generation of dormant cells may follow the parallel progression model; that dormant DTCs exist long before any treatment is administered to the patient (**Figure 1.2**) [60]. In a spontaneous mouse model of melanoma, for example, metastases were observed throughout the body in addition to the primary tumour, before the primary tumour was clinically detectable [61]. In this model, disseminated tumour cells also remained dormant for extended periods of time, mediated in part by cytostatic CD8+ T cells [61]. It has also been shown that breast cancer cells can disseminate systemically from the earliest epithelial

alterations in HER2 and PyMT mouse models [62]. Mechanistically, phenotyping of a population of early DTCs within the bone marrow determined these cells to be HER2<sup>+</sup>/p-p38<sup>lo</sup>/p-ATF2<sup>lo</sup>/ TWIST1<sup>hi</sup>/E-CAD<sup>lo</sup> [63]. Her2<sup>+</sup> early DTCs also activated a Wnt-dependent epithelial-mesenchymal transition (EMT), which was reversible upon inhibition of HER2 or Wnt [63]. This may provide a mechanism as to how early dissemination occurs and may imply a potential mechanism for the induction of a dormant state.



**Figure 1.2. Cancer progression and dormancy.** A patient who first develops breast cancer will begin with an initial tumour in the breast. The mass, before it is a size that can be detected clinically, may initiate early dissemination of cancer cells from the breast to distant tissues. In the primary disease stage, the aggregate tumour mass will be large enough for clinical detection. Here, the patient will have standard treatment, typically a lumpectomy, potentially followed by mastectomy, chemotherapy, radiotherapy, or a combination. Late dissemination of cancer cells may occur. There may be residual disseminated cells that have evaded treatment and exist in an asymptomatic, reversible state of quiescence or cell cycle arrest. After an extended period of time, these cells may undergo potentially lethal metastatic relapse. Adapted from [44, 54]. *Cancer Research*. 2015; 75(23): 5014 – 5022, Copyright 2015, DOI: 10.1158/0008-5472.CAN-15-1370. *Nature Reviews Cancer*. 2014; 14: 611 – 622, Copyright 2014, DOI: 10.1038/nrc3793.

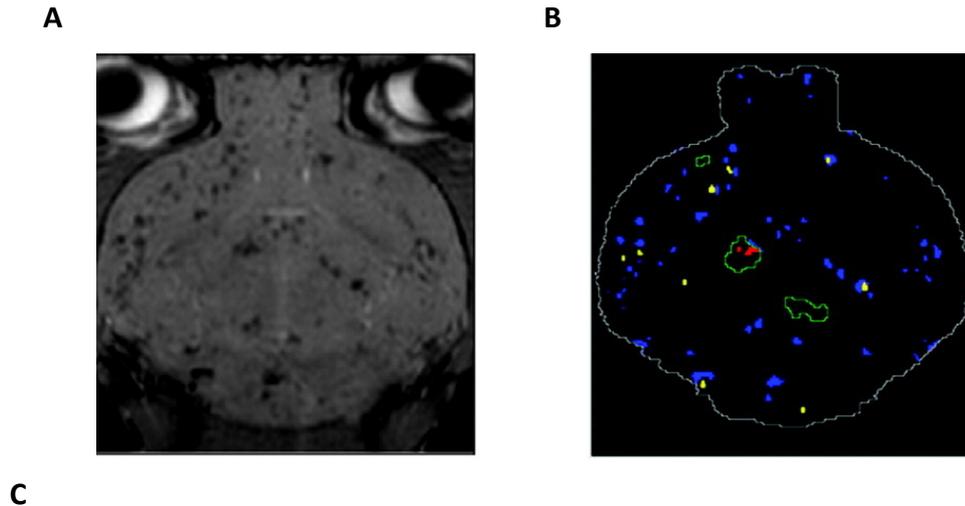
In clinical settings, DTCs have been detected in patients before the primary tumour would be considered as invasive [62]. Such cells were reported to display increased expression of epithelial cytokeratins, and lacking chromosomal abnormalities, compared to the primary tumour [64]. Ductal carcinoma *in situ* (DCIS) is a non-invasive breast cancer, yet DTCs have been detected within the bone marrow in some patients [65]. In addition, patients have relapsed with contralateral carcinoma and contralateral DCIS, despite initially being diagnosed with non-invasive DCIS [66]. Despite its rarity,

some patients with DCIS, even with complete mammary tissue resection, metastasis was observed [67]. This has been documented in other reports of metastatic development, following the complete resection of stage M0 primary breast cancer [68]. It is therefore reasonable to hypothesise that the establishment of cellular dormancy could also occur early, generating clinically occult single cells long before the time of first diagnosis.

### **1.3.3 Cancer Cell Fate following Dissemination to the Brain**

Experimentally, previous studies have demonstrated the ability to track cancer cell fate upon dissemination, whether they perish, proliferate or transition into dormancy. In an *in vivo* model of breast cancer dissemination, Micron-sized superparamagnetic Iron Oxide (MPIO)-labelled brain-homing MDA-MB-231 cells labelled with enhanced green fluorescent protein (eGFP) were administered by intracardiac injection into nu/nu mice (**Figure 1.3**). MPIO contains iron which is detectable by magnetic resonance imaging (MRI). Interestingly, 93.9% of cells injected were no longer present by day 28, whereas 4.5% remained dormant in the brain. Only 1.6% began actively proliferating, which formed the seeds for subsequent metastatic lesions [69].

Similar results were obtained in an experimental metastasis model involving the intraportal injection of B16F1 melanoma cells, to target mouse liver. This study found that 80% of injected cells persisted in the liver microcirculation for 3 days before extravasating into liver tissue. However, only 1 in 40 extravasated cells formed micrometastases, with 1 in 100 continuing to grow by day 13. Interestingly, using Ki67 and terminal deoxynucleotidyl transferase deoxyuridine triphosphate (dUTP) nick end labelling (TUNEL) staining as markers of proliferation and apoptosis, respectively, they found that the majority of the 36% of injected cells which survived to day 13 were dormant [70].



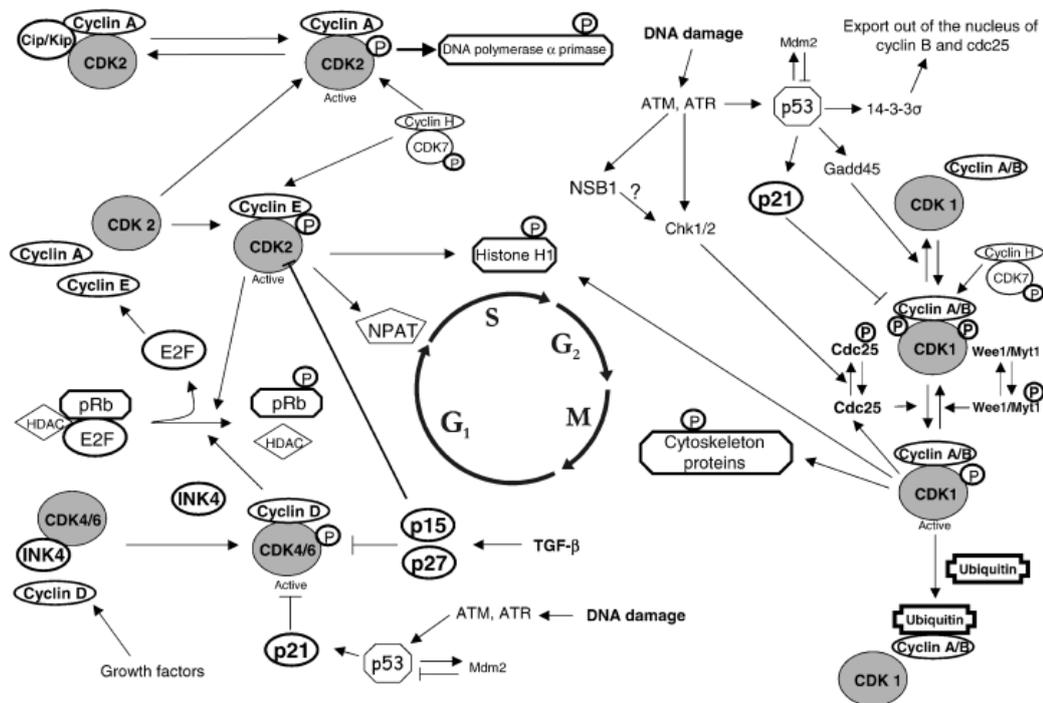
	Tumour contrast volume at day 28/ mm <sup>3</sup>	Percentage signal drop from background brain compared to day 0
<b>Transient cells</b>	33.5 +/- 3.3	93.9 +/- 0.07
<b>Non-proliferating cells</b>	1.60 +/- 0.13	4.5 +/- 0.8
<b>Proliferating cells</b>	0.56 +/- 0.07	1.6 +/- 0.06

**Figure 1.3. Tracking of breast cancer cell fate in the brain by magnetic resonance imaging.** A)  $1 \times 10^5$  brain-homing MDA-MB-231/eGFP cells were introduced by intracardiac injection into female nu/nu mice. Black voids represent metastatic cancer cells in the brain, as observed by MRI imaging on day 0. B) Cancer cell fate 28 days post-injection. Blue: transient cells which were absent at day 28 but present at day 0; yellow: dormant, non-dividing cells; green outline: established lesions within the brain, of which the associated cancer cells are shown in red. C) Table displaying cancer cell fate in terms of tumour contrast volume and expressed as a percentage drop of day 0 tumour volume. Figure adapted from [69]. *Magnetic Resonance in Medicine*. 2006; 56: 1001 – 1010, Copyright 2006, DOI: 10.18383/j.tom.2016.00151

#### 1.4 The Cell Cycle, Cancer and Dormancy

All dividing cells transition through phases of growth and division; this is known as the cell cycle and it is critical for long-term survival [71]. The function of the cell cycle is to duplicate genetic material [72]. Cell division in somatic cells produces two daughter cells in a process called mitosis [72], whilst meiosis in germline cells produced four haploid daughter cells from one diploid oocyte [73]. Mitosis, like meiosis, relies on accurate

deoxyribonucleic acid (DNA) replication, chromosome segregation and DNA repair mechanisms [72]. In cancer, a failure of surveillance mechanisms which monitor the order, integrity and fidelity of processes which occur during the cell cycle leads to unchecked proliferation, or replicative immortality [74, 75]. It is not clear as to whether dormant cells exist in sustained quiescence, outside of the cell cycle, or whether intrinsic or extrinsic mechanisms are holding cancer cells long-term in mid-cycle arrest. The cell cycle is a vastly complex process, characterised as a transition through distinct phases as a cell grows and replicates (**Figure 1.4**). At its most basic level, the eukaryotic cell cycle is divided into four distinct phases. Gap 1 (G1) phase is the first growth phase, followed by entry into synthesis (S) phase, in which the aforementioned duplication of genetic material occurs [72, 76]. A secondary growth phase occurs in the following gap 2 (G2) phase, followed by mitosis (M) phase, which involves the separation of sister chromatids and separation into daughter cells by cytokinesis [76]. Pertinent to dormancy and cell cycle arrest, cell cycle progression, in terms of the order, integrity and fidelity of major processes, is strictly monitored by cell cycle checkpoints [75].



**Figure 1.4. An overview of some of the critical regulators in cell cycle regulation.** Phosphorylated sites are denoted by 'P'. Activation and inhibition are denoted by  $\rightarrow$  and  $\dashv$ , respectively [77]. *Trends in Cell Biology*. 2013; 23(7): 345 – 356, Copyright 2013, DOI: 10.1016/j.tcb.2013.03.002

The transitions through cell cycle checkpoints and entry into each cell cycle phase depends on the time-sensitive activation of specific serine/threonine protein kinases, the cyclin-dependent kinases (CDKs) [75, 78]. This occurs through their association with various cyclins (**Table 1.3**) [77]. Unlike CDKs, the cyclins are synthesised and degraded according to cell checkpoint. At the critical time, synthesised cyclins bind and activate CDKs to drive cell cycle progression [75, 76, 78]. The major checkpoints are those controlling entry into the G1/S and G2/M phase transitions [76]. Cell cycle arrest can occur at a number of timepoints depending on context, for example S, sub-G1, G0/G1, G0/G1 and G1 [79]. In non-cancerous cells, cell cycle checkpoints become activated, for example during the DNA damage response (DDR) [80]. For example, this may lead to stabilisation of p53, a potent tumour suppressor, which induces p21 activation, inhibiting cyclin A/CDK2 and cyclin E/CDK2 complexes, leading to cell cycle arrest and apoptosis [81]. Cancer cells are defective in such checkpoint mechanisms, for example p53 may be lost, leading to uncontrolled proliferation [78, 80]. A further example of uncontrolled cell cycle progression in cancer is the aberrant activity of the oncogene, *MYC*. In non-cancerous cells L-, N- and C-Myc are a family of transcription factors involved in the expression of a number of genes involved in cellular growth and proliferation. In cancer, gene amplification, chromosomal translocation, or deletions of upstream *MYC* regulators leads to sustained oncogenic *MYC*-induced proliferation. In combination with a loss of p53, *MYC* is a highly potent promotor of carcinogenesis [82].

CDK	Cyclin	Cell Cycle Phase Involvement
CDK4	D1, D2, D3	G1 phase
CDK6	D1, D2, D3	G1 phase
CDK2	E	G1/S phase
CDK2	A	S phase
CDK1	A	G2/M phase
CDK1	B	Mitosis
CDK7	H	CAK (all phases)

**Table 1.3. Association of cyclin-CDK complexes within each point of the eukaryotic cell cycle.** CDK activating kinase = CAK [77].

All cells, including cancer cells, also have the ability to enter into another phase of replicative dormancy, known as quiescence [72]. Cells in G1 phase can remain in the cell cycle or withdraw from it and enter a fifth, and non-proliferating, quiescent phase, G0 [72, 83]. In the current literature, cancer dormancy is commonly characterised by a reversible state of quiescence, specifically in G0 phase [84]. Entry into the quiescent G0 phase is not limited to cancer cells but is known to occur in a number of differentiated cell types, typically stem cells [85]. Exiting from the active cell cycle may be in response to environmental changes, such as cell density, adhesion or nutritional stress [86]. The potential mechanisms of cancer cell quiescence will be explored further in this chapter. However, there is no evidence to suggest that dormancy may be characterised by mid-cycle arrest. In an *ex vivo* hepatic microphysiologic system, consisting of human hepatocytes and non-parenchymal cells, spontaneous growth arrest was achieved in MCF7 and MDA-MB-231 cells [87]. In this system, growth arrest was determined via the addition of 5-ethynyl-2'-deoxyuridine (EdU) for 96h, to which some cells remained negative after this time point [87]. This indicates that those cells had not undergone cell division and were quiescent [87]. However, there was no indication as to whether cells were arrested in G0 or mid-cycle.

In cancer, a number of chemotherapeutic agents exert their effects via cytotoxic mid-cycle arrest. In the MDA-MB-468 breast cancer cell line, Cucurbitacin E (CuE) induces G2/M arrest, with subsequent onset of apoptosis by activation of c-Jun N-terminal kinase (JNK) and an inhibition of protein kinase B (PKB/AKT) and extracellular regulated kinase (ERK) [88]. In these cases, it was independent drug-modulated mechanisms which induced cell death. However, the long-term (5 day) administration of Gallotannin (GltN), for example, to a number of breast cancer cell lines, resulted in a resistance of progression through S phase, without inducing apoptosis [89]. Therefore, it is possible for cancer cells to exist in the mid-cell cycle arrest for a prolonged time and this possibility should be taken into consideration.

## 1.5 Major Factors and Pathways Involved in Cancer Cell Dormancy

Cancer cells are the result of aberrant clonal outgrowth of malignant cell populations, whose uncontrolled proliferation defines the disease [90]. By the time a tumour is detected, a single cancer cell may have given rise to potentially billions of progenies [91]. In normal human cells, DNA is sometimes incorrectly replicated during cell division, which on average occurs at a rate of  $2.3 \times 10^{-8}$  mutations, per nucleotide, per cell, per generation [92]. However, DNA instability is a common characteristic of cancer, which in tumour growth results in accumulating genetic changes, leading to subclonal evolution of cancer cells [91]. As a result, tumours usually present with a wide range of intratumoral heterogeneity [91]. In addition to cancer cells, which may account for as few as 30% of cells in the tumour cell population, tumours represent complex ecosystems of a variety of heterogeneous cell types [93]. This is referred to as the tumour microenvironment and its composition has major implications for a tumour's biology, influencing carcinogenesis, cancer proliferation, invasion and metastasis [94]. The tumour microenvironment may therefore be critical for the induction and maintenance of cellular dormancy.

### 1.5.1 The Extracellular Matrix

The cells of the tumour microenvironment, and of normal tissues, are supported by a rich protein scaffold, the extracellular matrix (ECM) [94, 95]. The ECM is composed of fibrous proteins, glycoproteins, proteoglycans and polysaccharides [96]. As well as its composition biochemically, the functions of the ECM are defined by its topology [97], its density [98], stiffness [99] and tension [100]. Cancer cells are known to actively remodel their local ECM, a process which aids their survival and progression [101]. The microenvironment of organs containing disseminated cancer cells is also critical to the formation of metastases. In particular, the brain is very different to the breast, and other organs, in terms of the composition of its ECM [117]. The brain may not be immediately hospitable to cancer cells originating from the breast, and the differences in microenvironment composition may be a driver in dormancy induction. In an *in vivo* model of tumour cell invasion to the brain, for example, brain-seeking MDA-MB-231

cells display increased invasion in astrocyte-conditioned medium [103]. Astrocytes were observed to secrete matrix metalloproteinase-2 (MMP-2) and MMP-9, tumour invasion factors [103]. Injection of astrocyte-conditioned medium *in vivo* increased cancer cell invasion to the brain, whilst administration of broad spectrum MMP inhibitors reverses this effect [103]. Some of the ways in which ECM components cancer cell fate are outlined in **Table 1.4**.

ECM Component	Example	Effect on cancer
Molecular tracks	Examples include F-actin and myosin, Rho family members, adhesion proteins and adaptor proteins [104]	Cancer cells migrate along tracks as they undergo epithelial-to-mesenchymal transition [104]
Embedded fibroblasts	Tumour-induced cancer-associated fibroblasts (CAFs) [104]	CAFs cross-link ECM components, increasing stiffness and therefore promotes cancer proliferation [104, 106]
Signalling molecules	TGF- $\beta$ ligands [104]	TGF- $\beta$ and other growth factors are stored within the ECM and released under tension, promoting proliferation [104, 107]
Adhesion molecules	Integrins [104]	As well as binding cancer cells to the ECM, interactions with, and signal transduction from, ECM components are mediated through adhesion molecules [104]
Scaffold proteins	Multifunctional SLRPs, such as biglycan, decorin and versican [104, 105]	Regulators of signalling pathways, many of which are crucial for cancer growth and proliferation [104, 105]

**Table 1.4. The influence of ECM components on cancer cells**

### 1.5.1.1 Extracellular Matrix Stiffness

The mechanical properties of tissues reflect the composition and arrangement of the ECM, cytoskeleton, cellular component tension, fluid dynamics and tissue organisation [108]. Individual cells within tissues possess the ability to sense the physical stiffness of their local microenvironment, applying contractile force through transcellular structures, with implications for development, disease and regeneration [109].

Effect	Mechanism
Women with the relatively densest breasts are 4 to 6 times more likely to develop breast cancer [110].  Cancer progression in DCIS [128].	Transformation and increased rigidity are typically due to an increase in collagen I deposition, whilst interstitial collagen I presents with progressive linearisation and thickening, promoting carcinogenesis, invasion and metastasis [106].
Proliferation promotion in MDA-MB-231 cells <i>in vitro</i> in stiff ECM [112].	Nuclear localisation of Yes-associated protein (YAP) [112].
Proliferation promotion of hepatocellular carcinoma cells <i>in vitro</i> in stiff ECM [113].	Stiffness sensing by integrin- $\beta$ 1 and subsequent focal adhesion kinase (FAK) signalling [113]
Induction of malignant phenotype in normal mammary MCF10A cells in stiff ECM [114].	Increased ECM stiffness sensed through cell surface $\beta$ 4 integrin, which induces increased Ras-related C3 botulinum toxin substrate 1 (Rac1) and phosphoinositide 3-kinase (PI3K) signalling [114].

**Table 1.5. The effect of ECM stiffness on cancer proliferation**

Tissue stiffness is determined by the amount of force required to deform it, measured as a unit of elastic modulus,  $E$  (Pa). The lung and brain are amongst the softest tissues in the body ( $E \leq 10^2$  Pa), with muscle tissue as mid-range ( $E \leq 10^4$  Pa) and bone as the stiffest ( $E \leq 10^9$  Pa) [108]. The primary sites of interaction between cells and the surrounding ECM are through focal adhesions, as well as focal complexes, fibrillar adhesions, podosomes and three-dimensional matrix adhesions [115, 116]. All of the described cell-ECM adhesion structures connect integrins at the cell surface to the actin cytoskeleton [116]. Effectively, the stiffness of tissues and surrounding ECM translates to intracellular signalling [117]. The highest levels of stiffness are observed at the invasive front of the tumour [106]. As a result, the induction and/or maintenance of

cellular dormancy may be regulated by the stiffness of the ECM at sites of dissemination, as stiffness has shown to be highly influential on cancer proliferation (**Table 1.5**) and growth suppression/dormancy (**Table 1.6**) in a number of models.

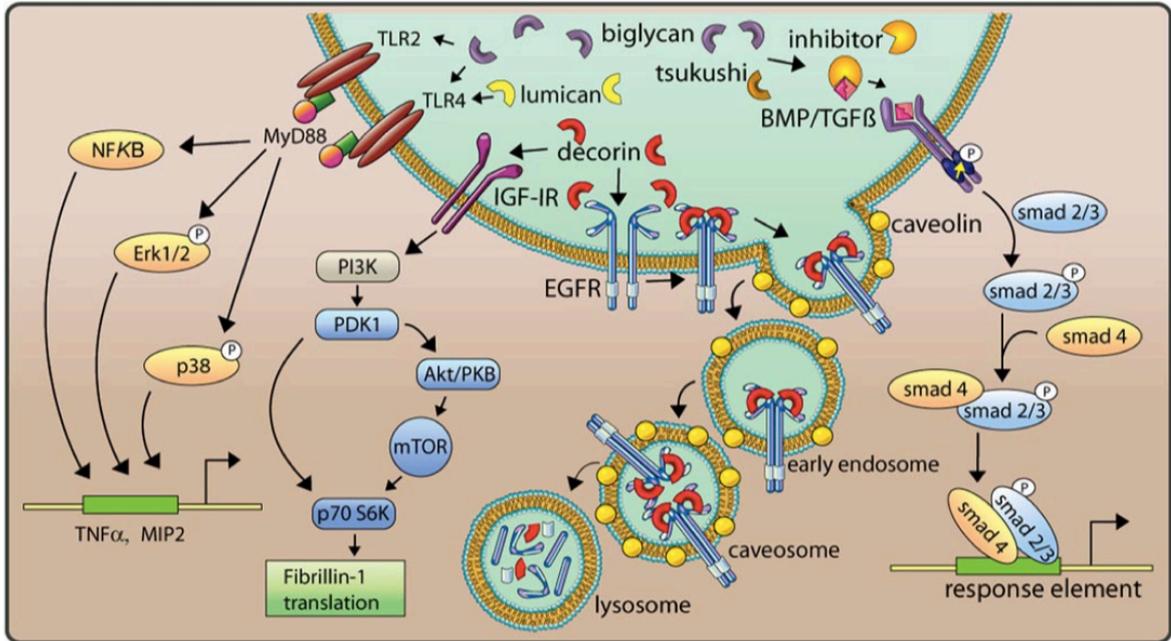
Effect	Mechanism
Dormancy induction in murine B16 and human A375 melanoma cells <i>in vitro</i> , in stiff ECM [118].	Nuclear translocation of cell division protein 42 (CDC42), activating p21 and p27 CDK inhibitors, in addition to downregulation of integrin- $\beta$ 3 [118]
Reduced proliferation of glioma cell lines U373-MG, U87-MG, U251-MG, SNB19 and C6, in soft polymeric matrix <i>in vitro</i> [119].	Reduced signalling through integrins, focal adhesion proteins, Rho GTPases and the cytoskeleton. Results in alterations to shape, plasticity, motility and proliferation [119]
Reduced proliferation of lung cancer A549 cells on semi-solid growth substrate, compared to plastic, and a significant increase in the IC <sub>50</sub> of anticancer drugs [120].	Inhibition of ERK1/2 and PI3K/Akt pathways, leading to decreased expression of genes associated with tumour progression (urokinase ( <i>uPA</i> ), urokinase receptor ( <i>uPAR</i> ), <i>MMP2</i> , <i>MMP7</i> , <i>MMP9</i> and <i>CXCR4</i> ) [120].

**Table 1.6. The effect of ECM stiffness on growth inhibition and dormancy in cancer**

### 1.5.1.2 Biglycan

Biglycan is a multivalent short leucine rich proteoglycan (SLRP), which constitute a family of proteoglycans characterised by their low molecular weight, shared structural motifs and leucine-rich repeats [105]. Extracellular biglycan can interact with receptors at the cell surface, modulating a number of intracellular signalling pathways (**Figure 1.5**) [121]. Notably, biglycan has been commonly observed as a ubiquitous ECM component, and is a potent inhibitor of TGF- $\beta$  receptor activity, synthesised downstream of an active TGF- $\beta$  signal transduction pathway, acting as a TGF- $\beta$  negative regulator [122, 123]. Many of the biglycan-induced downstream effectors are well described in dormancy, for example Smads, ERK, p38 and NF- $\kappa$ B [121]. Furthermore, there is evidence to suggest biglycan may localise to the nucleus to regulate cell division and survival [122, 124]. The presence of potential nuclear localisation signals meant that biglycan is postulated to play a role in cell division regulation, through its interaction with nuclear processes [124]. To my

knowledge, there has been no studies to date specifically focused on the role of intracellular or nuclear biglycan in the context of cancer.



**Figure 1.5. Signal transduction pathways activated downstream of extracellular biglycan.** A number of distinct pathways are influenced by extracellular SLRPs. Shown are biglycan and decorin (class I SLRPs), lumican (class II SLRP) and tsukushi (class IV SLRP). A number of SLRP family members can bind to and modulate the TGF- $\beta$ /BMP pathways, for example biglycan and decorin. Decorin can also modulate signalling via epidermal growth factor receptors (EGFRs) and insulin-like growth factor 1 receptors (IGF-IRs). Additionally, biglycan and lumican can activate the NF- $\kappa$ B and members of the mitogen-activated protein kinase family, via binding to and modulation of Toll-like receptors (TLRs) [121]. *Journal of Biological Chemistry*. 2008; 283(31): 21305 – 21309. Copyright 2008, DOI: 10.1074/jbc.R800020200

Biglycan expression is typically associated with endothelial cells, fibroblasts and blood vessels, and has been shown to be distributed throughout rat brain parenchyma, particularly after injury. In particular, regarding the brain, biglycan has been shown to be expressed by astrocytes and meningeal cells, presenting with neurotrophic activity [124]. X-ray crystallography has shown that biglycan core protein, and the closely related decorin, in solution forms stable dimers, which is crucial for protein folding mediated by parallel  $\beta$ -sheet faces [125]. It is the core protein of biglycan which is responsible for its biological activity and interaction with ECM components and cell surface receptors [126]. Under normal physiological conditions, biglycan is usually sequestered by the ECM, typically through binding to components such as collagen and

elastin. This process is crucial for organising ECM assembly [127]. However, in this state, biglycan is insoluble and cannot participate in the activation or inhibition of signal transduction pathways [127, 128]. Partial proteolytic cleavage of biglycan is required in order for it to be released from the ECM and become soluble [128].

### **1.5.2 The Perivascular Niche**

The late steps of metastasis constitute an inefficient process, with less than 0.02% of circulating tumour cells becoming DTCs [129]. Relative sizes of cancer cells and capillaries lead to the arrest of most circulating cancer cells within the first available capillary bed. However, most cancer cells perish before they can generate metastatic outgrowth at secondary sites, dictated by the molecular interactions of the cancer cells with the environment of the distant organ [130]. The perivascular niche (PVN) is characterised by the interactions of cancer cells with the endothelial cells of tissue vasculature and the surrounding extracellular matrix [131]. In brain tumours, the PVN forms one of the most important microenvironment hubs [132]. For DTCs, there is an increasing amount of literature which evidences the PVN as an essential regulator of their growth [133, 17].

In a previous study, upon injection of GFP-tagged MDA-MB-453 cells into the internal carotid artery of BALB/c mice, solitary cancer cells were observed scattered throughout the brain parenchyma and actively proliferating metastases were observed 30 – 70 days post-injection [134]. Only 12% of solitary cells were positive for the proliferation marker Ki67, whereas for larger lesions 63% were Ki67-positive and thus the majority of solitary cells were considered as quiescent, in G<sub>0</sub>-phase of the cell cycle. Solitary, Ki67-negative cancer cells were observed within approximately 10µm of brain microvasculature endothelium, implicating the PVN as supportive of a growth-arrest phenotype. No significant difference in the percentage of Ki67 cells was detected between the solitary cells and macroscopic lesions in the MDA-MB-231/brain model [134]. The MDA-MB-231/brain cell line in this study possessed a brain homing phenotype; such cells possess selectivity for brain colonisation [135] and therefore this may account for this observation. As further evidence that the PVN is supportive of a growth-arrest

phenotype, it has been shown that a critical step in metastatic outgrowth following latency is the spreading of cancer cells on host capillaries [136]. This activates the mechanotransduction effector YAP via cell adhesion molecule 1 (L1CAM), promoting proliferation. Whilst this study implicates L1CAM as a critical mediator of metastatic outgrowth but not dormancy regulation, this study highlights that impaired cancer cell interaction with vascular endothelial cells may be supportive of dormancy induction.

Tumour vessels and blood vessels consist of endothelial cells and mural cells (either pericytes or smooth muscle cells), surrounded by a basement membrane [139]. The basement membrane, otherwise referred to as the extracellular milieu, has previously been evidenced as playing a key role in tumour dormancy and cytotoxic resistance [138]. This is unsurprising given that the process of metastasis requires that cancer cells must cross vascular and non-vascular basement membranes [148]. Basement membranes possess vastly complex supramolecular structures. The basic structure of vascular basement membranes consists of layered proteins, glycoproteins and proteoglycans [137]. They are constructed as a strong covalently bonded mesh, formed by a number of laminins and type IV collagen polymers, bridged by nidogens; they are also bonded to netrins, perlecan and agrins [139].

Laminins constitute major components of both the vascular and parenchymal basement membranes [140]. They are large (400 – 900 kDa) heteromeric glycoproteins constructed of an  $\alpha$ ,  $\beta$  and  $\gamma$  chain, enabling a number of isoforms with hugely diverse functions in adhesion, differentiation, proliferation, polarity, tissue development, cellular communication, migration, phenotype stability and resistance to anoikis [141, 142]. The effects of laminin typically occurs via binding with cell surface integrin receptors, which consist of heteromeric  $\alpha$  and  $\beta$  subunits [143]. Laminins may also signal through the non-integrin laminin, dystroglycan and heparan sulphate receptors [143]. The integrins are one of the primary points of cellular adhesion to the ECM/basement membrane, presenting with the accumulation of phosphorylated proteins at their cytoplasmic domain upon laminin-binding and activation of a wide range of signal transduction pathways, such focal adhesion kinase (FAK), Rho guanosine triphosphate (GTP)ases, mitogen-activated protein kinases, phospholipase D, intracellular calcium signalling and G-proteins [143]. In breast cancer,  $\alpha 5 \beta 1$  integrin is one such receptor

upregulated in malignant cells and it is associated with cancer proliferation and survival [101]. Whilst the major model of integrin signalling in cellular proliferation is via positive regulation upon ligation and integrin crosslinking, additional direct growth inhibitory properties have also been observed by unoccupied  $\alpha 5\beta 1$  in non-adherent cells [144]. As such, dormancy may occur as a result of failure to adhere to the basement membrane of the PVN upon dissemination. With regards to specific laminins, laminin-8 in brain tumours is associated with the highly invasive and malignant glioblastoma multiforme, whereas an alternative isoform, laminin-9, was associated with low grade astrocytomas [145]. The overexpression of laminin-8 was also associated with a significantly shorter time to tumour recurrence in human glial tumours than those overexpressing laminin-9 [146]. In relation to breast cancer, the presence of laminin-2, laminin-8 and laminin-10 within vascular basement membranes are associated with metastatic progression in the brain [140]. It has further been suggested that blocking the binding of  $\alpha 3\beta 1$  integrin to laminin-111 could prevent the reactivation of dormant cells [147]. However, the proliferation effects of  $\alpha 3\beta 1$  could only occur when thrombospondin 1 (TSP-1) was degraded [147]. Interestingly, TSP-1 is secreted in response to inflammation, abrogating inflammatory processes which then enables phagocytosis of damaged cells to occur [148]. As such, the activity of TSP-1 may need to be overcome in order for inflammatory signals to either re-awaken dormant cells or promote proliferation upon dissemination of cancer cells to distant organs.

TSP-1 has been further implicated in dormancy induction in the context of the PVN. In a study of microvasculature and dormancy, MDA-MB-231 breast cancer cells were injected into the inguinal mammary gland of NOD-SCID mice. After 3 weeks, the tumours were resected. In mice that displayed no relapse after a further 6 weeks, the mice were culled. Using absence of Ki67 expression as a metric of growth arrest, dormant DTCs were observed localised along lung and bone marrow microvasculature endothelium [149]. In the same study, an alternative model using intracardiac injection of T4-2 cells, Ki67-negative DTCs were observed localised to the microvasculature endothelium of the lungs, bone marrow and brain [149]. Proteomic and functional analyses identified that TSP-1 was acting as a tumour suppressor in the proximity of stable vasculature, whereas

at the neovascular tips, sites of metastatic outgrowth, TSP-1 was degraded [149]. This suggests a niche in which tumour cells may escape growth regulation.

Furthermore, in breast and lung adenocarcinoma models, serpins are associated with brain metastatic phenotypes [150]. Serpins, serine proteinase inhibitors, represent the largest and diverse group of protease inhibitors, exerting their effects through the inhibition of serine proteinases or inhibition of catalytic activity, in processes such as hormone transport [151]. In the brain, plasmin can suppress metastatic outgrowth either through conversion of astrocytic FasL into a cancer cell paracrine death signal, or inactivation of L1CAM, a protein critical for cancer cell spreading on brain vasculature [150]. Breast and lung cancer cells which express neuroserpin and serpin B2, anti-plasminogen activator (PA) serpins, prevent plasmin generation and are thus able to undergo metastatic outgrowth [150]. Finally, the PVN also has implications for cancer therapy. There is consensus that dormant, Ki67-negative DTCs are protected from chemotherapy simply by nature of their non-dividing, quiescent phenotype [152]. However, medulloblastomas are an example of brain tumours which give rise to recurrence following radiotherapy [153]. Following  $\gamma$ -ionising irradiation, the majority of cells undergo p53-induced cell death. Despite this, a small number of cancer stem cells within the PVN, that are nestin-positive, undergo p53-induced reversible cell cycle arrest via activation of the PI3K/Akt pathway [153]. As well as protection from radiotherapy, the PVN also provides protection from chemotherapy. It has recently been shown that chemotherapy resistant dormant DTCs associate with the PVN [133]. Interestingly, the inhibition of integrin-mediated interactions between DTCs and the PVN allowed for sensitisation to chemotherapy, independent of cell cycle status [133]. This would allow for the treatment of dormant cells using existing chemotherapies, without the risk of having to first induce their reactivation.

### **1.5.3 Hypoxia**

Hypoxia is defined as a region of low oxygen tension [154]. Metabolic adaptations of cancer cells in response to a hypoxic microenvironment can result in the generation of both highly malignant cells and cancer cells that become non-proliferating [154, 155]. In

tumour expansion, cancer cell proliferation can rapidly exhaust available oxygen, and glucose as the primary cancer energy source [156]. Activated in response to hypoxia is the hypoxia-inducible factor 1 alpha (HIF-1A) transcription factor [157]. Under plentiful supply of oxygen, HIF-1A is hydroxylated at its proline residues by prolyl-4-hydroxylases (PHDs) and polyubiquitinated by von Hippel-Lindau (pVHL), resulting in 26S proteasome-mediated degradation [158]. Under hypoxic conditions, HIF-1A is stabilised, undergoes nuclear translocation, and dimerises with HIF-1 beta (HIF-1B) to initiate gene transcription [158]. HIF-1A induces expression of a wide variety of downstream genes, which exert effects in metabolism, neovascularisation, intracellular acidity, cellular survival, migration and proliferation [159]. HIF-1A activity, or its inhibition, may therefore be critical to dormancy induction and/or maintenance.

Exposure of serum-deprived MDA-MB-231 breast cancer cells *in vitro* to hypoxia enhances survival in a vascular endothelial growth factor (VEGF)-dependent manner [160]. Such conditions induced increased expression of *VEGF*, a known target gene of HIF-1A [160, 161]. This stimulated PI3K in an autocrine fashion, resulting in phosphorylation and activation of Akt/PKB, with subsequent intracellular signal transduction leading to increased cell survival [160]. Therefore, if hypoxia is a driver of dormancy induction, a potential mechanism mediating the long-term survival of dormant cancer cells may be through the autocrine effects of VEGF. Furthermore, hypoxia-resistant breast cancer cells (chMDA-MB-231), generated by exposure to three cycles of 1% O<sub>2</sub> followed by reoxygenation, enter a state of G<sub>0</sub>/G<sub>1</sub> arrest and are characterised by low metabolism [162]. Under hypoxic conditions, a stem-like phenotype was observed within the cancer cells, marked by a CD24<sup>-</sup>/CD44<sup>+</sup>/epithelial specific antigen (ESA)<sup>+</sup> phenotype. The proposed mechanism was a hypoxic stress-induced selection of MDA-MB-231 cells, of which the surviving population may persist via an upregulation of autophagy. Exposure to a normoxic environment induced re-proliferation of the cells after 2 weeks [162]. Reversible growth arrest is characteristic of the dormancy phenomenon, and therefore dormancy may be induced by transient exposure to a hypoxic microenvironment.

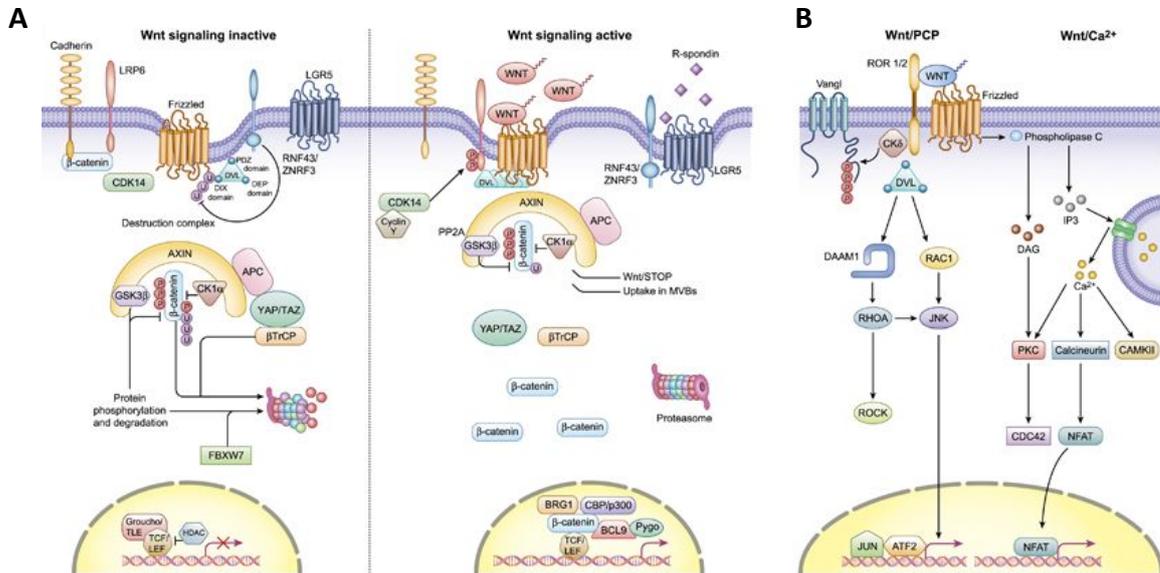
Interestingly, the microenvironment of cancer cells within the primary tumour influences cancer cell fate upon dissemination. In head and neck squamous cell carcinoma and breast cancer, the primary tumour microenvironments exhibited increased expression of HIF-1A and glucose transporter 1 (GLUT1) [173]. These were accompanied by an upregulation of key dormancy factors (nuclear receptor subfamily 2 group F member 1 (NR2F1), basic helix-loop helix family member e41 (BHLHE41/DEC2) and p27) [163]. This phenotype persists in cancer cells which disseminate post-hypoxia and were commonly dormant and NR2F1<sup>hi</sup>/DEC2<sup>hi</sup>/p27<sup>hi</sup>/TGFβ<sup>hi</sup> [164]. Therefore, the microenvironment to which primary breast cancer cells are exposed to may influence the fate of cancer cells which undergo dissemination.

#### 1.5.4 Wnt Signalling Cascades

The Wnt signalling cascades, regulators of cellular proliferation and polarity, are divided into the canonical (β-catenin-dependent) and the non-canonical (β-catenin-independent) (**Figures 1.6 A, B**) [165]. In cancer, aberrant Wnt signalling promotes tumorigenesis in a variety of tissues, including the breast [166]. Wnt signalling is also known to crosstalk with other pathways, for example fibroblast growth factor (FGF), Notch, Hedgehog and TGF-β signalling [167]. As such, the influence of Wnt signalling on the induction of dormancy may be context dependent. There are a number of reported mechanisms in which active Wnt signalling may induce dormancy. In interleukin-23 (IL-23) receptor-positive human oesophageal squamous carcinoma cells, cells which were negative for CD133 acquired a dormant-like G0/G1 phase arrest, through IL-23-mediated Wnt/Notch signalling [168]. Inhibition of IL-23 via small interfering ribonucleic acid (siRNA) or small molecule inhibitors abolished these effects [168]. This also suggests that dormancy may only be induced in a subset of cancer cells.

Conversely, in a mammary tumour model, amplification/overexpression of both Wnt and FGFR1 was associated with poor disease-specific survival [169]. Treatment with an FGR1 inhibitor, NVP-BGJ398, induced substantial tumour regression. Interestingly, some cancer cells persisted in a reversible state of dormancy for as long as treatment was

administered [169, 170], indicating that dormancy induction may arise from an inhibition of Wnt co-operation with other intracellular pathways.



**Figure 1.6. Overview of the canonical and non-canonical Wnt signalling pathways.** A) In canonical Wnt signalling, when Frizzled is unbound by Wnt, the cytoplasmic destruction complex, containing glycogen synthase kinase  $\beta$  (GSK3 $\beta$ ) and Axin, binds and phosphorylate  $\beta$ -catenin.  $\beta$ -catenin subsequently dissociates from the complex and is ubiquitinated by  $\beta$ -transducing repeats-containing protein (TrCP), followed by degradation by the proteasome. Wnt binding to Frizzled induces association of Axin with phosphorylated low-density lipoprotein receptor-related protein 1 (LRP), dissociating the destruction complex and stabilising  $\beta$ -catenin. This allows  $\beta$ -catenin to enter the nucleus to facilitate gene transcription [183]. B) Non-canonical Wnt signalling is defined by  $\beta$ -catenin-independent mechanisms. Wnt ligands bind to ROR-frizzled receptor complex. This activates dishevelled (Dvl), which binds the Rho guanosine triphosphate (GTP)-ase through an activation of Dvl associated activator of morphogenesis 1 (DAAM1). Rac1 and Rho GTPases can then trigger rho-associated protein kinase (ROCK) and JNK, leading to a number of transcriptional effects. Alternatively, adjacent to Dvl, the cell polarity protein and negative regulator of the canonical Wnt signalling pathway, Van Gogh-like (Vangl), may also be activated in a Wnt5a-dependent manner. Finally, Wnt/Ca<sup>2+</sup> signalling can be activated via phospholipase C, inducing intracellular calcium fluxes and subsequent transcriptional activation. [165, 171]. *Oncogene*. 2017; 36: 1461 – 1473. Copyright 2017, DOI: 10.1038/onc.2016.304

Furthermore, an inhibition of canonical Wnt signalling in the osteoblastic niche has been observed to induce dormancy. Secreted from dormant cells within the bone microenvironment is Wnt5a, which induces Siah E3 ubiquitin protein ligase 2 (SIAH2) expression, which represses canonical Wnt signalling [172]. Thus, there is some

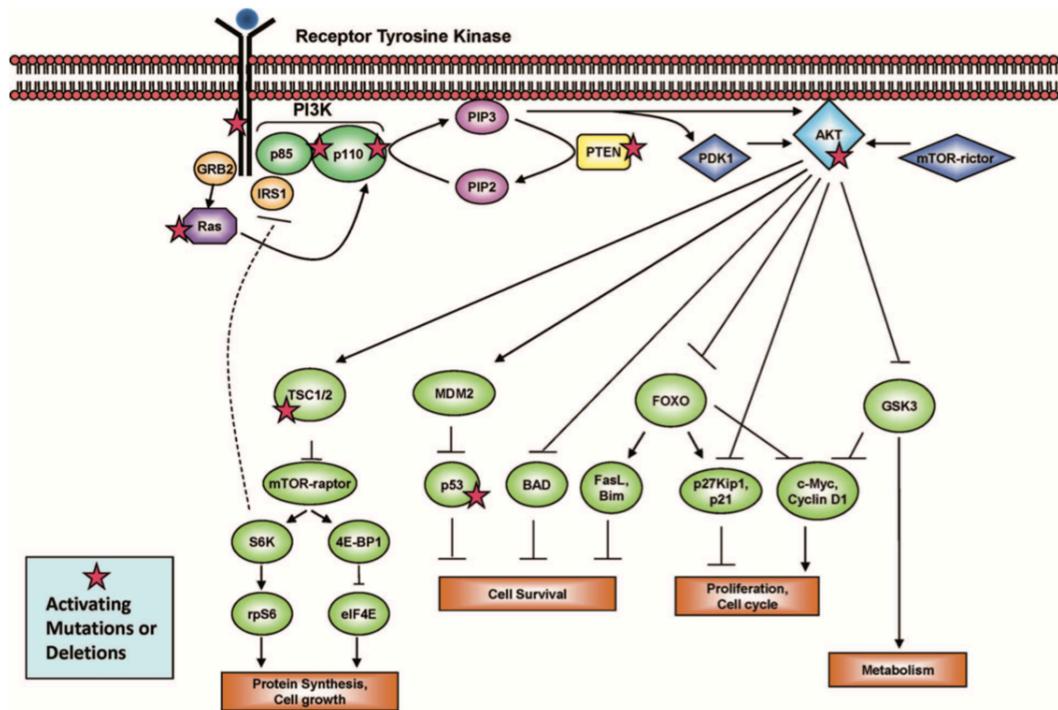
evidence of potential cross-pathway regulation in dormancy. Finally, expression of Sry-type high mobility group (HMG) box 2 and 9 (SOX2 and SOX9) survival factors have been observed in lung- and breast-derived latency competent cancer cells (LCCs). LCCs also downregulate UL16-binding protein (ULBP) ligands for natural killer cells, via expression of Dickkopf-related protein 1 (DKK1), a Wnt inhibitor [173]. Wnt signalling inhibition in dormancy may therefore assist in the evasion of local immune populations.

### 1.5.5 PI3K-Akt-mTOR Pathway

The phosphoinositide 3-kinase (PI3K)-protein kinase B (AKT)-mammalian target of rapamycin (mTOR) pathway (PI3K-AKT-mTOR pathway), is frequently hyperactivated in human cancers [174]. Its activity is associated with a range of cellular processes, including cell growth, survival, proliferation, metabolism, motility and cancer cell dissemination [174, 175]. The PI3Ks constitute a family of lipid enzymes which exert their effector functions through the phosphorylation of 3'-OH group of plasma membrane phosphatidylinositols [176].

As shown in **Figure 1.7**, activation of the PI3K pathway mediates its effects via phosphorylation of the second messenger, AKT [177]. Inhibition and reactivation of the PI3K pathway may be critical for dormancy induction and cancer relapse, respectively [178]. In laryngeal squamous cell carcinoma (LSCC) cells, activation of aurora kinase A (AURKA) induces transition from G<sub>0</sub> phase into active cell cycle entry, whilst its expression has been observed to be persistently low in dormant LSCC cells [178]. AURKA induces activation of the FAK/PI3K/Akt pathway to promote proliferation and migration, as well as reactivation of dormant cells [178]. Inhibition of the PI3K pathway may also directly influence cell cycle checkpoints. In a study of floating tumour spheroids of ovarian cancer cell lines, detached from the ECM, the cells entered a state of quiescence, in G<sub>0</sub> [179]. This was attributed to increased AKT inhibition, reduced expression of cyclin D1 and an upregulation of p130 and p27, mediated via autophosphorylation of epidermal growth factor receptor (EGFR)-Y1086, downstream of an inhibited PI3K-AKT signalling pathway [179]. Finally, Insulin-like growth factor 1 (IGF-1) plays a key role in breast cancer development. The IGF-1 receptor (IGF-1R), a receptor tyrosine kinase, is

activated upon IGF-1 binding [180]. This induces activation of the PI3K pathway, in addition to the MAPK pathway [181]. IGF-1 is mostly secreted by stromal cells and influences breast cancer growth via paracrine signalling [182]. Under serum-deprived stress, cancer cells have been shown to secrete clusterin, which binds and inhibits the binding of IGF-1 to its receptor, inducing quiescence [183]. This negatively regulates the PI3K-Akt pathway and therefore a potential dormancy mechanism may be due to the interaction between clusterin and IGF-1 upon cancer cell dissemination [183].

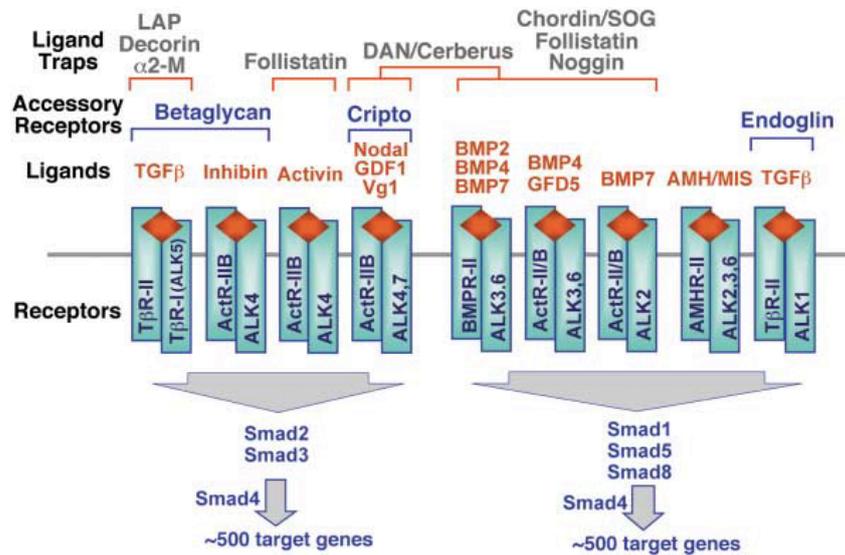


**Figure 1.7. Simple schematic of the PI3K-Akt-mTOR signalling pathway in cancer.**

The PI3K pathway, activated downstream of cell-surface receptor tyrosine kinases (RTKs), has important downstream widespread effector functions, including cell growth, cell survival, proliferation, differentiation and metabolism. Some components within the pathway, denoted by a red star, can be subject to activating mutations or deletions in cancer. Abbreviations: Bcl-2-associated death promotor (BAD); growth factor receptor-bound protein 2 (GRB2); insulin receptor substrate 1 (IRS1); murine double minute (MDM2); mammalian target of rapamycin (mTOR); 3-phosphoinositide-dependent protein kinase 1 (PDK1); phosphoinositide-3-kinase (PI3K); phosphatidylinositol bisphosphate (PIP2); phosphatidylinositol triphosphate (PIP3); phosphate and tensin homologue deleted on chromosome ten (PTEN); regulatory associated protein of TOR (RAPTOR); rapamycin-insensitive companion of mammalian target of rapamycin (RICTOR); tuberous sclerosis (TSC) [175]. *The Oncologist*. 2011; 16(1): 12 – 19. Copyright 2011. DOI: 10.1634/theoncologist.2011-S1-12

### 1.5.6 The Transforming Growth Factor- $\beta$ (TGF- $\beta$ ) Superfamily

There are a growing number of studies that suggest that members of the TGF- $\beta$  superfamily (**Figure 1.8**) play a central role in cellular dormancy. In cancer, the role of TGF- $\beta$  signalling is complex, and whether the TGF- $\beta$ -related pathways activate or inhibit tumour proliferation is context-dependent. In early tumour progression, TGF- $\beta$  is frequently observed as a tumour suppressor, whereas it has been more often identified as a tumour promotor in the later stages of cancer [107].



**Figure 1.8. The TGF- $\beta$  superfamily: ligands, receptors and SMADs.** A schematic representation of TGF- $\beta$ , BMP and activin signalling. Shown are ligands, ligand-binding traps, accessory receptors and type I/II receptors. SMADs are grouped according to signalling specificity. At the cell surface, TGF- $\beta$  ligands bind serine/threonine kinase receptors [197]. Binding of ligands to their cognate receptors induces the activation of type I receptors by phosphorylation of serine/threonine residues in their GS region by type II receptors, propagating the signal by SMAD protein phosphorylation [197, 200] Receptor-activated SMADs (SMADs 1-8) form heteromeric complexes with co-mediator SMADs (co-SMADs) which translocate to the nucleus to regulate gene transcription [197, 200]. Abbreviations: transforming growth factor  $\beta$  (TGF- $\beta$ ); transforming growth factor  $\beta$  receptor (T $\beta$ R); activin receptor (ActR); anaplastic lymphoma kinase (ALK), bone morphogenic protein (BMP); Anti-Müllerian hormone (AMH/MIS); growth differentiation factor 1 (GDF1); leucyl aminopeptidase (LAP); short gastrulation (SOG) [184]. *Cell*. 2003; 113: 685 – 700. Copyright 2003, DOI: 10.1016/s0092-8674(03)00432-x

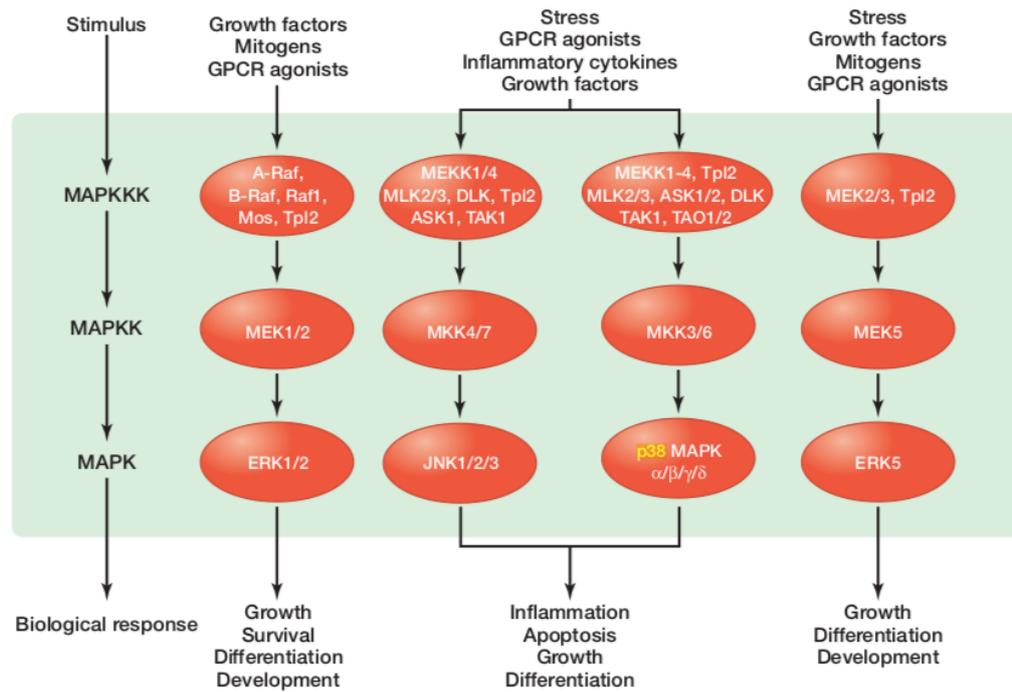
The majority of breast cancer metastases are positive for phosphorylated SMAD2, suggesting TGF- $\beta$  pathway activation [185]. TGF- $\beta$ , when secreted from astrocytes in the brain, induces expression of angiopoietin-like 4 (ANGPTL4) in MDA-MB-231 breast cancer cells via the induction of SMAD signalling, fuelling metastatic outgrowth [186]. TGF- $\beta$  may also directly regulate cancer cell cycle progression. TGF- $\beta$  signalling rapidly induces p21 and p27 transcription, cyclin kinase partners negative regulators [187, 188]. Furthermore, TGF- $\beta$  signalling induces expression of p15<sup>INK4B</sup>, an inhibitor of CDK4/6 activity [189]. BMP7 also induces cell cycle arrest in prostate cancer stem-like cells. This effect is mediated via the activation of p38, leading to increased expression of p21 and the metastasis suppressor, N-myc downstream-regulated 1 (NDRG1), and was reversible upon withdrawal of BMP7 [190]. TGF- $\beta$  also presents as part of a cross-pathway regulation of the dormancy phenomenon. With respect to TGF- $\beta$ 2, in a head and neck squamous cell carcinoma model, TGF- $\beta$ 2 signalling within the bone marrow was shown to regulate dormancy of DTCs [191]. TGF- $\beta$ 2 induced a state of dormancy via activation of the mitogen-activated protein kinase (MAPK) p38 $\alpha/\beta$ , thereby facilitating an ERK/p38<sup>Low</sup> signalling ratio [191]. As shown in Hep3 human head and neck carcinomas, the fate of DTCs i.e. whether they proliferate or enter a state of dormancy, is determined by the (ERK)<sup>MAPK</sup>/p38<sup>SAPK</sup> activity ratio and is regulated by extracellular signalling [192].

### 1.5.7 Mitogen-Activated Protein Kinase Family and Cancer Dormancy

The Ras-Raf-MEK-ERK pathway is widely associated with cellular differentiation, proliferation and apoptosis [193]. The MAP kinases sit within a number of kinase cascades, transmitting a diverse range of extracellular signals to cellular responses (**Figure 1.9**) [194]. Activation of MAPK signalling pathways potentially has a number of effects in dormant cancer cells.

Firstly, Rho-guanosine triphosphate (GTP)-ases have been shown to be hyperactivated in dormancy, leading to JNK phosphorylation and activation of autophagy [195]. Secondly, tumour suppression and dormancy induction is particularly associated with the stress activated MAPK, p38 [196]. Computational analysis of human squamous cell

carcinoma identified that quiescence was regulated by the co-ordination of 16 transcription factors, mediated by p38 activity [196]. In a 4T1 spontaneous metastasis model, in Balb/c mice, inhibition of lysophosphatidic acid receptor (LPA1), a G protein coupled receptor (GPCR), induces a state of metastatic dormancy [197]. This state was characterised by reduced expression of Ki67, with downregulation of pERK and an increase in the phosphorylation of the p38 stress kinase [197].



**Figure 1.9. An overview of the MAPK pathways.** A wide range of extracellular cues act through one of the many MAPK pathways to induce a biological response. At the top level, a stimulus activates a MAPK kinase kinase (MAPKKK), typically via small GTPase interactions or phosphorylation downstream of cell surface receptors. This is followed by phosphorylation of a MAPK kinase (MAPKK), which phosphorylates a conserved tripeptide TXY motif in the activation segment of the final MAPK. The activated MAPK is then able to induce cytosolic changes to affect protein function or participate in the induction of gene expression in the nucleus. Abbreviations: rapidly accelerated F ibrosarcoma (Raf); moloney sarcoma (Mos); tumour progression locus 2 (Tpl2); mitogen activated protein kinase kinase (MEK); mixed-lineage protein kinase (MLK); dual leucine zipper kinase (DLK); apoptosis signal-regulating kinase (ASK); thousand-and-one amino acid (TAO); G protein-coupled receptor (GPCR) [198]. *Cold Spring Harbor Perspectives in Biology*. 2012; a011254. Copyright 2012, DOI: 10.1101/cshperspect.a011254

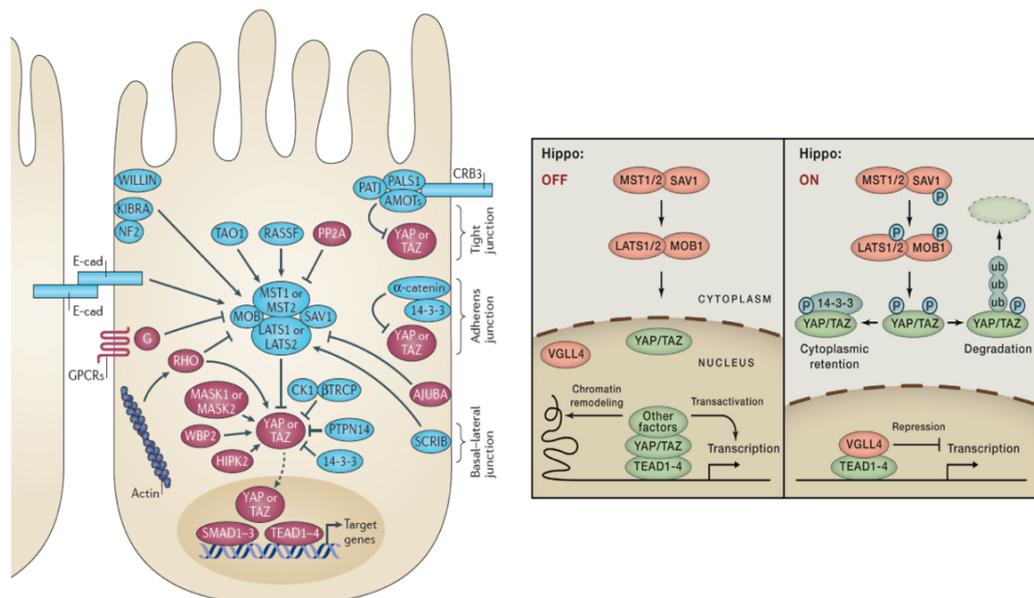
Finally, when metastatic human squamous carcinoma cells (T-HEp3) are passaged for >40 generations, they reprogram into a phenotype characterised by reversible dormancy upon inoculation *in vivo* [199]. Dormancy in this model was marked by G0/G1 growth-arrest, mediated by mitogen-activated protein kinase 6 (MKK6) and p38 $\alpha$ / $\beta$  [199]. In turn, these contribute to the nuclear translocation and transcriptional activation of ATF6 transcription factor, a prominent promotor of cellular survival [199]. As such, p38 activity may therefore be critical in the long-term survival of dormant cells. Cancer dormancy may also be regulated by crosstalk between integrins and MAPK signalling. In human carcinoma HEp3 cells, inhibition of urokinase plasminogen receptor (uPAR) induces G0/G1 cell cycle arrest *in vivo*, leading to reduced ERK1/2 activation [200]. It was found that uPAR forms complexes with  $\alpha$ 5 $\beta$ 1 integrins, which bind extracellular fibronectin [200]. This further suggests that detachment from the basement membrane, and perhaps the PVN, as previously discussed, may be critical in the induction of dormancy.

#### 1.5.8 HIPPO Signalling Pathway and Cancer Dormancy

The HIPPO signalling pathway (**Figure 1.10**) has increasingly been associated with cancer proliferation and dormancy in a number of tissues [201, 202]. The major HIPPO component frequently cited with respect to cancer is YAP. As discussed in Chapter 1.5.2, metastatic outgrowth of cancer within the brain is mediated via L1CAM activation of YAP, upon spreading of cancer cells on vascular endothelium. In 5-fluorouracil (5FU)-resistant colon cancer cells, exposure to high concentrations of 5FU results in reversible G0-phase growth arrest, and was associated with a proportional depletion of nuclear YAP, inhibiting its transcriptional activity [203]. Furthermore, YAP-dependent cellular proliferation has been shown to be mediated by an inhibition of PTEN, whilst YAP knock-down was strongly associated with growth-arrest [204]. In a heterogenous population of breast cancer cells, some cells may therefore be primed to enter dormancy upon administration of adjuvant therapies, potentially via YAP nuclear expulsion.

Similarly, YAP expression in clear cell renal cell carcinoma (ccRCC) positively correlates with cancer stage [205]. In 786-0 ccRCC cell lines, treatment with YAP-short hairpin RNA

(shRNA) lentiviral vectors results in G1-phase arrest, with some cells undergoing apoptosis [205]. YAP appears to be critical for cancer cell proliferation, and blockage of its transcriptional ability may be a driver of dormancy. Finally, YAP interactions with many of the beforementioned dormancy pathways have been described. For example, under hypoxic conditions, nuclear YAP stabilizes HIF-1A, activating pyruvate kinase M2 (PKM2) which accelerates glycolysis [206]. Furthermore, YAP has also been observed to function co-operatively with transcription factor AP-1 and SMAD7 to suppress TGF- $\beta$  signalling [207]. Together with TAZ, YAP has also been shown to work synergistically with the PI3K signalling pathway to promote mammary tumorigenesis [208]. Finally, cellular adhesion with the basal extracellular matrix is a major promoter of YAP nuclear localisation, which itself then activates transcription of a number of integrins and focal adhesion docking proteins [209, 210]. Importantly, YAP and its downstream targets may therefore represent very interesting drug targets in the therapeutic targeting of dormant cells, or the induction of dormancy in established lesions.

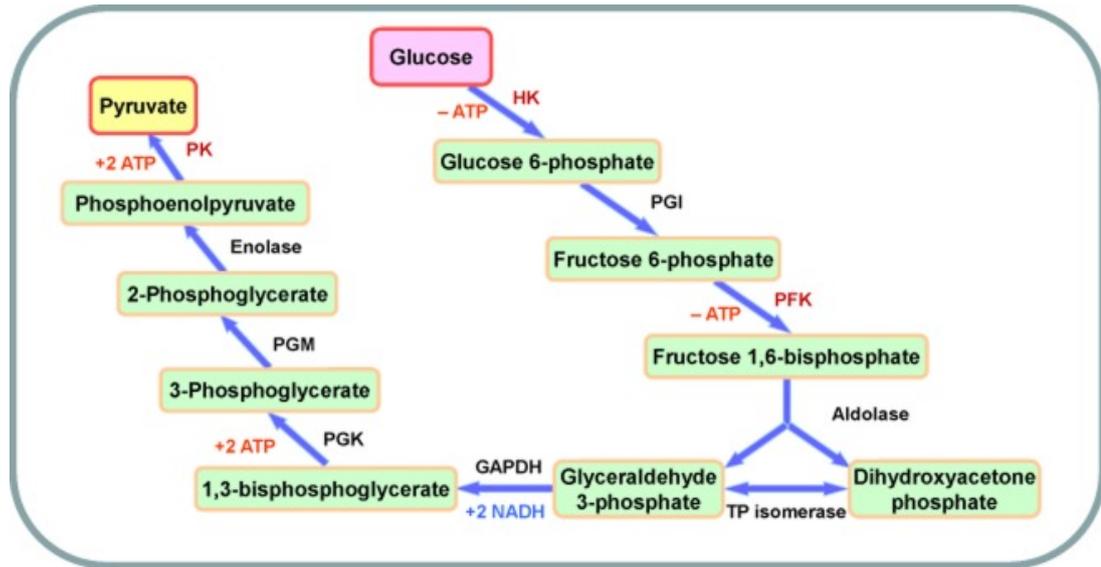


**Figure 1.10. HIPPO signalling pathway cascade in humans.** The core of the pathway consists of a serine/threonine kinase cascade [230]. In the centre lies macrophage stimulating protein 1/2 (MST1/2), followed by downstream kinases, large tumour suppressor 1/2 (LATS1/2), exerting their effects through phosphorylation and inhibition of the transcriptional co-activators yes-associated protein (YAP) and tafazzin (TAZ) [230, 231]. Active YAP/TAZ act with akyrin repeat and HK domain containing 1 (MASK1) or MASK2 co-factors to convert TEA domain 1-4 (TEAD1-4) to transcriptional activators, which in turn initiates gene transcription [211, 212]. Some upstream components may also interact directly with YAP/TAZ [201, 202]. *Nature Reviews Cancer*. 2013; 13: 246 – 257. Copyright 2013, DOI: 10.1038/nrc3458

### 1.5.9 Cancer Cell Respiration

Glycolysis is a process which takes part in the cytoplasm, involving 10 steps of enzyme-mediated reactions which converts glucose into pyruvate, generating ATP (**Figure 1.11**) [213]. Increased glycolysis activity, favoured over oxidative phosphorylation, even under aerobic conditions, is a hallmark of cancer and particularly favourable to the formation of breast cancer brain metastases, a phenomenon known as the Warburg effect [214, 215]. Increased metabolic gene signatures are a strong predictor of metastasis in breast cancer, lung cancer and melanoma [215]. As such, decreased expression of glycolysis enzymes would be unfavourable to cancer proliferation, and favourable to the induction and maintenance of dormancy. For example, as an effect of reduced AKT phosphorylation, induced via chronic hypoxia [216]. The Warburg effect is characterised by increased activity of HIF-1A and c-Myc transcription factors, with an upregulation of glucose transporters and glycolytic enzymes which are mediated by loss of p53 function [214].

The conversion of glucose ( $C_6H_{12}O_6$ ), to lactate ( $C_3H_6O_3$ ) during glycolysis, creates 2 molecules of ATP:  $C_6H_{12}O_6 + 2(ADP^{3-}) + 2(HPO_4^{2-}) \rightarrow 2(ATP^{4-}) + 2(C_3H_6O_3^-)$  [217]. It is the subsequent ATP breakdown which releases acidic  $H^+$  ions, protons [217]. In cancer, high rates of glycolysis correlate with high rates of proton transport into the extracellular space, facilitated by proton transporters such as  $Na^+-H^+$  exchanger (NHE), vacuolar-type  $H^+-ATPase$  (V-ATPase),  $H^+-K^+-ATPases$  and carbonic anhydrases (CAs) [218]. This acidifies the extracellular environment, changing pH from approximately 7.4 to 6.8, and alkalinises the intracellular environment, changing pH from approximately 7.2 to 7.6, creating a pH gradient which is critical for cancer cell survival and propagation; this has emerged as another hallmark of cancer [218].



**Figure 1.11. Schematic overview of glycolysis.** In the cytoplasm, glycolysis involves the sequential conversion of glucose to pyruvate, generating ATP. Specific enzymes catalyse each step of the glycolytic pathway. These are hexokinase (HK), phosphoglucose isomerase (PGI), phosphofructokinase (PFK), aldolase (ALDOA), triphosphate isomerase (TPI), glyceraldehyde 3 phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK), phosphoglycerate mutase (PGM), enolase (ENO) and pyruvate kinase (PK) [213]. *Thoracic Cancer*. 2015; 6(1): 17 – 24, Copyright 2015, DOI: 10.1111/1759-7714.12148

## 1.6 Exploitation of Dormancy Mechanisms – The Future of Cancer Therapy?

Despite emerging experimental data on the biology of dormant DTCs, the industrial response to the development of novel therapeutics specifically targeting dormant cells is limited. In 2019, HiberCell Inc. (USA) was the first private company constructed to exclusively focus on the detection and targeting of cancer dormancy [219]. Additionally, a small number of clinical trials have focused on the eradication of residual DTCs in breast cancer (**Table 1.7**). At the time of writing, there appears to be no clinical trials for therapies which target breast cancer cell dormancy in the brain. As such, this is an area of great unmet medical need. The identification of novel molecular mechanisms responsible for the induction and maintenance of breast cancer dormancy could give rise to novel therapeutics for the treatment of breast cancer brain metastases. There are three potential treatment strategies, each with their own advantages and disadvantages. The first strategy may involve the chronic maintenance of a dormant state within quiescent DTCs through the upregulation of key quiescence mechanisms. Alternatively, the same quiescence mechanisms may be used to induce and maintain a

state of dormancy in established lesions. For example, this may be via increased expression of p38 MAPK [192], or administration of ligands present within the dormant niche, for example TGF- $\beta$  [191]. This strategy would prevent the progression of breast cancer to lethal metastasis; however, this would also require chronic lifelong therapy and intra-/inter-tumoral heterogeneity may mean that not all cells within established metastases would respond.

Secondly, critical survival genes could be inhibited prior to progression to metastatic relapse, clearing the patient of residual dormant cells and significantly reducing the risk of relapse. For example, in *in vivo* pre-clinical and 3-dimensional (3D) *in vitro* models of breast cancer, pharmacological inhibition of Src family kinase (SFK) induces nuclear localisation of p27, preventing the reactivation of dormant cells and the formation of metastases [220]. As dormancy in these models required ERK1/2 activation, combinatorial treatment with Src and MEK1/2 inhibitors induced apoptosis in a large number of dormant cells [220]. However, if not all cells respond, as described here, this may transform the residual disseminated tumour cells to a more aggressive phenotype.

Finally, and perhaps the least suitable therapy, is to awaken dormant cells as such that they may be targeted by anti-proliferative chemotherapies. This approach depends on a highly effective therapy post-awakening of dormant cells. If the therapy is inefficient, or the cells do not respond, awakening dormant cells earlier than their natural progression may worsen a patient's prognosis. To support this argument, inhibition of TGF- $\beta$ -p38-MAPK signalling in head and neck squamous cell carcinoma induced reactivation of dormant cells and entry into the cell cycle [191]. This was accompanied by multi-organ metastasis [191]. In breast cancer, whole exome sequencing shows that brain metastases are highly divergent to that of the primary tumour, particularly with respect to sensitivity to PI3K/Akt/mTOR, CDK and HER2/EGFR inhibitors [221]. Therefore, it may not be the most logical approach to potentially induce multi-organ metastasis with increased chemoresistance. As such, whilst each of the described strategies may offer a novel therapeutic opportunity for the treatment of metastatic breast cancer, it is critical to first understand the dormancy mechanisms.

Study Title	Identifier	Therapy	Study summary
Gedatolisib, hydroxychloroquine or the combination for prevention of recurrent breast cancer (“GLACIER”)	NCT03400254 <i>Recruiting</i>	Drug (mechanism explored): <b>Gedatolisib</b> (PI3K and mTOR kinase inhibitor)  Drug (mechanism explored): <b>Hydroxychloroquine</b> (Numerous potential inhibitory targets – autophagy, TLR9, CXCL12/CXCR4 signalling, p53-associated pathways)	Targeting of residual disease in early-stage breast cancer patients with evidence of bone marrow DTCs
CLEVER pilot trial: A phase II pilot trial of hydroxychloroquine, everolimus or the combination for prevention of recurrent breast cancer	NCT03032406 <i>Recruiting</i>	Drug (mechanism explored): <b>Hydroxychloroquine</b> (as row above)  Drug (mechanism explored): <b>Everolimus</b> (mTOR kinase inhibitor)	Targeting of residual bone marrow DTCs following primary breast cancer therapy
Secondary adjuvant treatment for patients with isolated tumour cells in bone marrow	NCT00248703 <i>Active</i>	Drug (mechanism explored): <b>Docetaxel</b> (Inhibitor of microtubule disassembly)	Targeting of residual DTCs in bone marrow and following therapy (post-standard treatment for breast cancer)
Pilot study to evaluate the impact of denosumab on disseminated tumour cells (DTC) in patients with early stage breast cancer.	NCT01545648 <i>Active</i>	Drug (mechanism explored): <b>Denosumab</b> (Receptor activator of nuclear factor kappa-B ligand (RANKL) inhibitor – prevents osteoclast formation)	Targeting of residual DTCs in bone marrow following 12 months of therapy (post-standard treatment for breast cancer)

**Table 1.7. Clinical trials for the targeting of residual disseminated tumour cells in Breast cancer.** All studies are listed on the ClinicalTrials.gov online registry [222].

## 1.7 Hypothesis

The hypothesis for this study is that dormant and proliferating breast adenocarcinoma cells in the brain have distinct transcriptomic profiles. I reasoned that this would give rise to targetable molecular players and pathways which underpin the dormancy phenomenon.

### 1.7.1 Aims and Objectives

Targeting of identified dormancy-associated genes and pathways could lead to strategies for the induction of growth arrest in established tumours, or, alternatively, targeting of dormant cells prior to their exit from growth arrest. To identify such targets, this project aims to:

#### 1) Isolate dormant and proliferating human breast adenocarcinoma cells

This will be achieved through the use of an established *in vivo* experimental model of metastasis, using human breast adenocarcinoma cells to generate brain tumour xenografts in mice. This model will also be used to characterise the microenvironment of dormant breast cancer cells in the brain.

#### 2) Identify molecular differences between dormant and proliferating breast cancer cells *in vivo* by transcriptomic profiling

Isolated dormant and proliferating breast cancer cells will be harvested for RNA, followed by mRNAseq. Molecular differences, and thereby potential therapeutic candidates, will be identified by differential gene expression analysis, and by Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment analysis.

#### 3) Identify genes and pathways functionally implicated in breast cancer cell dormancy

This will be achieved by functional studies of candidate genes and pathways identified in the second aim, using *in vitro* models. In addition, my mRNAseq data, in particular, functionally identified genes, will be compared to publicly available gene expression data to validate my findings.

## Chapter 2

### Materials and Methods

#### 2.1 Mammalian Cell Culture

Human breast adenocarcinoma (MDA-MB-231) and human embryonic kidney (HEK293T) cells were obtained from American Type Culture Collection (ATCC). MDA-MB-231 cells were cultured in standard medium (Eagle's Minimum Essential Medium (EMEM) (Sigma Aldrich, M2279), with supplements), unless otherwise stated (**Table 2.1A**). For glycolysis-sensitive experiments, cells were cultured in high glucose Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F-12) (Gibco™), with supplements (**Table 2.1B**). The appropriate medium is noted in-text. HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma Aldrich, D5671), with supplements (**Table 2.1C**). All media were filter-sterilised prior to use. Cells were maintained at 37°C and 5% CO<sub>2</sub> in a Sanyo incubator with a humidified environment. For doxycycline-inducible *in vitro* models, the culture medium was supplemented with doxycycline (Sigma Aldrich, D9891) at a working concentration of 1µg per ml.

All cells were cultured in either culture flasks (Corning) or culture plates (Corning) of varying surface areas. Cell seeding densities are stated where appropriate. All cells were cultured to approximately 70% confluency. Whenever cells were not confluent, they were washed every 2 days in phosphate buffered saline (PBS) (Sigma Aldrich, D8537) prior to addition of fresh medium. Cells at 70% confluency were washed in PBS and covered with 1ml trypsin (Thermo Fisher, SV30031.01) per 75cm<sup>2</sup>. Cells were incubated for up to 2 minutes in the incubator until they detached, followed by addition of serum-containing medium. The cells were pelleted by centrifugation at 290 x g for 5 minutes. For further propagation, cells were re-seeded at a ratio of 1:4. Trypsinised cells were assessed for viability by staining with the trypan blue (Sigma Aldrich, T8154) azo dye.

<b>A) Standard MDA-MB-231 culture medium</b>	
<b>Component</b>	<b>Concentration</b>
EMEM (Sigma Aldrich, M2279)	-
Foetal Bovine Serum (Labtech, FB1090/500)	10% v/v
L-glutamine (Thermo Fisher, SH30034.01)	1x
Vitamin mix (Thermo Fisher, 10453305)	1x
Non-essential amino acids (Thermo Fisher, 11140-050)	1x
Sodium pyruvate (Thermo Fisher, 11360-039)	1x
Penicillin/streptomycin (HyClone, SV0079.01)	1x
<b>B) High glucose MDA-MB-231 culture medium</b>	
<b>Component</b>	<b>Concentration</b>
DMEM/F-12 (Gibco™; 11320033)	-
Foetal Bovine Serum (Labtech, FB1090/500)	10% v/v
Penicillin/streptomycin/fungiezone (HyClone, SV0079.01)	1x
<b>C) HEK293T culture medium</b>	
<b>Component</b>	<b>Concentration</b>
DMEM (Sigma Aldrich, D5671)	-
Foetal Bovine Serum (Labtech, FB1090/500)	10% v/v
Penicillin/streptomycin/fungiezone (Thermo Fisher, 15410-122)	1x

**Table 2.1. Culture medium constituents for MDA-MB-231 and HEK293T cells**

Cells which were permeable to the dye, and visualised as blue, were deemed as non-viable. For cell counting experiments, cells were stained with trypan blue and visualised in a BRAND® counting chamber (Sigma Aldrich, BR19605). Cell counts were taken for each corner quadrant of the counting chamber, before taking a mean across the values. Counts were adjusted relative to sample dilution.

For long-term cryopreservation,  $1 \times 10^6$  cells were pelleted as described. Following removal of supernatant, cells were re-suspended in 700µl basal medium and 200µl foetal bovine serum (FBS) (Labtech). Cells were pre-cooled on ice for 5 minutes before addition of 100µl dimethyl sulfoxide (DMSO) (Sigma Aldrich, D2650). Vials of cells were frozen at  $-80^{\circ}\text{C}$  for 24 hours prior to transferral to long-term cryopreservation at  $-196^{\circ}\text{C}$  in liquid nitrogen.

### **2.1.1 2-Deoxy-D-Glucose**

Powdered 2-deoxy-D-glucose (2DG) (Sigma Aldrich, D8375) was reconstituted in DMEM/F-12 basal medium. MDA-MB-231 cells were cultured for at least 1 week in high glucose Dulbecco's Modified Eagle Medium, with 10% FBS, prior to glycolysis inhibition studies. High glucose medium was used throughout all glycolysis experiments. At 24 hours before addition of 2DG, cells were seeded into 6- or 12-well plastic tissue culture plates (Corning) and cultured for 24 hours to allow the cells to adhere. All cells in 2DG-based studies were washed daily, followed by the addition of fresh medium supplemented with 2DG at the indicated concentration. For 2DG titration studies, spent medium was centrifuged at  $290 \times g$  for 5 minutes, prior to discarding. Fresh medium with 2DG was used to re-suspend the cell pellet, consisting of floating cells, which was added back to the corresponding well. Where required, cells were split at approximately 80% confluency and re-seeded at 1:4 dilution. The dilution factor was applied to downstream counts of floating and adherent cells.

### **2.1.2 Recombinant Human Biglycan**

Lyophilised recombinant human biglycan (rhBGN) (R&D Systems, 2667-CM-050) was reconstituted in DMEM/F-12 basal medium. All cells were cultured on plastic 12-well tissue culture plates (Corning). For the coating of wells with rhBGN, 100µl rhBGN was added to each well at a concentration of 5µg per ml, followed by incubation at 37°C for 8 hours. After 8 hours, rhBGN was aspirated and wells were washed 3 times in 300µl PBS for 5 minutes each, at room temperature. MDA-MB-231 cells were added to rhBGN-coated and non-coated wells at  $1 \times 10^5$  cells per well. Where required, soluble biglycan was supplemented to the medium at 5µg per ml. All cells were cultured in high glucose Dulbecco's Modified Eagle Medium, with 10% FBS. Cells were cultured for 72h prior to counting, without replenishment of media.

## **2.2 Cell Membrane Staining**

In some experiments MDA-MB-231 cells were stained using either the CellVue Claret® (CV) Far Red Fluorescent Linker Kit (Sigma Aldrich, USA), or Vybrant™ Dil Cell-Labeling Solution (Dil) (Thermo Fisher; V22885).

### **2.2.1 Staining of MDA-MB-231 Cells with CellVue® Claret Membrane Label**

Adherent MDA-MB-231 cells were washed twice with EMEM and covered with approximately 1ml Diluent C per 75 cm<sup>2</sup>. A 2x Dye Solution was prepared according to manufacturer's instructions and mixed by vortexing. The 2x stock was added to the Diluent C covering the cells in a 1:1 ratio, and distributed evenly. Cells were incubated at 37°C for 5 minutes to incorporate the dye. Following incubation, FBS was added to the staining solution at a 1:1 ratio for 1 minute at ambient temperature (AT). CV-labelled cells were washed 3 times for 10 minutes each at 37°C in standard serum-containing medium.

### 2.2.2 Staining of MDA-MB-231 Cells with Dil Long-Term Tracer

Adherent MDA-MB-231 cells were washed 1x in PBS. For staining of cells with Dil, staining medium was prepared according to manufacturer's instructions. Cells were covered in staining medium and incubated at 37°C for 8 minutes. Cells were washed according to manufacturer's instructions before addition of complete growth medium.

### 2.3 *In Vivo* Experiments

All procedures were approved by the University of Leeds Animal Welfare & Ethical Review Committee (AWERC) and performed under the approved UK Home Office project licence in line with the Animal (Scientific Procedures) Act 1986 and in accordance with the UK National Cancer Research Institute Guidelines for the welfare of animals [233].

The mouse strain used for *in vivo* experiments were C.B.17 severe combined immunodeficiency (SCID) mice (C.B-*Igh-1<sup>b</sup>/IcrTac-Prkdc<sup>scid</sup>*). All mice were female and aged between 6 and 8 weeks at the start of each experiment. All mice were bred and caged at the St James' Biological Services (SBS) Facility, in individually ventilated cages at a maximum of 5 mice per cage.

#### 2.3.1 Preparation of Cancer Cells for Surgery

For intra-carotid injections, cultured CV-labelled MDA-MB-231 cells were washed with PBS, trypsinised and gently pipetted up and down to dissociate any cell clumps. Cells were transferred to standard medium (containing serum) and strained through a 70-100µm cell strainer (Greiner Bio-One, 542-070). Cells were counted before centrifugation at 290 x *g* for 5 minutes. The supernatant was discarded, and the pellet

was re-suspended in EMEM without supplements to wash cells, prior to centrifugation at  $290 \times g$ . The supernatant was discarded, and cells were re-suspended in an appropriate volume of EMEM without supplements, to a density of  $1 \times 10^5$  cells per ml.

### **2.3.2 Generation of Mouse Brain Tumour Xenografts**

Brain tumour xenografts were generated by injection of  $1 \times 10^5$  CV-labelled MDA-MB-231 cells into the internal carotid artery of 6-8 weeks old female C.B.17-SCID mice. Procedures involving injection into the internal carotid artery were kindly carried out by Mihaela Lorgier (University of Leeds, UK) as previously published [134]. At 28 days post-injection of cancer cells, the mice were euthanised by terminal perfusion. Upon observation of adverse symptoms prior to day 28, mice were euthanised by schedule I culling and excluded from all downstream analyses.

### **2.3.3 Terminal Perfusion and Mouse Brain Isolation**

For organ harvesting, mice were anaesthetised with 5% isoflurane, followed by intra-peritoneal injection of 300 $\mu$ l Lethobarb/pentobarbitone (Merial). Upon cessation of respiration, the thoracic cavity was cut open and an incision was made to the heart, distal to the apex. Terminal intra-cardiac perfusion was performed by flushing through the circulatory system with 15ml PBS. For brains destined for cell sorting or processing as frozen sections, the cranium was opened with scissors, followed by gentle removal of the brain and placing on ice or in 4% paraformaldehyde (PFA) (Sigma Aldrich, P6148), respectively. For PFA fixation, tissue was submerged in PFA at 4°C for 24 hours, followed by transfer into phosphate buffered sucrose solution (**Table 2.2**) at 4°C until the tissue sunk to the bottom of the vessel. The brains were then frozen and stored at -80°C, or sectioned into floating sections for immunofluorescence (see Chapter 2.7).

Component	Concentration
Sucrose (Sigma Aldrich; 50389)	25% w/v (730.361mM)
NaH <sub>2</sub> PO <sub>4</sub> (Sigma Aldrich; 71505)	23mM
Na <sub>2</sub> HPO <sub>4</sub> (Alfa Aesar; A11817)	77mM
ddH <sub>2</sub> O (Veolia, PURELAB Option, OQ007XXM1)	-

**Table 2.2. Constituents of phosphate buffered sucrose solution.**

### 2.3.4 Isolation of Cancer Cells from Mouse Brain

The right hemisphere and brain stem were removed with scalpels prior to processing. The remaining brain (left hemisphere, expected location of xenograft lesions) was chopped finely by scalpel and dissociated with collagenase/hyaluronidase solution (**Table 2.3**), for 2 x 10 minutes at 37°C. The volume was increased to 20-30ml with pre-chilled incubation buffer (filter-sterilised PBS (Sigma Aldrich, D8537), 0.5% w/v bovine serum albumin (BSA) (Sigma Aldrich, A4503) and 2mM ethylenediaminetetraacetic acid (EDTA) (Fisher Scientific, BP2482)) and strained with a 70-100µm cell strainer (Greiner Bio-One, 542-070). Samples were pelleted by centrifugation at 290 x g for 5 minutes at 4°C. Myelin was removed from samples using Myelin Removal Beads II (Miltenyi Biotec; 130-096-733), according to manufacturer's instructions. Two rounds of myelin removal were employed (200 and 100µl myelin removal beads per mouse, respectively), using 3 MAC LS Removal Columns (Miltenyi Biotec, 130-042-401) in round 1, and 2 MAC LS Removal Columns in round 2.

Component	Concentration
Eagle's Minimum Essential Medium (Sigma Aldrich, M2279)	-
Collagenase (Roche, 1088793)	3mg per ml
Hyaluronidase (Sigma Aldrich, H-3506)	250 U per ml
Penicillin/streptomycin/fungiezone (Thermo Fisher, 15140-22)	1X

**Table 2.3. Components of collagenase/hyaluronidase solution for mouse brain dissociation.**

## 2.4 Flow Cytometry

### 2.4.1 Flow Cytometry Analysis

Flow cytometry analysis was carried out using the CytoFLEX Flow Cytometer with accompanying CytoFLEX CytExpert Software, V1.2.11.0 (Beckman Coulter). Where antibody staining was not required, cells were pelleted (see *Chapter 2.1 Mammalian Cell Culture*), washed in PBS and re-suspended in further PBS at  $1 \times 10^6$  cells per ml in polystyrene tubes (SLS, 352054). Samples were then strained through a 70-100 $\mu$ m cell strainer (Greiner Bio-One, 542-070). Individual cells were analysed according to their relative size and fluorescence intensity. Laser and fluorescent channel wavelength data is shown in **Table 2.4**. The number of events analysed was 10,000-30,000 at a rate of 10 $\mu$ l per minute. For the flow cytometry protocol involved in cell cycle analysis, see Chapter 2.4.3.

Marker	Laser (nm)	Fluorescent Channel
Green fluorescent protein (GFP)	488	525/40 BP
CellVue Claret® (CV)	638	660/20 BP
Propidium iodide (PI)	488	690/50 BP
Allophycocyanin (APC)	638	660/20 BP

**Table 2.4. Laser and fluorescent channel configurations for flow cytometry (CytoFlex)**

#### **2.4.2 Fluorescence Activated Cell Sorting (FACS) – Isolation of Dormant and Proliferating MDA-MB-231 Cells from Mouse Brain Tumour Xenografts**

Proliferating and dormant cancer cell populations were sorted from the suspension of dissociated metastases-bearing brains using the BD inFlux v7 Sorter (BD Biosciences). The cells were separated according to relative size and fluorescence intensity, using a 488nm excitation laser with 530/40nm emission filter, and 655nm excitation laser with 670/30nm emission filter, for detection of GFP and CV dye, respectively. Cells were sorted directly into 100µl extraction buffer from the ARCTURUS® PicoPure™ RNA Isolation Kit (Thermo Fisher, KIT0204). Fluorescence channel data was acquired in log scale, for  $1 \times 10^7$  cells using BD FACS™ Software v1.0.0.641 (BD Biosciences). Analysis of sort data was performed using FlowJo v10.2 (Tree Star Inc.) software. The percentage of dormant cells within mouse brain tumour xenografts was calculated as follows:

$$\frac{\text{Total number of dormant cells}}{\text{Total number of all cells (dormant + proliferating)}} \times 100$$

### 2.4.3 Cell Cycle Analysis

Prior to harvesting of adherent cells, 5-bromo-2'-deoxyuridine (BrdU) (Sigma Aldrich, B5002) was added directly to culture medium at a working concentration of 10 $\mu$ M. Cells were incubated under standard culture conditions for 30 minutes. Cells were collected by trypsinisation (see Chapter 2.1) into a round-bottomed polystyrene tube (SLS, 352054) and washed in PBS. All centrifugation steps were carried out at 290  $\times$  g for 5 minutes. Cells were re-suspended in 100 $\mu$ l of cold PBS. Whilst vortexing, 900 $\mu$ l of cold 70% v/v ethanol (Merck, 32221) was added dropwise. Ethanol fixation was carried out on ice for 30 minutes. Cells were pelleted by centrifugation and re-suspended in 1ml of denaturation buffer (**Table 2.5A**), with incubation for 30 minutes at AT. Cells were centrifuged, denaturation buffer was discarded and cells were re-suspended and incubated in 1ml of neutralisation buffer (**Table 2.5B**) for a further 30 minutes at AT. Cells were then pelleted by centrifugation, neutralisation buffer was discarded. Cells were re-suspended in 50 $\mu$ l antibody solution (**Table 2.5C**) containing 1:9 dilution of anti-BrdU-APC antibody (**Table 2.9**) and incubated for 1 hour at AT. Cells were washed 3 times in PBS and re-suspended in 500 $\mu$ l PBS containing 5 $\mu$ g per ml RNase A (Qiagen, 1031301) and 10 $\mu$ g per ml propidium iodide (PI) (Sigma Aldrich, P4864), before passing through a 70 $\mu$ m cell strainer (Greiner Bio-One, 542-070). After 30 minutes incubation at AT, cells were analysed by flow cytometry (see Chapter 2.4.1).

<b>A) – Denaturation solution</b>	
<b>Component</b>	<b>Concentration</b>
Hydrochloric Acid	2M
Triton X-100	0.5% v/v
PBS	-
<b>B) Neutralisation solution</b>	
<b>Component</b>	<b>Concentration</b>
Na <sub>2</sub> B <sub>4</sub> O <sub>7</sub> ·10H <sub>2</sub> O (pH 8.5)	0.1M
PBS	-
<b>C) Antibody blocking solution</b>	
<b>Component</b>	<b>Concentration</b>
Bovine serum albumin	1% w/v
Tween 20	0.5% v/v
PBS	-

**Table 2.5. Solution constituents for cell cycle FACS analysis**

## **2.5 RNA Isolation and Purification**

### **2.5.1 RNA Isolation from Cells Isolated by FACS**

RNA was isolated according to the manufacturer's protocol for the isolation of RNA from cell pellets (ARCTURUS® PicoPure™ RNA Isolation Kit, Thermo Fisher, KIT0204). As described in *Chapter 2.4.2*, cells were sorted directly into extraction buffer and immediately lysed. All samples were treated with DNase using the RNase-free DNase Set (Qiagen, 79254), according to manufacturer's instructions. Samples were eluted in 11µl manufacturer-supplied elution buffer. Quality control analysis was performed using the

RNA 6000 Pico Kit (Agilent Technologies, 5067-1513) and 2100 Bioanalyser Instrumentation (Agilent Technologies), according to manufacturer's instructions.

### **2.5.2 RNA Isolation from Adherent Cells**

Adherent cells were trypsinised and washed in PBS. All samples were processed using the RNeasy Mini Kit (Qiagen, 74104), according to manufacturer's instructions. RNA was eluted in 50µl RNase-free water. RNA concentration quantification was determined using the Nanodrop (Labtech, ND1000), according to manufacturer's instructions.

## **2.6 cDNA Synthesis, Library Construction and mRNAseq**

We utilised Smart-seq2 [224] and Nextera XT DNA Library Prep (Illumina, FC-131-1024) protocols for the generation of full-length cDNA and sequencing libraries. The starting material was 114pg of total RNA in 3µl nuclease-free water. Oligo-dT primer (Invitrogen, 58063) and dNTP mix (Thermo Fisher, AB0196) was directly added to RNA in a thin-walled 0.2ml PCR tube (STAR LAB, I1402-8100). Initial PCR for cDNA synthesis was run for 15 cycles using a Veriti 96-Well Thermal Cycler (Thermo Fisher, 4375786). The quantity of cDNA was determined with the Qubit dsDNA High Sensitivity Assay Kit using a Qubit Fluoremeter (Thermo Fisher, Q33238). We carried forward 5µl of cDNA (0.2ng per µl) for the tagmentation reaction. The quality of the final cDNA library was checked using the 4200 TapeStation System (Agilent). Library pooling was performed by the Next Generation Sequencing (NGS) Facility (University of Leeds), which also ran the paired-end mRNA sequencing using the Hiseq3000 platform (Illumina).

### **2.6.1 mRNAseq Data Processing and Analysis**

Raw sequencing data was provided in fastq format. This work was in part undertaken on Medical Advanced Research Computer 1 (MARC1), part of the High Performance Computing and Leeds Institute for Data Analytics (LIDA) facilities at the University of Leeds, UK. This work was in part also undertaken on Advanced Research Computer 3 (ARC3), part of the High Performance Computing facilities at the University of Leeds, UK.

All custom-designed relevant scripts involved in mRNAseq data processing are displayed in the appendix. The R script for concatenating human and mouse reference genomes is shown in **Supplementary Figure S1**. The R script for the generation of read counts (processing stream II) is shown in **Supplementary Figure S2**. All other scripts employed in this study were constructed according to the published manual and instructions of each referenced software/package. Data in this study was processed via two independent analysis streams. The first (stream I) was performed as a service by the NGS Facility, University of Leeds, with alignment to human and mouse genomes consecutively. The second (stream II) was performed independently, with alignment to a concatenated human and mouse genome and using an alternative source of reference genomes and splice junction annotations. Supervision of stream II bioinformatics analyses was undertaken by Dr Alastair Droop, formerly of the University of Leeds.

### 2.6.2 Alignment to Reference Genomes

For both processing streams, assessment of significant trends in sequence quality and diversity within Fastq files was carried out using FastQC, version 0.11.8 [225]. Trimmomatic read trimming software version 0.39 [226] was employed to trim poor quality bases of Phred score < 20, as were sequences pertaining to primers, poly-A regions and sequences homologous to Nextera Transposase Adaptors. For processing stream II, this was carried out using Cutadapt software [227] under the same parameters. For processing stream I, genome annotation data was downloaded from the publicly available University of California Santa Cruz (UCSC) genome browser [228] for human (GRCh38/hg38) and mouse (GRCm38/mm10) genome reference sequences (**Table 2.6**). For processing stream II, primary assemblies (fasta) and comprehensive gene annotations for human, release 31 (GRCh38.p12), and mouse, release M22 (GRCm38.p6), genomes were downloaded from the GENCODE project website [229]. Concatenation of reference assemblies and renaming of human and mouse chromosomes for downstream distinction was performed using Bioconductor R package Biostrings [230]. Generation of genome indexes and alignment of sample data to the generated indexes was performed using STAR version 2.7 [231]. For alignment, STAR was run in 2-pass mode. Binary Alignment Map (BAM) files of aligned data were sorted

and indexed using SAMtools software [232], exporting to the Sequence Alignment Map (SAM) format. Quality control of BAM files was carried out using QualiMap [233] and Picard Tools [234] softwares. Aligned sequencing data was visualised using Integrative Genome Viewer (IGV) software [235]. Consensus sequences for genes of interest and identification of read-specific polymorphisms was determined using IGV.

<b>Assembly</b>	December 2013 (GRCh38/hg38)	December 2011 (GRCm38/mm10)
<b>Group</b>	Genes and Gene Predictions	Genes and Gene Predictions
<b>Track</b>	NCBI RefSeq	NCBI RefSeq
<b>Region</b>	Genome	Genome
<b>Table</b>	RefSeq All (ncbiRefSeq)	RefSeq All (ncbiRefSeq)

**Table 2.6. Download parameters for human (GRCh38/hg38) and mouse (GRCm38/mm10) reference genomes (NCBI RefSeq), processing stream I.**

### 2.6.3 Read Counting

For processing stream I, assembly of transcriptomes and quantification of gene expression was carried out using Cufflinks software [236]. Merging of data into a single transcriptome was carried out using Cuffmerge software [236], on Cufflinks-generated data. Quantified reads mapping to each transcript was obtained using Cuffdiff software [236]. For processing stream II, read quantification, transcriptome merging and mapping of counts to transcript was performed using the *featureCounts* function of the R package RSubread [237]. Multi-mapping reads were included.

### 2.6.4 Differential Gene Expression Analysis

Prior to differential gene expression analysis between dormant (GFP+CV-) and proliferating (GFP+CV-) cells, the total number of reads mapping to human and mouse

genomic positions for each sample were quantified. This was given as the total number of raw reads across all genes for each species, within each sample. Any reads mapping to the mouse genome were discarded as the cancer cells of interest were of human origin. Read counts were converted to integer values for compatibility with R package DESeq2 [238]. Size factor normalisation was performed using DESeq2. Normalised gene expression data was exported as a data matrix, with expression values and fold changes  $\log_2$ -transformed by DESeq2. Heat maps (generated by unsupervised hierarchical clustering) and Principle Component Analysis (PCA) plots were generated using ClustVis, an open-sourced web tool, using pre-filtered gene expression datasets [239]. For the heat maps, colour scaling was performed at the level of each transcript. Where specific groups of genes were of interest, data matrices of normalised read count data annotated with a gene were used as input. Ensembl and Refseq gene IDs were converted to gene symbols using the Ensembl Gene ID Converter, part of the online Bioinformatics Tools for genomics and transcriptomics analyses [240]. Visualisation of gene expression data, given as the mean of normalised counts against the log of fold change, was performed using the plotMA function of DESeq2 to generate an MA plot. For both data processing streams, genes converged on  $y = 0$ . Therefore, normalisation was applied correctly as the majority of genes were not differentially expressed and that any genes marked as differentially expressed was likely to be a true representation. No further normalisation was required.

### 2.6.5 Calculation of Allele Frequencies

Allele frequencies were counted using custom Java code (Next Generation Sequencing Facility, University of Leeds) by querying alignments from BAM files at known SNP positions. The complete list of known MDA-MB-231 cell line mutation variants were sourced from the Catalogue of Somatic Mutations In Cancer, COSMIC [241], for genome version GRCh38. Java htsjdk library was used for BAM file random access [232]. Allele frequencies were then imported into R and pairwise Pearson correlations were computed using base R.

### 2.6.6 GO and KEGG Enrichment Analysis

The Benjamini-Hotchberg method of statistical significance adjustment (False Discovery Rate (FDR)) was employed by DESeq2 to account for the effects of multiple testing. Differentially expressed genes were identified under a number of parameters to increase and decrease the gene pool for varying levels of gene coverage for downstream analyses. These were  $p \leq 0.05$  and  $FDR \leq 0.05$ , combined with or without a gene expression fold change of  $> 2$ . Identified differentially expressed genes were used as input to identify functionally enriched groups of genes and pathways. Enriched Gene Ontology (GO) terms for biological processes, molecular functions and cellular components were identified using the Bioconductor R package EnrichGO function of the R package, clusterProfiler [242]. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis within specified gene sets was carried out using the enrichKEGG function of the R package, clusterProfiler. Statistically significant analyses were visualised using the dotplot function of the R package, clusterProlifer. Enriched pathways identified by KEGG analysis were visualised using the pathview function of the R package ReactomePA [243].

### 2.6.7 Transcription Factor, Protein Network and Cell Cycle Analyses

Functionally implicated transcription factors, predicted as active or repressed based on an input list of differentially expressed genes, was predicted using TFactS software [244]. Functional protein annotation networks for single or groups of selected genes was visualised using STRING [245]. Cell cycle phase was predicted using an input of differentially expressed genes for each sample into R package, Scran [246]. The R script for cell cycle phase analysis of mRNAseq data is shown in **Supplementary Figure S3**.

### 2.6.8 Analysis of Publicly Available Gene Expression Datasets

Publicly available mRNAseq data [247] was employed for the comparison of patient-matched primary breast cancers and breast cancer brain metastases. *BGN* expression was first normalised to that of a suitable housekeeping gene, RNA polymerase II subunit

A (*POLR2A*). For the analysis of gene expression in merged gene expression datasets, data from each study (Affymetrix platform) was first independently normalised using R package *affy* [248]. The R script for the normalisation of microarray data, using CEL file inputs, is shown in **Supplementary Figure S4**. All genes were then normalised to an appropriate housekeeping gene, *POLR2A*. All datasets were then merged prior to the calculation of z scores. The data accession numbers for each study were as follows: GSE2034, GSE5327, GSE12276, GSE14017, GSE14018, GSE43837 [249, 250, 251, 252, 253, 254]. The analyses of the correlation of biglycan expression with distant metastasis-free survival (DMFS) and relapse-free survival (RFS) was performed by Kaplan Meier plotter [255]. The Affy ID for the *BGN* gene was 201261\_x\_at. Cut-offs for gene expression quartiles, used in the determination of high and low gene expression, was calculated automatically by the software.

## 2.7 Immunofluorescence

For frozen sections, PFA-fixed tissue was cut into 40µm floating sections using the CM3050 S Cryostat (Leica) and transferred into Walter's antifreeze (**Table 2.7**). Sections were stored at -20°C. Sections were transferred from Walter's antifreeze into 700µl PBS in a 12-well tissue culture plate (Corning) and incubated with gentle shaking at AT for 5 minutes for washing. The PBS was discarded and sections were washed a further 2 times. Antibody blocking buffer of 500µl volume (**Table 2.8**) was added to sections and incubated for 1 hour at AT with gentle shaking. All sections were first incubated with a primary antibody, diluted in 200µl blocking buffer, followed by incubation with a fluorophore-conjugated secondary antibody, diluted in 200µl blocking buffer. Details of each antibody with optimal concentrations and conditions for immunofluorescence are described in **Table 2.9**. Sections were washed 3 times for 5 minutes each, in 700µl PBS, at AT, with gentle rocking between antibodies. Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) (Thermo Fisher, D1306), diluted 1:5000 in 700µl PBS for 10 minutes at AT, with gentle shaking. Sections were again washed 2 times in 700µl PBS, before transferring to glass slides (Thermo Fisher, J1800AMNZ) and mounted using ProLong Gold Antifade Mountant (Invitrogen, P36930), with glass coverslips (SLS, MIC3246). Images were acquired using AxioCam MRm (Zeiss) and AxioVision software

(Zeiss), or with the A1R confocal microscope and Confocal NIS-Elements software (Nikon).

For adherent/cultured cells, Growth medium was aspirated from 12-well tissue culture plates and cells washed in 700µl PBS. Cells were fixed in 300µl 4% w/v PFA for 10 minutes at room temperature. Cells were then washed 3 times in 700µl PBS, for 5 minutes, at room temperature. The staining protocol is the same as described in Chapter 2.7.2, without gentle shaking. After the final wash, ProLong Gold Antifade Mountant was added directly to the cells in the well, followed by the addition of a glass coverslip. Images were acquired using the EVOS M5000 Imaging System (Thermo Fisher).

Component	Concentration
NaH <sub>2</sub> PO <sub>4</sub> (BDH Laboratory Supplies, 102454R)	0.157% v/v
Na <sub>2</sub> HPO <sub>4</sub> .2H <sub>2</sub> O (BDH Laboratory Supplies, 102494C)	0.82% v/v
ddH <sub>2</sub> O (Veolia, PURELAB Option, OQ007XXM1)	40% v/v
Ethylene glycol (Acros Organics, 146750010)	30% v/v
Glycerol (Sigma Aldrich, G5516)	30% v/v

**Table 2.7. Components of Walter's antifreeze for storage of PFA-fixed floating frozen sections**

Component	Concentration
PBS (Sigma Aldrich, D8537)	-
Goat serum (Thermo Fisher, 16210064)	10% v/v
Triton™ X-100 (Sigma Aldrich, P1379)	0.3% v/v

**Table 2.8. Components of antibody blocking buffer for immunofluorescence**

Antibody	Species	Dilution	Conditions	Vendor
<b>α-biglycan</b>	Rabbit	1 in 200 (IF) 1 in 800 (WB)	1 hour at AT 1.5 hours at AT	Proteintech (16409-1-AP)
<b>α-GFP</b>	Chicken	1 in 1000 (IF)	1 hour at AT	Abcam (Ab13970)
<b>α-chicken-FITC</b>	Donkey	1 in 200 (IF)	1 hour at AT	Jl (703-096-155)
<b>α-rabbit-TRITC</b>	Donkey	1 in 200 (IF)	1 hour at AT	Jl (711-025-152)
<b>α-BrdU-APC</b>	Mouse	1 in 10 (5μl per test)	1 hour at AT	eBioscience (17-5071-42)
<b>α-Yap1</b>	Mouse	1 in 50 (IF) 1 in 1000 (WB)	1 hour at AT O/N at 4°C	SCB (sc- 101199)
<b>α-Yap1</b>	Rabbit	1 in 50 (IF)	1 hour at AT	NB (NB110- 58358)
<b>α-beta-catenin</b>	Rabbit	1 in 1000 (WB) 1 in 50 (IF)	1.5 hours at AT 1 hour at AT	GeneTex (GTX61089)
<b>α-mouse-HRP</b>	Horse	1 in 5000 (WB)	1 hour at AT	CST (7076P2)
<b>α-rabbit-HRP</b>	Goat	1 in 5000 (WB)	1 hour at AT	CST (7074P2)
<b>α-phospho-Yap1 (Ser127)</b>	Rabbit	1 in 50 (IF) 1 in 1000 (WB)	O/N at 4°C 1 hour at AT	Fisher Scientific™ (PA5-17481)
<b>α-human Vimentin</b>	Mouse	1 in 1000 (WB)	1 hour at AT	DAKO (M0725)
<b>α-alpha-tubulin</b>	Mouse	1 in 1000 (WB)	O/N at 4°C	GeneTex (GT114)

**Table 2.9. Summary table of antibodies.** Abbreviations: green fluorescent protein (GFP); immunofluorescence (IF); Western blot (WB); room temperature (AT); tetramethylrhodamine (TRITC), fluorescein isothiocyanate (FITC); horseradish peroxidase (HRP); allophycocyanin (APC); Jackson Immunoresearch (Jl); Santa Cruz Biotechnology (SCB), Novus Biologicals (NB), Cell Signalling Technology (CST).

## 2.8 Quantitative Polymerase Chain Reaction (qPCR)

For cDNA synthesis, we utilised the Superscript III cDNA Synthesis Kit (Invitrogen, 18080-044), except in cases of low RNA input in which we employed Smart-seq2 protocol [224]. Up to 1µg of RNA was used as input for cDNA synthesis. Mastermix I (**Table 2.10**) was assembled in polymerase chain reaction (PCR) tubes (STAR LAB, I1402-8100) and heated to 65°C for 5 minutes, followed by placing on ice for 2 minutes.

Constituent	Amount
Oligo dT (Thermo Fisher, 58063)	1µl
dNTPs (20mM) (Thermo Fisher, AB0196)	1µl
ddH <sub>2</sub> O (Sigma Aldrich, RNBF2734)	Variable (Total reaction volume = 7µl)
RNA (1µg)	Variable

**Table 2.10. Mastermix I components for cDNA synthesis**

Constituent	Amount
5X First Strand Buffer (Thermo Fisher, Y02321)	4µl
0.1M DTT (Thermo Fisher, Y00147)	1µl
RNase OUT (Thermo Fisher, 100000840)	1µl
Superscript III (Thermo Fisher, 56575)	1µl

**Table 2.11. Mastermix II components for cDNA synthesis**

To each sample, 7µl of mastermix II (**Table 2.11**) was added and mixed. Samples were incubated at 50°C for 60 minutes and 70°C for 15 minutes to generate cDNA. Where possible, qPCR reactions were performed in biological triplicate, with three technical

replicates for each individual sample. To each well of a 96-well qPCR plate (Applied Biosystems, N8010560), 7.5µl of Taqman® Universal Mastermix II (Applied Biosystems, 4440040), 0.75µl of Taqman® qPCR probe (**Table 2.12**) and 3.75µl ddH<sub>2</sub>O (Sigma Aldrich, RNBF2734). To this, 3µl of template cDNA was added. Where 384-well qPCR plates were employed (Thermo Fisher, 4309849), all reaction volumes were scaled down by a factor of 1/3. All qPCR reactions were performed for 40 cycles using a comparative  $\Delta\Delta\text{CT}$  assay. The instruments used were the QuantStudio™ 5 Real-Time PCR System (Thermo Fisher) and QuantStudio™ 7 Flex Real-Time PCR System (Applied Biosystems). Post-qPCR analysis was performed using Excel (Microsoft).

Probe (gene)	Target Species	Company (Cat. No./Assay ID)
Biglycan ( <i>BGN</i> )	Human	Thermo Fisher (4331182/Hs00959143_m1)
Collagen I ( <i>COL1A1</i> )	Human	Thermo Fisher (4331182/Hs00164004_m1)
RNA polymerase II subunit A ( <i>POLR2A</i> )	Human	Thermo Fisher (4331182/Hs00172187_m1)
Phosphoglycerate kinase 1 ( <i>PGK1</i> )	Human	Thermo Fisher (4331182/Hs00943178_g1)
Glyceraldehyde-3-phosphate dehydrogenase ( <i>GAPDH</i> )	Human	Thermo Fisher (4331182/Hs02786624_g1)
L1 cell adhesion molecule ( <i>L1CAM</i> )	Human	Thermo Fisher (4331182/Hs01109748_m1)
Naked cuticle homolog 1 ( <i>NKD1</i> )	Human	Thermo Fisher (4331182/Hs00263894_m1)
Axis inhibition protein 2 ( <i>AXIN2</i> )	Human	Thermo Fisher (4331182/Hs00610344_m1)

**Table 2.12. Summary table of Taqman® probes.**

## 2.9 Protein Extraction and Quantification

For adherent cells, culture medium was aspirated, and cells were washed 1x in PBS. An appropriate volume of radioimmunoprecipitation assay (RIPA) buffer (**Table 2.13**) was added to the cells, followed by pipetting up and down 10x, avoiding any bubbles. Cells were incubated on ice for 10 minutes. The lysed cells were placed in a fresh tube and centrifuged at 13,000 *g* for 10 minutes. The protein-containing supernatant was collected. This was the total protein fraction.

For nuclear fractionation experiments, adherent cells were trypsinised and centrifuged at 290 x *g* for 5 minutes at AT. As fractionation requires the use of a cell suspension, the trypsinisation step was kept to a minimum to limit potential proteolytic damage to the samples. The supernatant was removed and cells were washed in PBS, followed by centrifugation at 290 x *g* for a further 5 minutes at AT. The supernatant was removed, and cells were re-suspended in hypotonic buffer (**Table 2.14**). Cells were left to swell on ice for 5 minutes.

Cells were then passed through a 27-gauge needle (company) 10 times or until cells were lysed, followed by incubation on ice for 20 minutes. Samples were centrifuged at 500-720 *g* for 5 minutes at AT to pellet the nuclear fraction. The supernatant containing the cytoplasmic fraction was collected. Hypotonic buffer was added to the nuclear pellet for washing. The pellet was dispersed by pipetting and passing through a 25-gauge needle (company) 20 times. The samples were then centrifuged at 500-720 *g* for 10 minutes at AT. The pellet contains nuclei. The supernatant was discarded. To lyse the nuclear fraction, the pellet was re-suspended in RIPA buffer (**Table 2.13**) and pipetted up and down 10 times. The samples were incubated on ice for 10 minutes, followed by centrifugation at 13,000 *g*. The supernatant contains nuclear protein. Protein quantification for each sample was determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher; 23225), according to manufacturer's instructions. Colorimetry was performed using the Multiskan EX (Thermo Fisher). Absorbance readings from bovine serum albumin (BSA) standards were used to generate a standard curve. Absorbance readings for each sample were used to interpolate concentrations from the standard

curve using Excel (Microsoft). Prior to Western blot, all samples were standardised to the lowest concentration across the cohort, in an appropriate volume of RIPA buffer.

Component	Concentration
ddH <sub>2</sub> O (Veolia, PURELAB Option, OQ007XXM1)	-
Tris pH 7.5 (Merck, 348317)	20mM
Sodium chloride (NaCl) (VWR, 28710.364)	100mM
Triton X-100 (Sigma Aldrich, T91284)	1mM
Dithiothreitol (DTT) (Anaspec, AS-25185-5)	1% v/v
Sodium orthovanadate (Na <sub>3</sub> VO <sub>4</sub> ) (Merck, 450243)	1mM
β-glycerophosphate (Sigma Aldrich, G9422)	25mM
Complete Mini Protease Inhibitor Cocktail (Roche, 11836170001)	1 tablet per 10ml
Sodium dodecyl sulphate (SDS) (VWR, 442444H)	0.1% w/v
Sodium deoxycholate (DOC) (Sigma Aldrich, D6750)	0.5% w/v

**Table 2.13. Constituents of RIPA buffer**

Component	Concentration
Tris-HCl pH 7.4 (Fisher BioReagents, BP153-1)	10mM
EDTA (Fisher Scientific, BP2482)	0.1mM
Sucrose (Sigma Aldrich, S0389)	250mM
ddH <sub>2</sub> O (Veolia, PURELAB Option, OQ007XXM1)	-

**Table 2.14. Constituents of hypotonic buffer for cellular lysis**

### 2.10 Patient-Derived Tumour Xenografts

Tumour lysates from patient-derived tumour xenografts (PDTXs) were kindly donated by Yolanda Kartika (University of Leeds, UK). Xenografts were generated via intracranial implantation of human brain metastatic breast cancer cells into NOD SCID gamma (NSG) mice. The phenotypic characterisations of each breast cancer brain metastases (BCBM) PDTX model, BCBM2 and BCBM6, were HER2+ER-PR- and HER2+ER+PR+, respectively. BCBM2 and BCBM6 cells at the time of xenograft generation were at xenograft passage 3 and 7, respectively. BCBM2 and BCBM6 tumours were harvested at 48- and 100-days post-implantation, respectively, upon observation of terminal symptoms. Determination of tumour lysate protein concentration was carried out as described in Chapter 2.9, prior to Western blot.

### 2.11 Western Blot

Prior to loading of samples, 6X SDS loading buffer (**Table 2.15**) was added to each sample to a final concentration of 1X. All samples were run on polyacrylamide gels, constituted of a 10% resolving gel with a 4% stacking gel (**Table 2.16**) at 100V at AT, within SDS running buffer (**Table 2.17**), until the dye reached the bottom of the gel. All samples were standardised to the same protein concentration in RIPA buffer. SeeBlue™ Plus2

Pre-stained Protein Standard (Fisher Scientific; LC5925) was prepared according to manufacturer's instructions and loaded on the gel along with the samples.

Constituent	Concentration
Tris-HCl (Merck, 348317)	300mM
DTT (Anaspec, AS-25185-5)	600mM
SDS (VWR, 442444H)	12% w/v
Glycerol (Sigma Aldrich, G5516)	60% v/v
Bromophenol blue (Merck, B0126)	0.6% w/v
ddH <sub>2</sub> O (Sigma Aldrich, RNBF2734)	-

**Table 2.15. Constituents of 6X SDS loading buffer.**

Constituent	Resolving gel (10%)	Stacking gel (4%)
1M Tris, pH 8.8 (Merck, 348317)	1.875ml (pH 8.8)	0.25ml (pH 6.8)
ddH <sub>2</sub> O (Veolia, PURELAB Option, OQ007XXM1)	1.375ml	1.423ml
Protogel/30% acrylamide (Geneflow, A2-0074)	1.67ml	0.267ml
20% w/v SDS (VWR, 442444H)	25µl	10µl
10% w/v ammonium persulphate (APS) (Sigma Aldrich, A3678)	50µl	50µl
Tetramethylethylenediamine (TEMED) (Sigma Aldrich, T7204)	3µl	3µl

**Table 2.16. Polyacrylamide gel constituents (1 gel, 0.75mm thickness)**

Constituent	Concentration
Tris-HCl (Merck, 348317)	25mM
Glycine (Sigma Aldrich, G7403)	192mM
SDS (VWR, 442444H)	0.1% w/V
ddH <sub>2</sub> O (Veolia, PURELAB Option, OQ007XXM1)	-

**Table 2.17. SDS running buffer constituents**

Protein was wet-transferred from the polyacrylamide gel to a 0.45µm nitrocellulose membrane (Bio-Rad, 162-0113), for 2 hours at 360mA, at 4°C. Transfer buffer constituents are shown in **Table 2.18**. Following transfer, the membrane was washed 3 times in ddH<sub>2</sub>O, for 5 minutes each, to remove residual transfer buffer. All membranes were blocked for 1 hour at AT in blocking solution (TBST (2.42% w/v Tris-base, 8% w/w NaCl and 0.1% v/v Tween-20, in ddH<sub>2</sub>O (Sigma Aldrich, RNBF2734), pH 7.6) with 5% w/v non-fat dry milk (Marvel, 3023034)). Primary antibodies were diluted in blocking solution and incubated with the membrane at the concentrations and times indicated in **Table 2.9**. Membranes were washed 3x with TBST for 10 minutes each, before incubating with HRP-conjugated secondary antibody, in blocking solution, at the concentrations and times indicated in **Table 2.9**. Membranes were washed 3x in TBST for 10 minutes each. ECL substrate (Thermo Fisher, 32106) was prepared according to manufacturer's instructions and incubated with the membrane for 1 minute at AT. Membranes were used to expose autoradiography film (GE Healthcare, 28-9068-35) in a dark room, or developed digitally using a Chemidoc (Bio-Rad). Band intensity was quantified using Fiji image processing package [256] and subsequent data was analysed using Excel (Microsoft).

Constituent	Concentration
10x SDS running buffer ( <b>Table 2.13</b> )	1x
ddH <sub>2</sub> O (Veolia, PURELAB Option, OQ007XXM1)	70% v/v
Methanol (Merck, 32213)	20% v/v

**Table 2.18. Transfer buffer constituents for Western blot**

## 2.12 Plasmid Propagation

The Biglycan cDNA insert (**Supplementary Figure S5**) was artificially synthesised and inserted into pFUW [257], pTREAUTO3 [258] and pLenti Puro [259] lentiviral vectors by Genscript, USA, with product sequencing. Maps for all lentiviral vectors are shown in the appendix (**Supplementary Figures S6 – S12**). Lyophilised vectors were reconstituted in ddH<sub>2</sub>O (Sigma Aldrich, RNBF2734), at 200ng vector DNA per  $\mu$ l, and heated to 50°C for 15 minutes. For plasmid propagation, 10 $\mu$ l of One Shot™ Stbl3™ chemically competent *Escherichia coli* (*E.coli*) (Thermo Fisher, C737303) were transformed using 100ng of input plasmid DNA, according to manufacturer's instructions. To each vial of transformed cells, 50 $\mu$ l of super optimal broth with catabolite repression (SOC) medium (Thermo Fisher, C737303) was added. Each vial of transformed cells was shaken at 37°C and 225rpm for 1 hour in an incubator shaker (Kuhner, ISF-1-W). Cells were spread onto a pre-warmed selective plate containing 2% w/v LB Broth (Powder) (Fisher BioReagents, BP1427-500), 1.5% w/v agar (Sigma Aldrich, A1296), in ddH<sub>2</sub>O, supplemented with 100 $\mu$ g per ml ampicillin (Fisher BioReagents, BP1760). Plates were incubated for 16 hours at 37°C. Individual colonies were selected and inoculated into 3ml bacterial growth medium (2% w/v LB, in ddH<sub>2</sub>O, supplemented with 100 $\mu$ g per ml ampicillin) to produce a starter culture. Starter cultures were incubated for 8 hours at 37°C and 225rpm, under aerobic conditions, in an incubator shaker (Kuhner, ISF-1-W). This was followed by inoculation of 1ml starter culture into 50ml bacterial growth medium and incubated for 16 hours at 37°C and 225rpm, under aerobic conditions, in an incubator

shaker (Kuhner, ISF-1-W). Plasmid DNA was harvested and purified using the S.N.A.P.<sup>™</sup> Plasmid DNA MidiPrep Kit (Thermo Fisher, K191001), according to manufacturer's instructions, and eluted in 750µl ddH<sub>2</sub>O (Sigma Aldrich, RNBF2734). Quantity and quality of plasmid DNA was verified using the Nanodrop (Labtech, ND1000).

### 2.13 Generation of Lentivirus Stocks

HEK293T cells were seeded one week prior to generation of lentiviral stocks in HEK293T culture medium (**Table 2.1**) and propagated as described under Chapter 2.1. On day 0, HEK293T cells were split and singularised by pipetting. After 24h, 10cm culture dishes (Corning, 430167) were coated with poly-L-lysine (Sigma Aldrich, P8920). Poly-L-lysine solution was diluted 1:40 in PBS, followed by addition of 3ml solution per 10cm dish and incubated for 15 minutes at AT. Dishes were washed 3 times for 5 minutes each with PBS. HEK293T cells were plated at  $2 \times 10^6$  cells per 10cm dish in 10ml DMEM medium. The following day, HEK293T cells were transfected by dropwise addition of solution containing packaging and transfer vectors (**Table 2.19**). Medium was replaced after 16 hours with DMEM medium containing 1x sodium pyruvate, at 5ml per 10cm dish, before returning cells into culture under standard conditions. At 24h post-media change, lentivirus-containing medium was collected and filtered through a 0.45µm filter (Rephile, RJP3222SH). New DMEM medium with 1x sodium pyruvate was added to the cells and lentivirus was collected again after a further 24h, with filtering. Virus was stored at -80°C until required. For lentiviral stocks requiring concentration to enhance multiplicity of infection (MOI), the appropriate volume of lentiviral stocks were pooled and centrifuged using an Avanti J-26S Series high-speed centrifuge (Beckman Coulter) at  $49,000 \times g$ , for 2.5h, at 4°C. Supernatant was completely discarded and an appropriate volume of basal DMEM basal medium was added to the cell pellet. After incubating on ice for 10 minutes, the lentiviral pellet was re-suspended.

Component	Amount per 10cm dish
0.1x TE, pH 8.0 (Table 2.20)	450µl – total plasmid volume
Gag-pol packaging plasmid [260]	6.5µg
VSV-G packaging plasmid [260]	3.5µg
Prsv-rev packaging plasmid [260]	2.5µg
Transfer plasmid	10µg
2.5M CaCl <sub>2</sub> (Sigma Aldrich, C2536)	50µl
2x HBS (Table 2.20)	500µl

**Table 2.19. Constituents of solution containing vectors for generation of lentiviral stocks via transfection of HEK293T cells.**

TE buffer (1X), pH 8.0	
Component	Concentration
Tris, pH 8.0 (Merck, 348317)	10mM
EDTA (Fisher Scientific, BP2482 )	1mM
ddH <sub>2</sub> O (Sigma Aldrich, RNBF2734)	-
2xHBS	
Component	Concentration
NaCl (VWR, 27810.364)	281mM
2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulphonic acid (HEPES) (VWR, 30487.297)	100mM
Na <sub>2</sub> HPO <sub>4</sub> , pH 7.12 (BDH Laboratory Supplies, 102494C)	1.5mM

**Table 2.20. Buffer constituents for generation of lentiviral stocks. All buffers were sterile and filtered prior to use.**

## 2.14 Lentiviral Transduction

All cells to be transduced were seeded at approximately 10,000 cells per cm<sup>2</sup>, 8 hours prior to transduction. All lentiviral suspensions were supplemented with polybrene (Sigma Aldrich, 107689) at a working concentration of 8µg per ml and thoroughly mixed by pipetting prior to transduction. Cells were transduced with lentivirus at a lentiviral suspension-to-medium ratio of 1:1. After 16 hours, cells were washed twice in PBS, followed by addition of complete medium appropriate to the cell line, as detailed in Chapter 2.1, Mammalian Cell Culture. Where cells contain a marker for antibiotic resistance, neomycin (Sigma Aldrich, N6386) was added to the medium at a working concentration of 500µg per ml until all non-transduced cells perished. Where required, lentiviral titre was obtained using the Lenti-X™ Provirus Quantitation Kit (Takara, 631239).

## 2.15 Statistical Analysis

Statistical analyses of data not derived from mRNAseq outputs were carried out using GraphPad Prism version 8.0.0 for Windows (GraphPad Software, La Jolla California USA, [www.graphpad.com](http://www.graphpad.com)). The error bars on all graphs represent the standard error of the mean (SEM). Statistical analyses for between-group comparisons were one-way analysis of variance (ANOVA) with post-hoc tests for multiple comparisons and paired t-tests. All tests were 2-tailed unless otherwise stated. Significance (p) values are given as \* ≤ 0.05, \*\* ≤ 0.01, \*\*\* ≤ 0.001 and \*\*\*\* ≤ 0.0001.

## Chapter 3

# Isolation and Transcriptomic Profiling of Dormant and Proliferating Breast Adenocarcinoma Cells

### 3.1 Introduction

In breast cancer, cellular dormancy represents a significant problem of unmet clinical need. Over the past three decades, the treatments for BC of all subtypes has significantly improved and, with it, the prognosis of patients [12, 13]. However, relapse within the brain, affecting 15-30% of breast cancer patients, is associated with an extremely poor prognosis [20, 21]. Late development of brain metastases may be caused by disseminated breast cancer cells, which have evaded treatment of the primary cancer and/or metastases to organs other than the brain, and persist in an asymptomatic, reversible state of quiescence. This phenomenon is known as cancer dormancy.

There are no therapies, or ongoing clinical trials, for the targeting of dormant breast cancer cells in the brain. Such cells will be resistant to chemotherapy, which target proliferating cancer cells [14, 53]. In addition, there is evidence to suggest that dormant cells are protected from chemotherapy and radiotherapy by endothelial cells of the perivascular niche within the brain [222]. As such, a deeper understanding of the molecular mechanisms surrounding breast cancer cell dormancy in the brain is warranted for the development of novel and specific therapeutics. Potential targets may be critical survival mechanisms in dormancy maintenance, for targeted cytotoxic therapy of dormant cells. Alternative targets include mechanisms which are involved in dormancy induction, to induce a state of growth arrest in established lesions. To identify such targets, the next step, and the basis for this chapter, was to determine the transcriptomic profiles of dormant and proliferating breast cancer cells in the brain.

### **3.1.1 Hypothesis**

The hypothesis for this chapter is that dormant and proliferating breast cancer cells within the brain display distinct transcriptomic profiles.

### **3.1.2 Aims and Objectives**

This chapter aims to identify the extent to which the gene expression profiles of dormant and proliferating breast cancer cells in the brain diverge. From this data, potentially implicated genes and pathways in the induction and maintenance of dormancy will be identified, representing the first step in uncovering potential novel therapeutic targets for the treatment of breast cancer. To achieve this, this chapter aims to:

**1) Identify distinct populations of dormant and proliferating breast adenocarcinoma cells within mouse brain tumour xenografts**

This will be achieved by first demonstrating that cancer cell proliferation can be tracked using a membrane dye. This will be followed by staining of breast adenocarcinoma cells with the dye, and utilizing a murine model previously published by our group. This will involve the intra-carotid injection of breast cancer cells into immunosuppressed mice to generate brain tumour xenografts [134].

**2) Isolate viable RNA from dormant and proliferating human breast adenocarcinoma cells**

This will be achieved by harvesting mice brains at 28 days post-injection of cancer cells, a time point as such that proliferating cancer cells have lost tracking dye through proliferation, and when non-dividing cancer cells have retained it. Brains will be dissociated, and pertinent cells will be isolated by flow cytometry for RNA extraction.

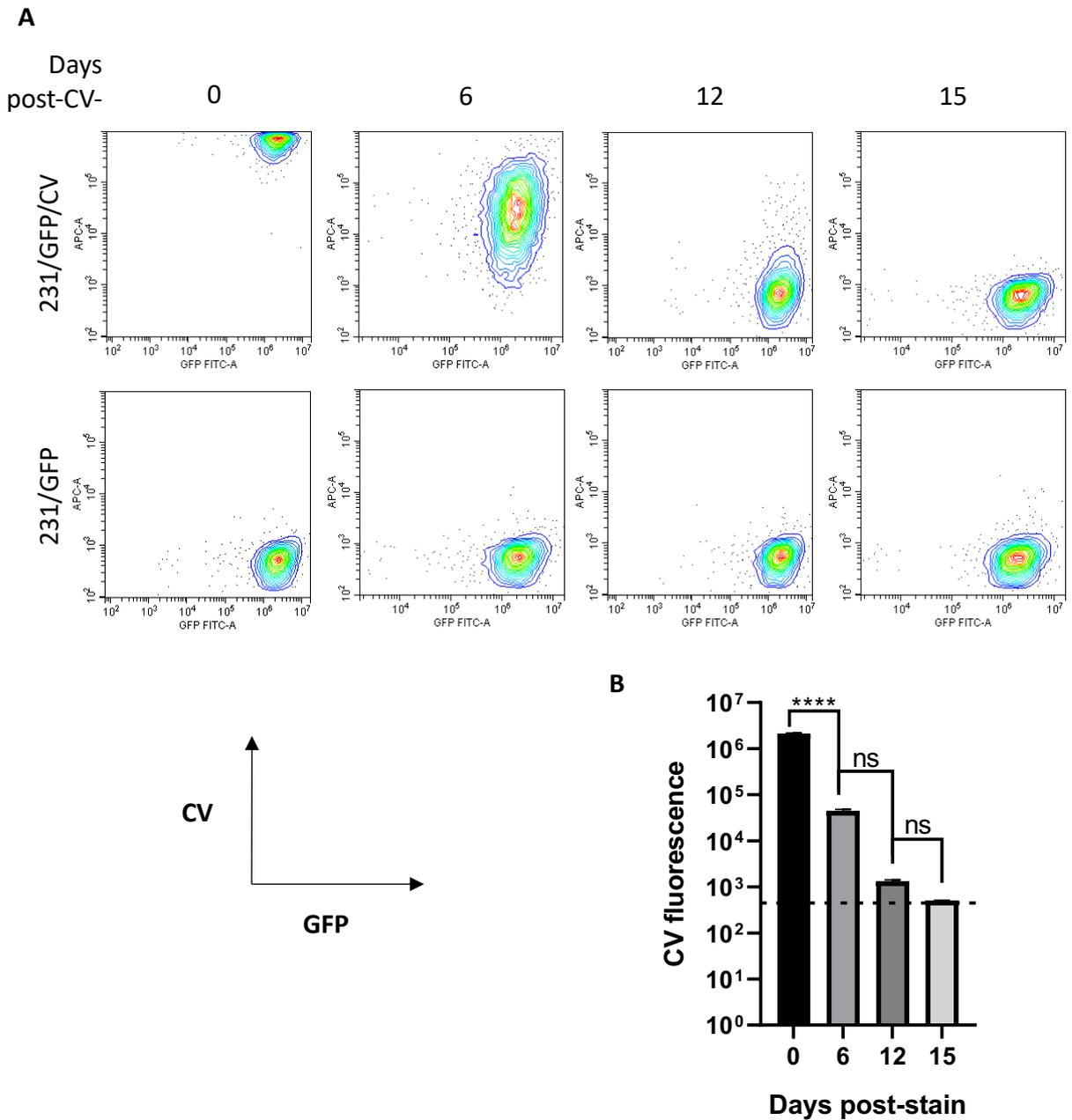
### **3) To identify genes and pathways potentially implicated in breast cancer cell dormancy in the brain**

To achieve this, mRNAseq will be performed using isolated RNA from dormant and proliferating breast cancer cells. Implicated genes and pathways will be identified using transcriptomic analysis of differentially expressed genes.

#### **3.2 Tracking of Dormant and Proliferating MDA-MB-231 Cell Proliferation using a Lipophilic Cell Membrane Dye**

To identify dormant and proliferating cancer cells for isolation of each population *in vivo*, the far-red general cell membrane dye, CellVue® Claret (CV) was used as a tracker of proliferation. CV incorporates into lipid regions of the plasma membrane and has previously been used as a proliferation-tracking dye of immune cells *in vitro*. As a membrane dye, CV is gradually lost over time as a cell divides and produces successive generations of daughter cells, losing a proportional amount of dye during each division [261].

The timeline for CV loss in MDA-MB-231 cells was determined through propagation of labelled cancer cells *in vitro*. MDA-MB-231 cells tagged with GFP (231/GFP) were stained with CV and the fluorescence intensity of CV+ cells were quantified by flow cytometry at 5 hours, then at 6, 12- and 15-days post-staining. At 5 hours, all 231/GFP cells were CV+. The CV fluorescence intensity, given as a percentage of the fluorescence intensity of cells at 5 hours post-staining, dropped rapidly and significantly to approximately 2.12% by day 6. This dropped further to 0.06% and 0.02% by days 12 and 15, respectively (Figure 3.1).



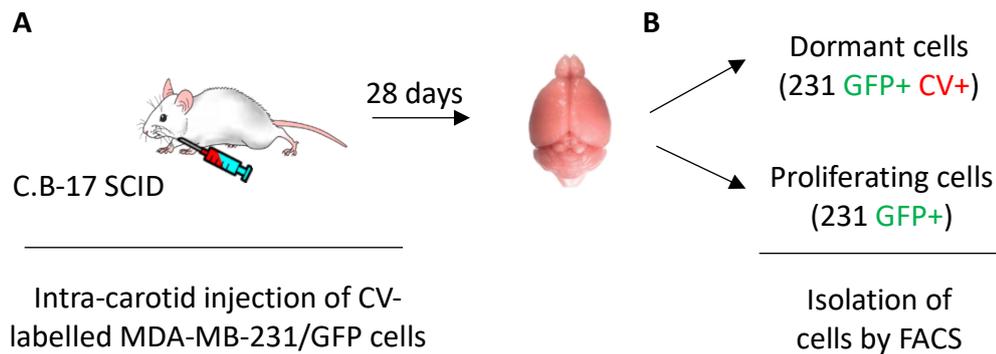
**Figure 3.1. Proliferation tracking of MDA-MB-231 cells *in vitro* with the vital membrane dye, CellVue® Claret.** GFP-tagged MDA-MB-231 (231/GFP) cells were labelled with CV dye. The fluorescence intensity of CV dye was recorded at the times indicated. A) Flow cytometry analysis of CV fluorescence intensity of 231/GFP/CV cells and unlabelled 231/GFP cells. B) Quantification of mean CV fluorescence of all intact cellular bodies at the given time points. Performed was an ordinary one-way analysis of variance (ANOVA) and Tukey's multiple comparisons test, with a single pooled variance.  $N=3$ . Dotted line represents baseline threshold for background APC fluorescence (used for CV detection) in 231/GFP cells.

### 3.3 Generation of Mouse Brain Tumour Xenografts

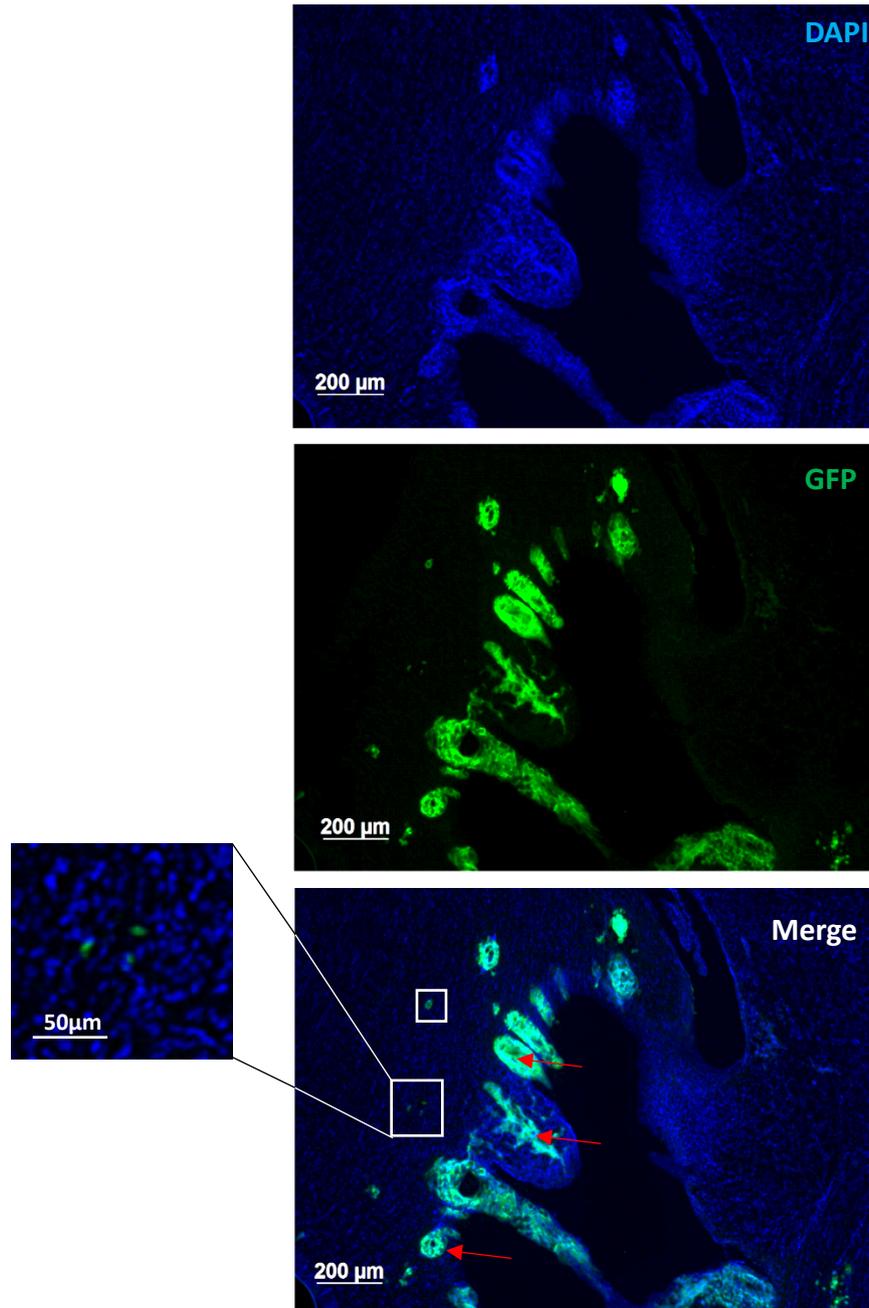
A previously established BrM model to study breast cancer cell dormancy *in vivo* was then employed [134]. The goal was to isolate dormant and proliferating cancer cells from mouse brains by fluorescence-activated cell sorting (FACS) and to perform molecular profiling of the two populations. To this end,  $1 \times 10^5$  231/GFP/CV breast cancer cells were administered via intra-carotid injection into C.B-17 severe combined immunodeficiency (SCID) mice. In the model employed, metastatic lesions within the brain were allowed to establish for 28 days (**Figure 3.2A**). At 28 days post-injection of 231/GFP/CV cells, the mice were terminally perfused, followed by brain isolation and dissociation as described in Chapter 2.3.4. Based on my *in vitro* data, it was expected that the proliferating cells would have lost CV dye by this time point. Proliferating cancer cells were therefore identified as those having lost CV dye (GFP+). Conversely, dormant cancer cells were identified as those having retained CV dye (GFP+CV+). Immunofluorescence imaging of brain sections, 28 days post-cancer cell engraftment, revealed large lesions of 231/GFP cells within the brain parenchyma, in addition to solitary cancer cells (**Figure 3.3**). Each population was isolated by FACS from 9 or 10 mice and pooled prior to isolation of RNA (**Figure 3.2B**).

Based on *in vitro* data, as shown in Chapter 3.2.1, it was expected that 231/GFP cells which retained the CV dye (231/GFP/CV) were dormant, either through slow cycling or an absence of proliferation. In contrast, cells in which CV was not detected (231/GFP) were considered as actively proliferating cells. Gating parameters were set using naïve mouse brain, brain isolated from a mouse injected with 231/GFP cells and a brain isolated from a mouse injected with 231/CV cells (**Figure 3.4**). *In vitro* cultured 231, 231/GFP, 231/CV and 231/GFP/CV cells were also employed to set the gates for cell sorting (data not shown). A representative plot of three independent experiments showing dormant and proliferating cancer cell populations during sorting is shown in **Figure 3.4**. Between each of the three experiments, approximately 0.5 to 2% of the total cancer cell population sorted were dormant. These are displayed in Table 3.1 along with other descriptive statistics for each cell sort. I also performed cell sorting on 231/GFP/CV

cells grown *in vitro* for 4 weeks, under identical gating thresholds. In contrast to the *in vivo* model, the percentage of cells identified as GFP+CV+ *in vitro* was approximately 0.004% (**Table 3.1**). Thus, the proportion of growth-arrested 231 cells is significantly lower in cell culture than in the brain, suggesting that growth arrest is induced by the tumour microenvironment.

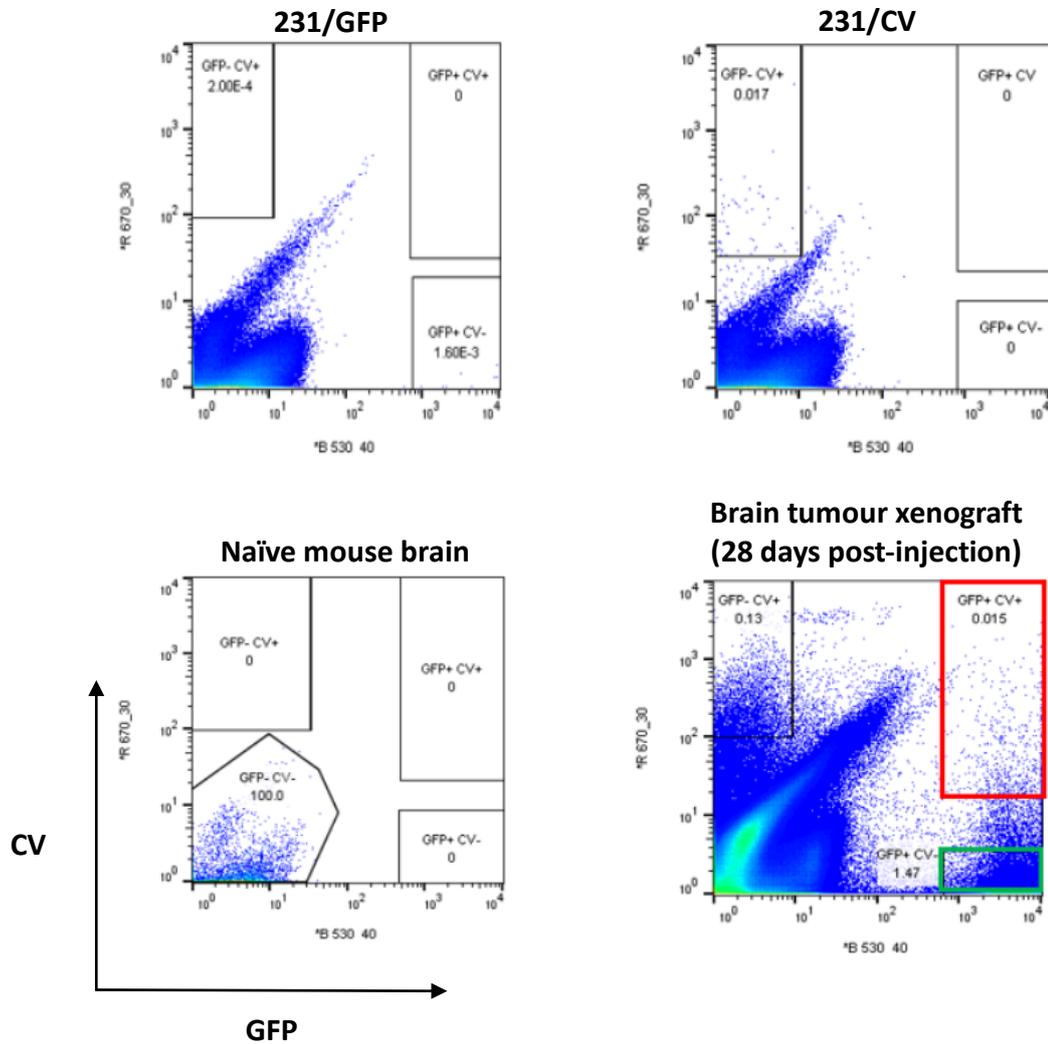


**Figure 3.2. *In vivo* model to study breast cancer cell dormancy in the brain.** A) GFP-expressing MDA-MB-231 cells were labelled with the vital dye, CellVue (CV) Claret. The cells were injected into the internal carotid artery of C.B-17 SCID mice to colonise the brain. After 28 days, mice were terminally perfused, and brains were dissociated. B) Dormant and proliferating breast cancer cells were sorted by FACS according to GFP+CV+ and GFP+ phenotypes, respectively.



**Figure 3.3. Identification of brain tumour xenografts *in vivo* by immunofluorescence staining.** MDA-MB-231/GFP/CV cells were injected into the internal carotid artery of C.B-17 SCID mice. Cells were allowed 28 days to colonise the brain before harvesting. Large lesions (green, red arrows) were observed alongside solitary 231 cells (green, white boxes) within the brain parenchyma. Abbreviations: green fluorescent protein (GFP); 3'6-diamidino-2-phenylindole (DAPI).

### 3.3.1 Sorting of Dormant and Proliferating MDA-MB-231 Cells



**Figure 3.4. Isolating dormant (231/GFP/CV) and proliferating (231/GFP) cancer cells by fluorescence activated cell sorting.** Mouse brain tumour xenografts were harvested at 28 days post-intra-carotid injection of 231/GFP/CV cancer cells into C.B-17 SCID mice. Following brain dissociation, cancer cells (GFP+) were sorted according to loss or retention of CV vital membrane dye. Gating thresholds were set using naïve mouse brain and brain tumour xenografts generated with 231/GFP and 231/CV cells, respectively. Data is representative of three independent experiments.

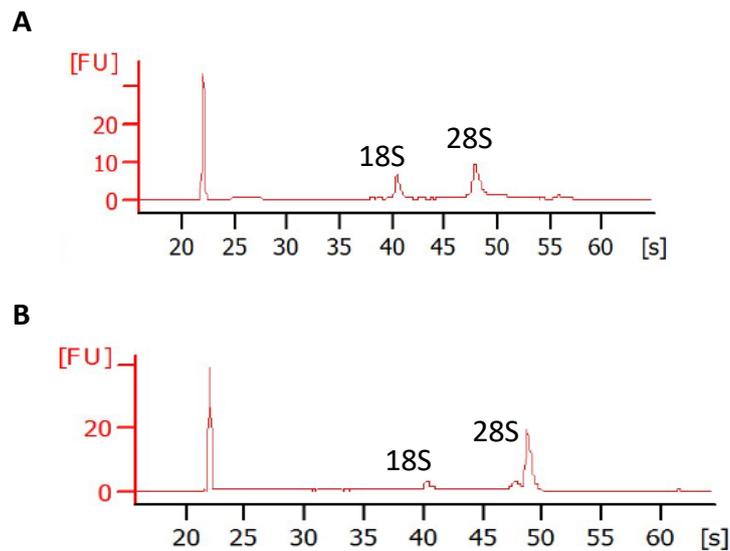
Experiment	Number of dormant cells (231/GFP/CV)	Number of proliferating cells (231/GFP)	Percentage of dormant cells / %
Xenograft – 10 mice	418	20,250	2.064
Xenograft – 10 mice	747	59,795	1.249
Xenograft – 9 mice	134	23,008	0.582
<b>231/GFP/CV</b> (4 weeks <i>in vitro</i> )	141	3,378,833	0.004

**Table 3.1. Dormant cells account for a small fraction of the total cancer cell population within mouse brain tumour xenografts.** Shown is the number of dormant (231/GFP/CV) and proliferating (231/GFP) cells isolated from mouse brain tumour xenografts and from 231/GFP/CV cells harvested after 4 weeks *in vitro*. Also shown is the percentage of dormant (231/GFP/CV) cells within the total sorted cancer cell population.

#### 3.4 Generation of mRNAseq Sequencing Library from a Small Starting Input of RNA

RNA was isolated from the few (134 – 747) dormant cells immediately following cell sorting to preserve its structural integrity. RNA was also extracted from 1,000 proliferating cancer cells so that any potential degradation and/or loss of RNA mimicked that of the dormant population, yielding RNA of comparable quality. RNA quantity and quality were checked by Bioanalyser with pico-scale sensitivity. The resulting traces displayed identifiable 18S and 28S ribosomal subunits with no visually apparent degradation products for both populations of cells (**Figures 3.5A, B**).

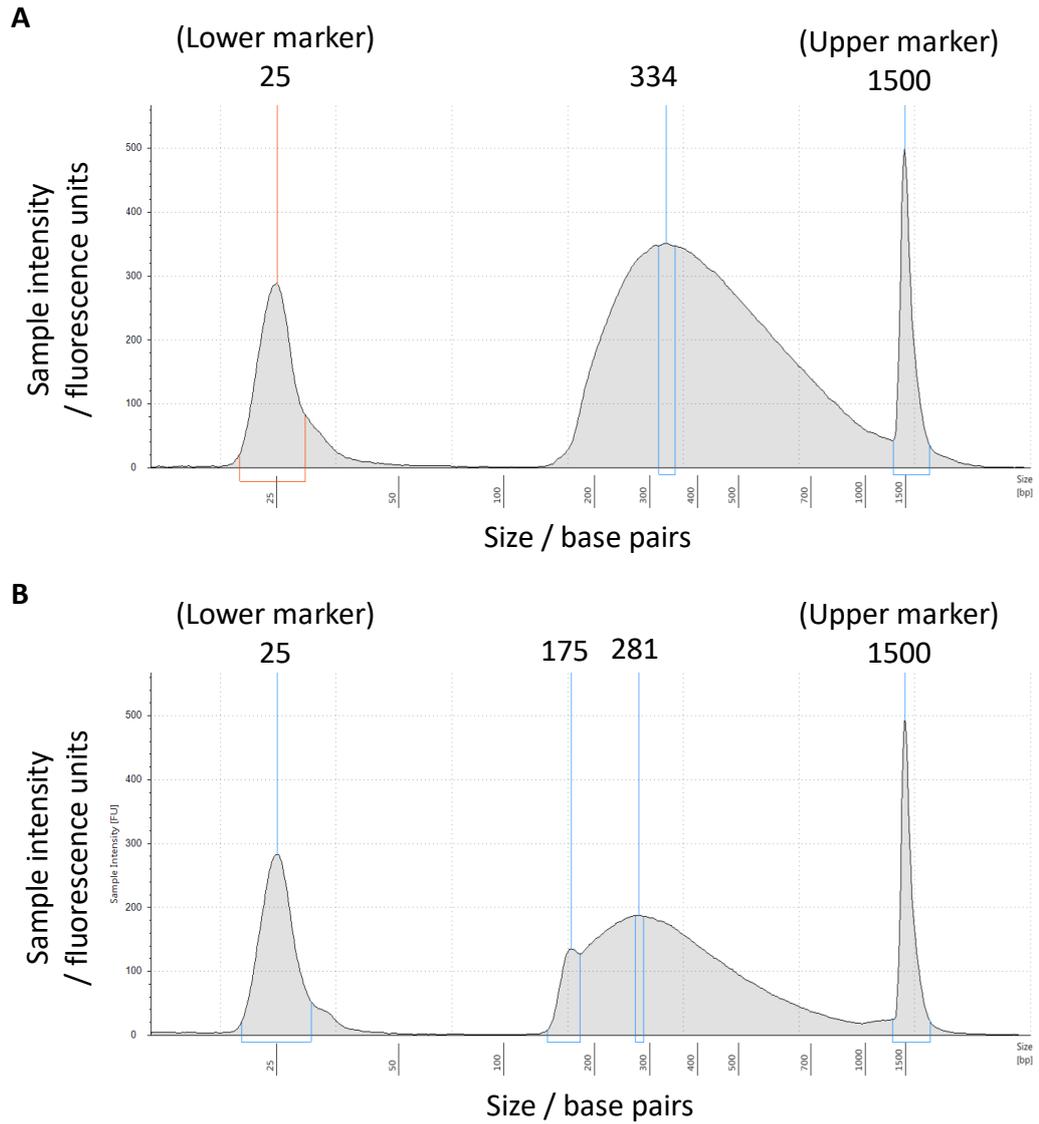
As anticipated, the yield of total RNA was small, at 0.4 to 3.9 ng from up to 747 dormant cells (**Table 3.2**). For the 1,000 proliferating cells, a yield of 0.7 to 2.5ng was observed (**Table 3.2**). For generation of cDNA sequencing libraries, the previously published Smart-seq2 protocol was employed [224]. This protocol was designed for library synthesis from single cells but can be applied to pooled cells, selecting for polyadenylated RNA with improved sensitivity, accuracy and full-length coverage across transcripts, compared to alternative high-throughput methods [224]. Owing to volumetric limitations of the Smart-seq2 library synthesis protocol, 114pg of RNA per sample was used for the synthesis of a polyadenylated sequencing library (**Figure 3.6**). This was followed by paired end mRNAseq, as described in Chapter 2.6.



**Figure 3.5. Quality control of RNA isolated from a small number of dormant and proliferating cancer cells.** Total RNA was isolated from 134 – 747 dormant (A) and 1,000 proliferating (B) MDA-MB-231 cells isolated from mouse brain tumour xenografts. Graphs representative of 3 independent experiments. Ribosomal subunits are denoted by 18S and 28S.

Cell sort	Population sorted	Number of cells	Total RNA yield (ng)
1	231/GFP/CV	418	3.9
1	231/GFP	1,000	1.2
2	231/GFP/CV	747	0.9
2	231/GFP	1,000	0.7
3	231/GFP/CV	134	0.4
3	231/GFP	1,000	2.5

**Table 3.2. Yields of total RNA derived from dormant and proliferating MDA-MB-231 cells *in vivo*.** Given is the total RNA yield with respective cell count from three independent experiments.



**Figure 3.6. Quality control of cDNA library for mRNAseq.** Representative cDNA size distribution of synthesised cDNA, as output from the Smart-seq2 reaction [224]. Plots shown are polyadenylated libraries generated from (A) proliferating and (B) dormant polyadenylated mRNA. Inputted DNA with corresponding peak sizes is flanked by lower and upper cDNA markers.

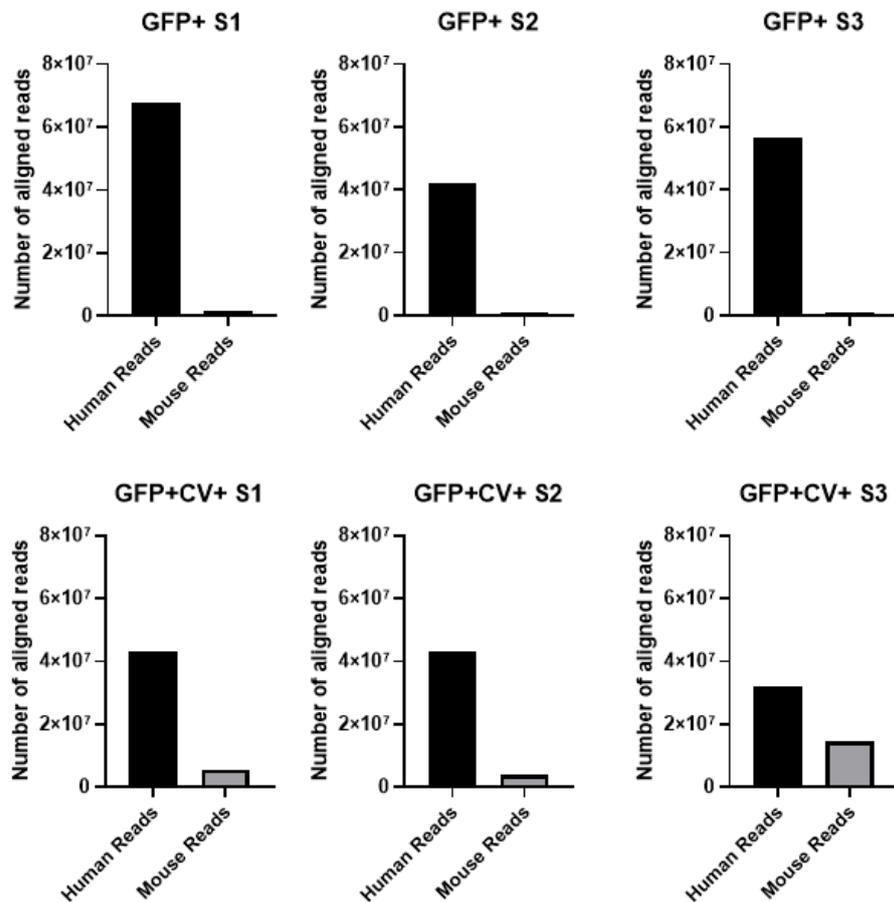
### 3.4.1 Processing Stream I: Pre-Processing of mRNAseq Data and Quality Control (Individual Genome Alignment)

First-round pre-processing of mRNAseq data (processing stream I) was carried out using in-house scripts (Next Generation Sequencing Facility, University of Leeds). The processing pipeline is detailed in Chapter 2.6. In processing stream I, raw sequencing reads were aligned individually to both human (GRCh38/hg38, December 2013 assembly, UCSC) and mouse (GRCm38/mm10, December 2011, UCSC) reference genomes. Analysis under processing stream I assumed minimal contamination with reads that were of mouse origin. The total read count per sample was at least 57 million reads, as shown with the descriptive statistics for the analysis in **Table 3.3**.

Sample	Library size (bp)	Read length (bp)	Total reads	Total base pairs
GFP+ S1	252	151	78,637,675	11,874,288,925
GFP+ S2	256	151	50,591,786	7,639,359,686
GFP+ S3	256	151	64,150,823	9,986,774,273
GFP+CV+ S1	231	151	63,557,464	9,597,177,064
GFP+CV+ S2	243	151	57,907,700	8,744,062,700
GFP+CV+ S3	246	151	58,946,470	8,900,916,970
			<b>373,791,918</b>	<b>47,541,662,648</b>

**Table 3.3. Descriptive statistics for mRNAseq alignment data (processing stream I).** Gene-wise read counts were generated from paired-end sequencing data for proliferating (GFP+) and dormant (GFP+CV+) MDA-MB-231 cells, with an allowance for multi-mapping reads. The data indicates library size and read length for each sample during the sequencing analysis. Given are the total number of reads sequenced for each sample and the corresponding base pair count, calculated as (*total reads x read length*). Data inclusive of all three independent experiments.

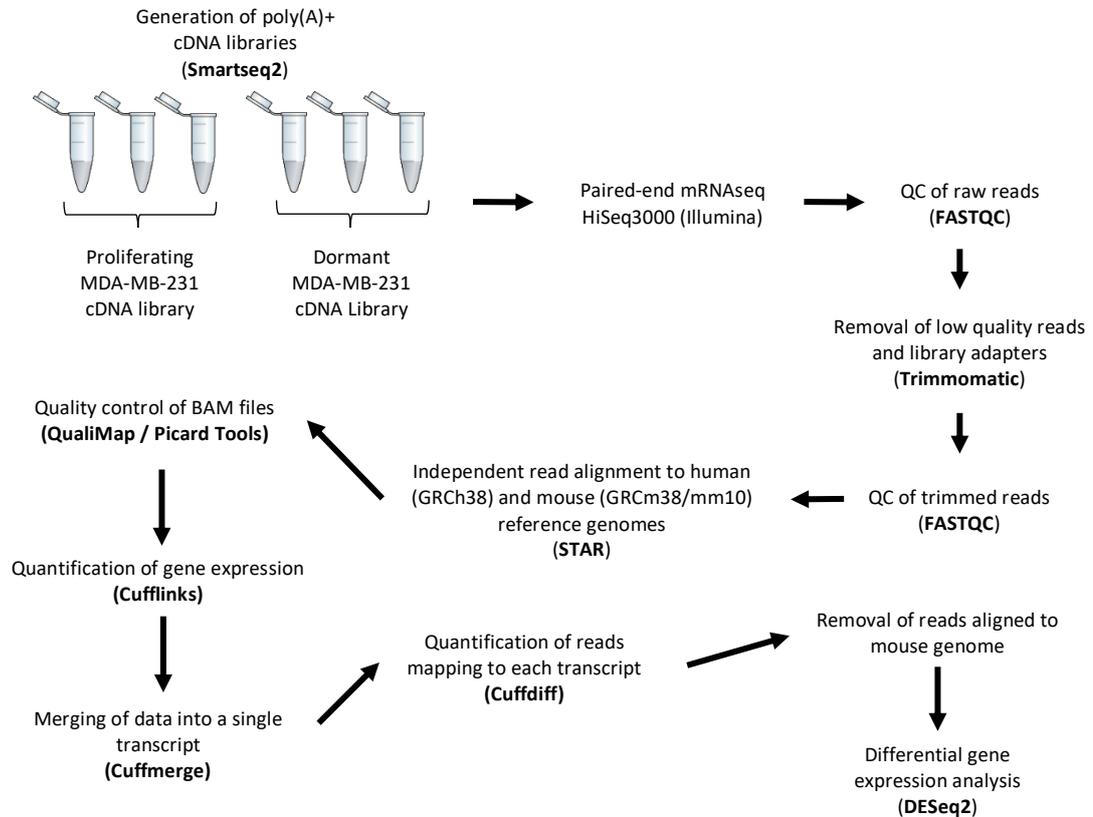
When aligned individually to human and mouse genomes, some contamination with mouse sequences was observed in all samples (**Figure 3.7**). This affected dormant (GFP+CV+) populations, in which low numbers of cells were collected from mouse brains, to a greater extent than proliferating populations. Despite this, all reads which aligned to human genomic positions were counted at over 30 million reads per sample. This is above the minimum threshold for effective differential expression analysis [262, 263]. Subsequent analysis pipelines were carried out using only human-aligned reads using DESeq2 size-factor normalisation to generate normalised read counts.



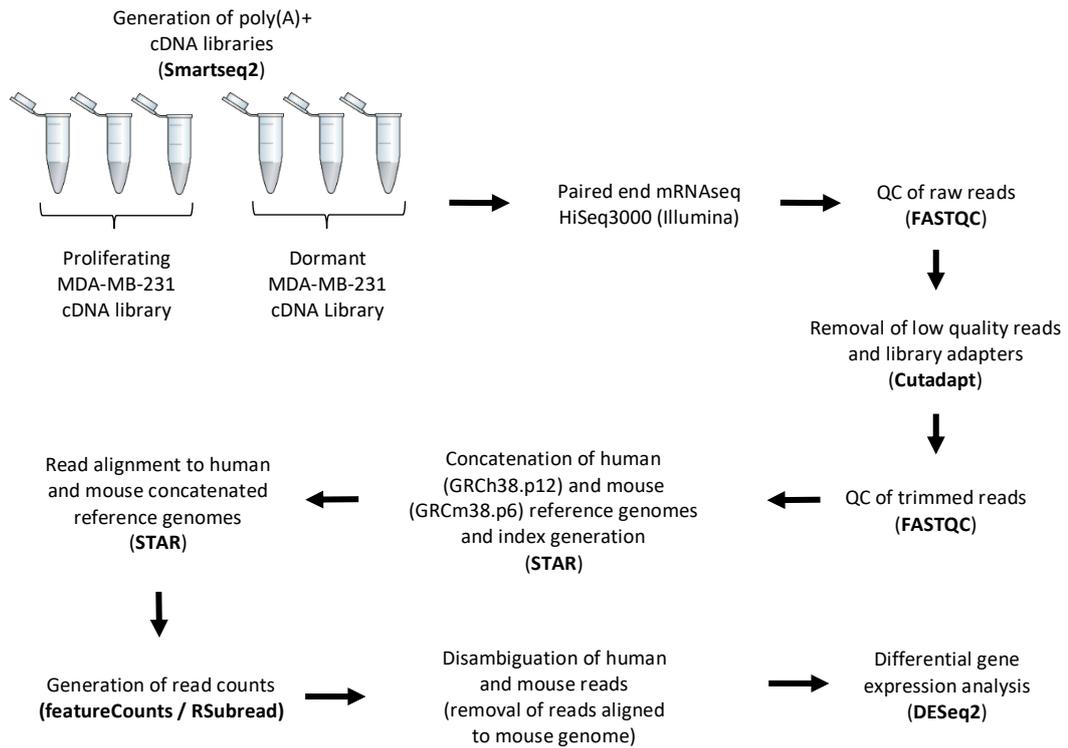
**Figure 3.7. Proportion of human- and mouse-aligned reads in dormant and proliferating MDA-MB-231 cells (processing stream I).** Gene-wide read counts were generated from paired-end sequencing data for proliferating (GFP+) and dormant (GFP+CV+) MDA-MB-231 cells, with an allowance for multi-mapping reads. Each sample was aligned individually to human (GRCh38/hg38) and mouse (GRCm38/mm10) genomes.

### 3.4.2 Processing Stream II: Pre-Processing of mRNAseq Data and Quality Control (Concatenated Genome Alignment)

It was important to recognise that some reads mapping to known human exons when using processing stream I (**Figure 3.8**) could actually be mouse-derived reads. This could be due to contamination from mouse brain cells incorrectly sorted as cancer cells during cell sorting of dissociated xenografts. In such cases, incorrect mapping may be a result of repetitive elements or homologs, or errors during sequencing. Despite the human-aligned total read count being optimal for differential gene expression analysis, such counts were generated from a small number of isolated cells. It was therefore important to determine, to a degree of confidence, the proportion of uniquely mapped genes to the human genome, and those mapped to both human and mouse genomes. The pre-processing of raw sequencing data was therefore repeated using an independent pipeline, processing stream II (**Figure 3.9**), using a merged human and mouse reference genome for alignment. The intention was to present the alignment software with competition, in cases where reads present with a high degree of homology between both genomes. This would effectively increase the probability of correct alignment. The resulting data was to be analysed in parallel to the dataset from processing stream I. Quality control of raw reads was performed using FastQC, before and after removal of low quality reads and those pertaining to sequencing adaptors. Prior to first-round trimming by Cutadapt, the per sequence GC content for both proliferating (GFP+)- and dormant (GFP+CV+)-derived reads was close to the theoretical distribution (**Figure 3.10A**). However, first-round trimming uncovered GC-rich regions in reads within all 3 dormant cell populations. This phenomenon did not occur in any of the proliferating cell populations. Observed was a higher number of reads with approximately 50% GC content, above that of the theoretical distribution, and many reads of close to 100% GC content (**Figure 3.10B**). Over-represented reads from FastQC analysis detailed many poly-GC sequences (data not shown). BLAST analysis did not associate such reads with any known human sequences and such reads were therefore trimmed. This returned the per-sequence GC content of dormant cell reads close to the theoretical distribution (**Figure 3.10C**).

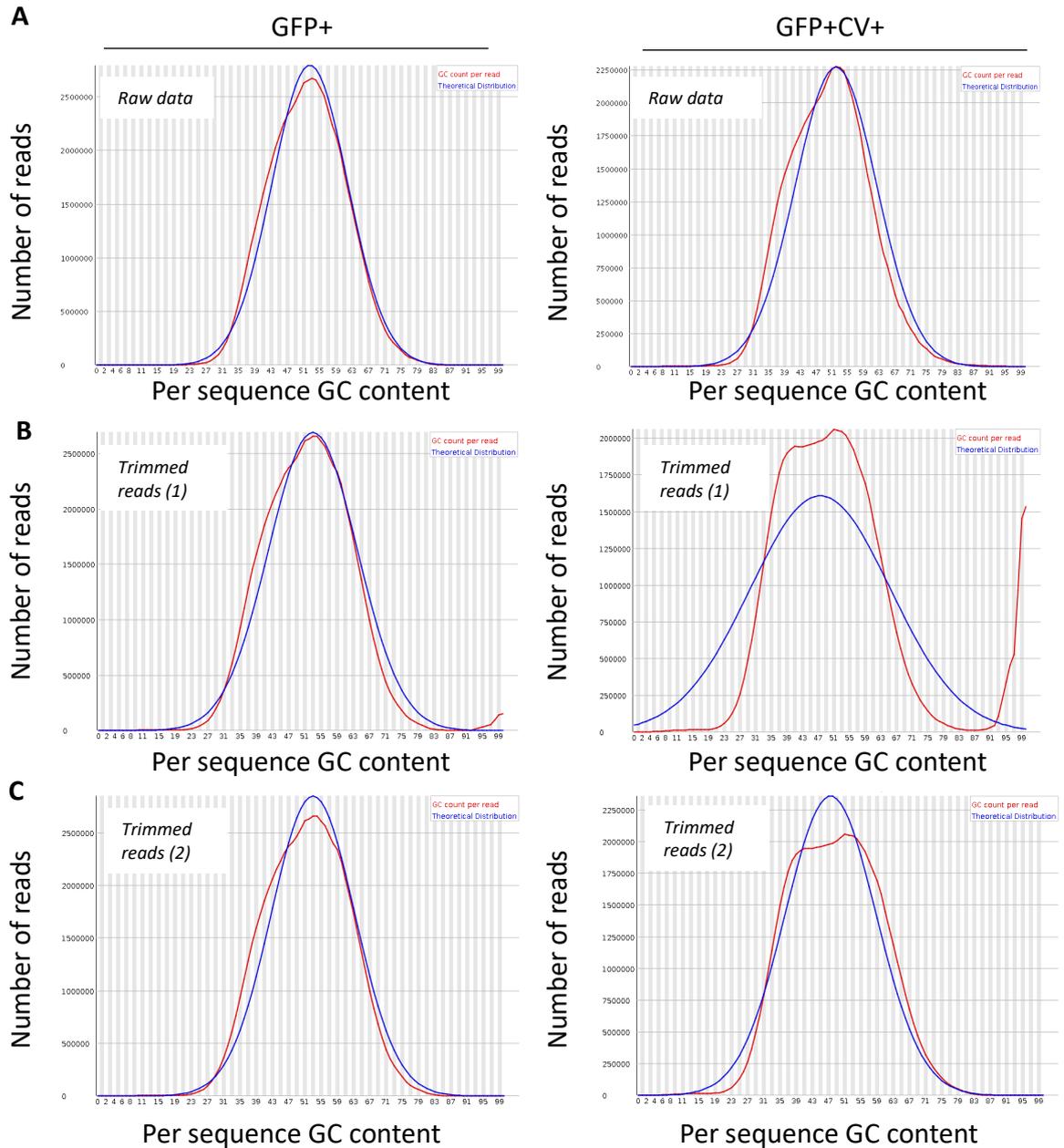


**Figure 3.8. Workflow of mRNAseq data processing – alignment to concatenated human and mouse reference genomes (processing stream I).** Polyadenylated (poly(A)+) mRNA was pulled down from total RNA derived from proliferating and dormant MDA-MB-231 cells *in vivo*. cDNA was synthesised using the published Smartseq2 protocol [224]. This was followed by paired end mRNAseq (Illumina platform). Quality control was performed before and after removal of low-quality reads and library adaptors (Cutadapt) by FastQC software. Human (GRCh38) and mouse (GRCm38/mm10) reference genomes were used independently in the alignment of gene expression data. Human and mouse chromosomal identifiers were labelled for downstream disambiguation. Quality control of BAM files following STAR alignment was carried out by both QualiMap and Picard Tools. Gene expression quantification, merging into a single transcript and quantification of final read counts were carried out by Cufflinks, Cuffmerge and Cuffdiff, respectively. Any reads aligned to the mouse genome were removed prior to analysis. Differential gene expression was carried out by DESeq2.

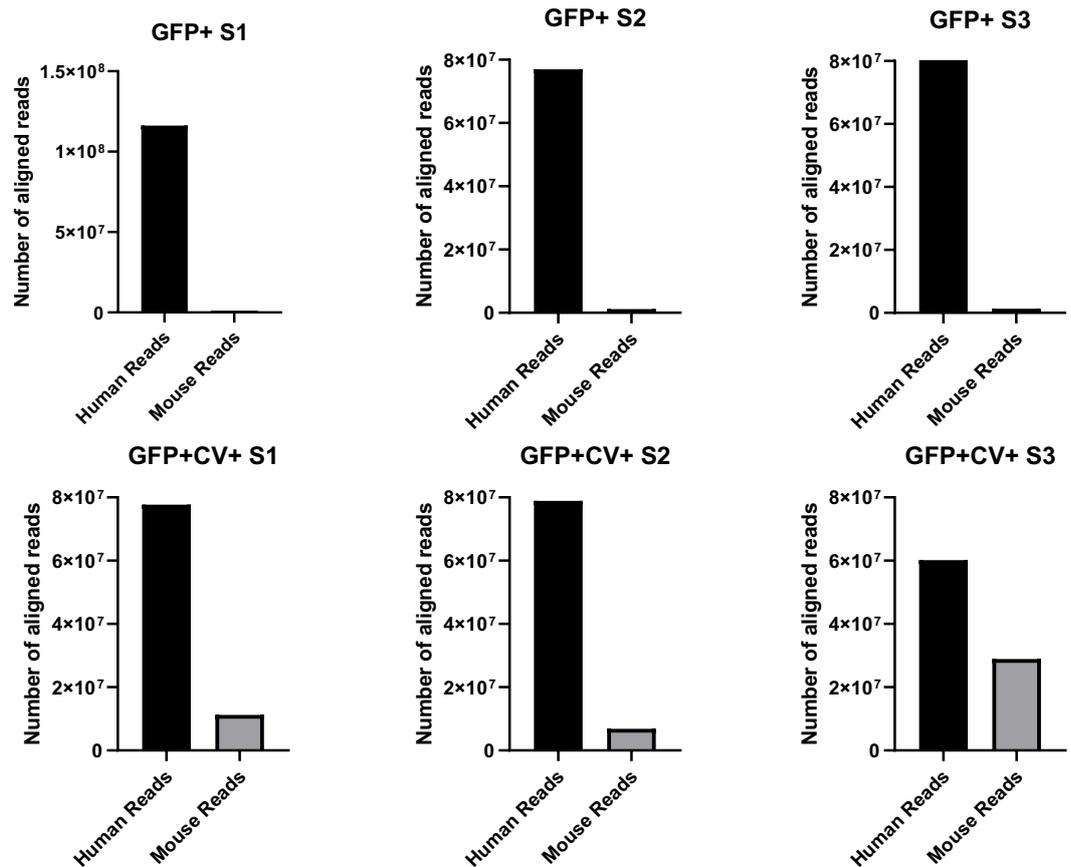


**Figure 3.9. Workflow of mRNAseq data processing – alignment to concatenated human and mouse reference genomes (processing stream II).** Polyadenylated (poly(A)+) mRNA was pulled down from total RNA derived from proliferating and dormant MDA-MB-231 cells *in vivo*. cDNA was synthesised using the published Smartseq2 protocol [224]. This was followed by paired end mRNAseq (Illumina platform). Quality control was performed before and after removal of low-quality reads and library adaptors (Cutadapt) by FastQC software. Prior to alignment, human (GRCh38.p12) and mouse (GRCm38.p6) reference genomes were concatenated and used to generate an alignment index. Human and mouse chromosomal identifiers were labelled for downstream disambiguation. Read counts of aligned reads were determined by the *featureCounts* function of the R package RSubread. The proportion of reads aligning to human and mouse reads was noted, and any reads aligning to mouse were discarded for downstream analysis. Differential gene expression analysis was performed by R package DESeq2.

Primary genome assemblies for human (release 31, GRCh38.p12, December 2017) and mouse (release M22, GRCm38.p6, September 2017), and accompanying annotations, were concatenated prior to generation of a reference index. Both genome assemblies are more current than the versions used for alignment using in-house scripts. Trimmed reads were then aligned to the merged index, using the Spliced Transcripts Alignment to a Reference (STAR) aligner. STAR is a publicly available read alignment software of which its algorithm performs seed clustering and stitching, using sequential mappable seed searching in uncompressed suffix arrays [231]. It improves on alternative alignment software given its unbiased *de novo* detection of canonical splice junctions, with improved sensitivity and precision. Concatenation of human and mouse primary genome assemblies was carried out to increase the likelihood of alignment to the correct species [231]. Where STAR is presented with a read overlapping with regions of high homology between human and mouse, the alignment algorithm, in practice, chose between human or mouse, or indicated that this particular read could not be aligned. The criteria for alignment to human or mouse genomes is guided by a scoring scheme. Built into the custom-built script for STAR aligner were penalties for matches, mismatches, insertions, deletions and splice junction gaps. This combination was stitched and assigned a score. Therefore, where regions of high homology exist within human and mouse genomes, the read was aligned to the region with the highest scoring stitched combination. This was also dependent on the score being within a user-defined range. Where reads mapped to the mouse genome, these were discarded from downstream analysis and only those with human chromosome identifiers were included.



**Figure 3.10. Per-sequence GC content of mRNAseq reads derived from dormant and proliferating breast cancer cells *in vivo* (processing stream I).** Quality control (QC) of reads derived from mRNAseq of proliferating (GFP+) and dormant (GFP+CV+) breast cancer cells, as determined by FastQC. Displayed is the per-sequence GC content against the number of reads. Images representative of three representative experiments. A) Top graphs: QC of raw data, directly from sequencing. B) Middle graphs: QC of first-pass trimming by Cutadapt, removing low quality and over-represented reads, along with those belonging to Nextera Transposase adaptors. C) Bottom graphs: QC of second-pass trimming following removal of poly-GC sequences by Cutadapt.



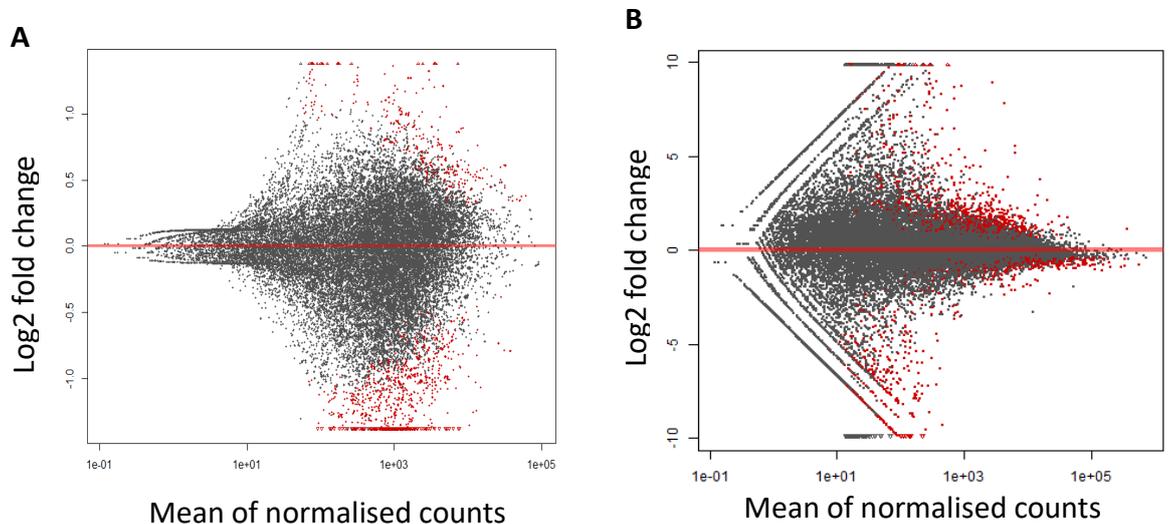
**Figure 3.11. Proportion of human- and mouse-aligned reads in dormant and proliferating MDA-MB-231 cells (processing stream II).** Gene-wise read counts were generated from paired-end sequencing data for proliferating (GFP+) and dormant (GFP+CV+) MDA-MB-231 cells, with an allowance for multi-mapping reads. Each sample was aligned to a concatenated human-mouse reference index.

The proportion of reads aligning to both human and mouse known exon positions in stream II is shown in **Figure 3.11**. For all 3 independent experiments, there was a very low number of mouse-aligned reads within the proliferating cell populations. There were higher, but still relatively low, levels of mouse contamination in two of the dormant cell populations (GFP+CV+ S1 and S2). Approximately one-third of reads in GFP+CV+ S3 were of mouse origin. However, it was important to recognise that some reads mapping to human exons could also be mouse-derived, as an effect of sequence homology. The same applies vice versa. However, sequence homology is also dependent on read-length [264]. In this mRNAseq dataset, the majority of reads were determined

to be greater than or equal to 151bp (**Table 3.3**), therefore one can be confident that the reads aligned to the correct species. Only those aligned to human sequences were carried forward for further analysis.

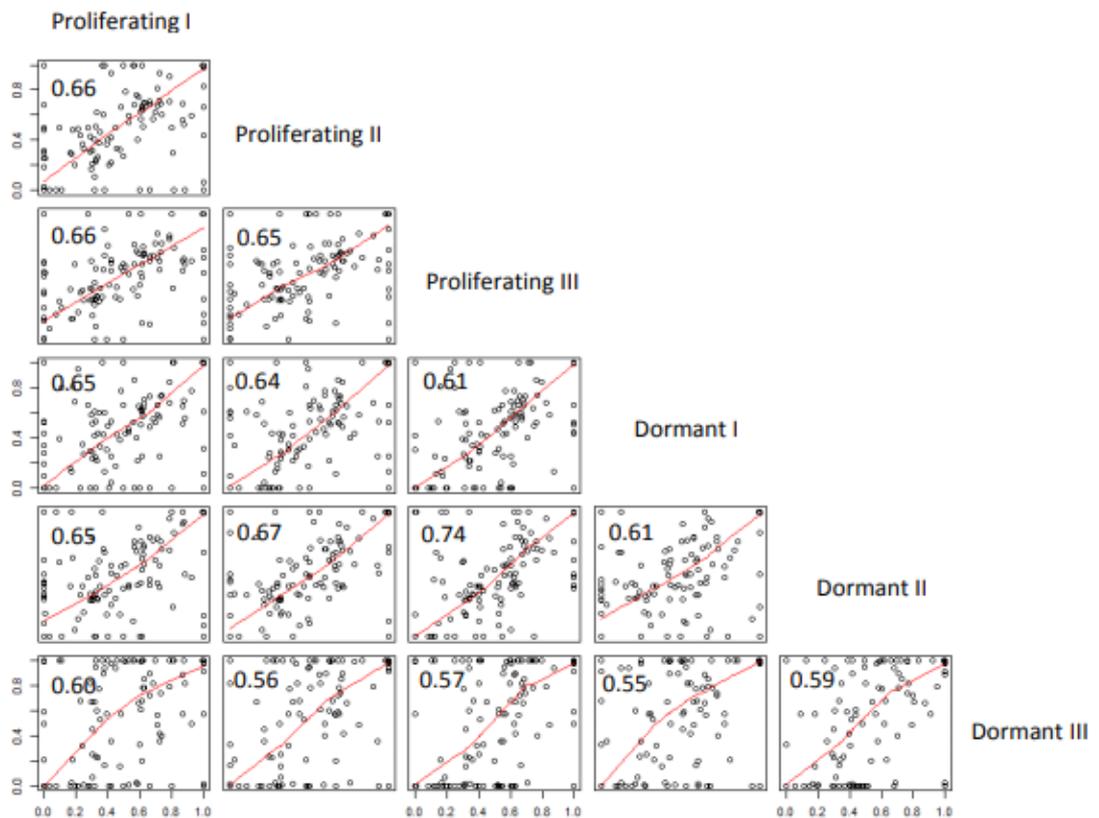
### 3.5 Analysis of Dormant and Proliferating MDA-MB-231 Cell Gene Sets

Most methods which normalise gene expression data make the assumption that most of the housekeeping genes are not altered in their expression across biological samples and are not influenced by variations in experimental conditions [265]. The MA plots, shown in **Figure 3.12**, depict log-intensity ratios (M-values) i.e. the  $\log_2$  of fold change in gene expression, against log intensity values (A-values) i.e. the mean of normalised expression. The MA plot given for processing stream I (**Figure 3.12A**) and processing stream II (**Figure 3.12B**) both demonstrate that the majority of the expression values cluster around a  $\log_2$  fold change of 0, and thus I could apply DESeq2 size-factor normalisation in downstream analyses.

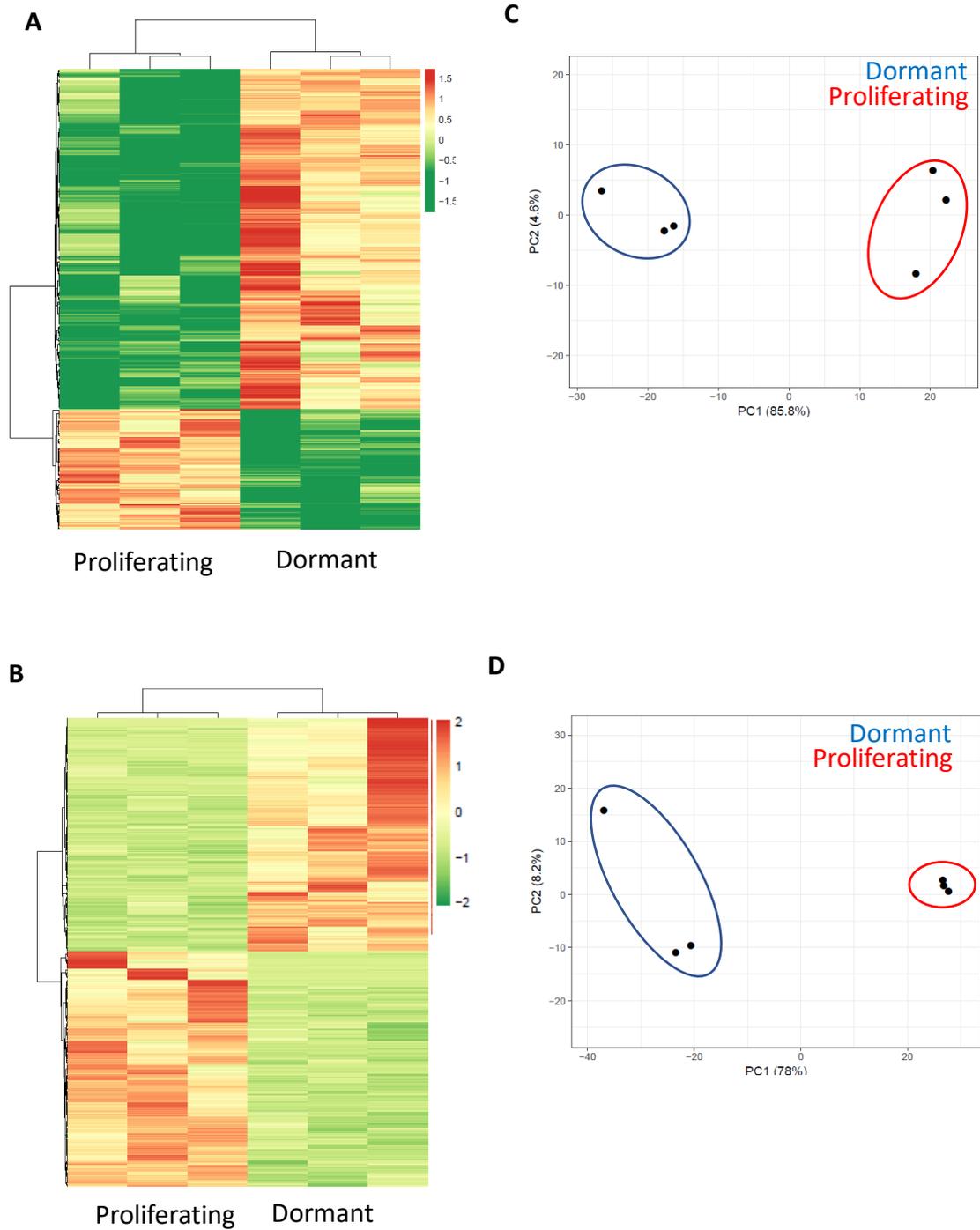


**Figure 3.12. MA-plots of normalised mRNAseq data.** Mean of normalised counts against  $\log_2$  fold change of read counts. Data represents a comparison between dormant and breast adenocarcinoma cells *in vivo*. A) mRNAseq processing stream I; B) mRNAseq processing stream II. Red plots highlight genes identified as differentially expressed ( $FDR \leq 0.05$ ).

In order to confirm that isolated cells originate from the 231 cell line, allele frequencies were counted by querying alignments from BAM files (reads aligned to the human genome; processing stream II) at known MDA-MB-231 single nucleotide polymorphism (SNP) positions. Pearson correlation coefficients for pairwise comparison of SNPs was calculated for all samples (**Figure 3.13**). The coefficient values for comparison of dormant-to-proliferating cell samples were similar to that of dormant-to-dormant and proliferating-to-proliferating samples. This provides confidence that the majority of sequenced reads for both dormant and proliferating cancer cell populations is derived from the same cell line with little contamination.



**Figure 3.13. Pairwise comparison of known MDA-MB-231 cell SNPs between dormant and proliferating breast adenocarcinoma cells (processing stream I).** Allele frequencies were counted by querying alignments from BAM files at known SNP positions. Frequencies were compared between all samples. Plotted are allele frequencies with corresponding correlation coefficients.



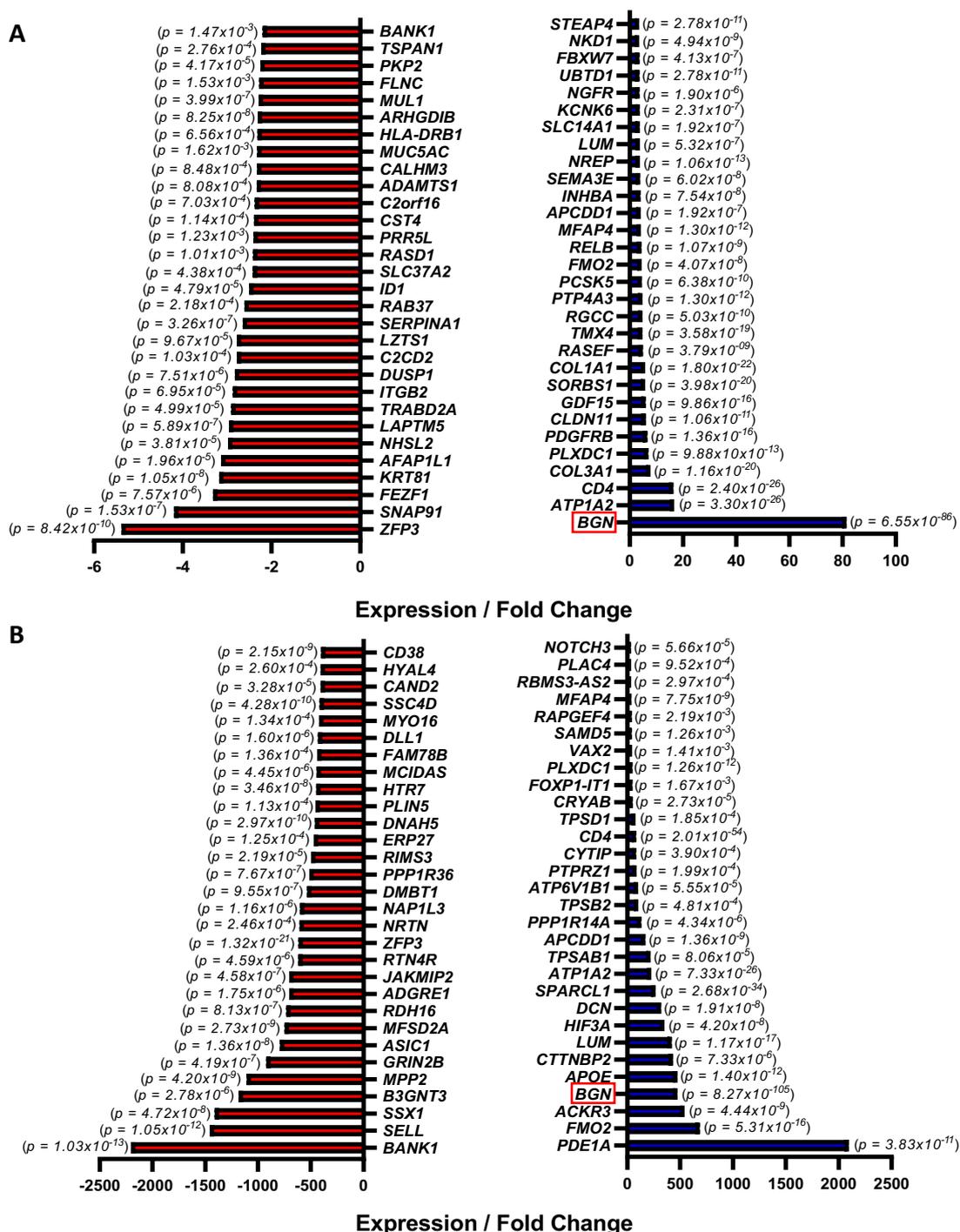
**Figure 3.14. Dormant and proliferating MDA-MB-231 cells display distinct gene expression profiles *in vivo*.** The heat map displays unsupervised hierarchical clustering of differentially expressed genes between dormant and proliferating MDA-MB-231 cells ( $FDR \leq 0.05$  cutoff). A) Heat map for genes identified by processing stream I; B) Heat map for genes identified by processing stream II. PCA plots comparing transcriptome clustering for dormant (blue) and proliferating (red) cancer cells ( $FDR \leq 0.05$  cutoff). C) PCA plot for genes identified by processing stream I; D) PCA plot for genes identified by processing stream II.

Unsupervised hierarchical clustering of differentially expressed genes (DEGs) ( $FDR \leq 0.05$ ) demonstrated distinct gene expression profiles of dormant and proliferating MDA-MB-231 cells derived from mouse brain tumour xenografts (**Figures 3.14A, B**). Observed was a bias for DEGs in processing stream I to be upregulated (**Figure 3.14A**). However, when aligned to a concatenated genome, this bias was depleted and there was a similar distribution between up- and downregulated genes (**Figure 3.14B**). Principle component analysis (PCA) of transcriptomes between samples displayed individual clustering of dormant and proliferating populations (**Figures 3.14C, D**). When compared to processing stream I (**Figure 3.14C**), alignment to a concatenated genome produced a much tighter cluster for the proliferating population (**Figure 3.13D**).

### 3.6 Analysis of Differentially Expressed Genes

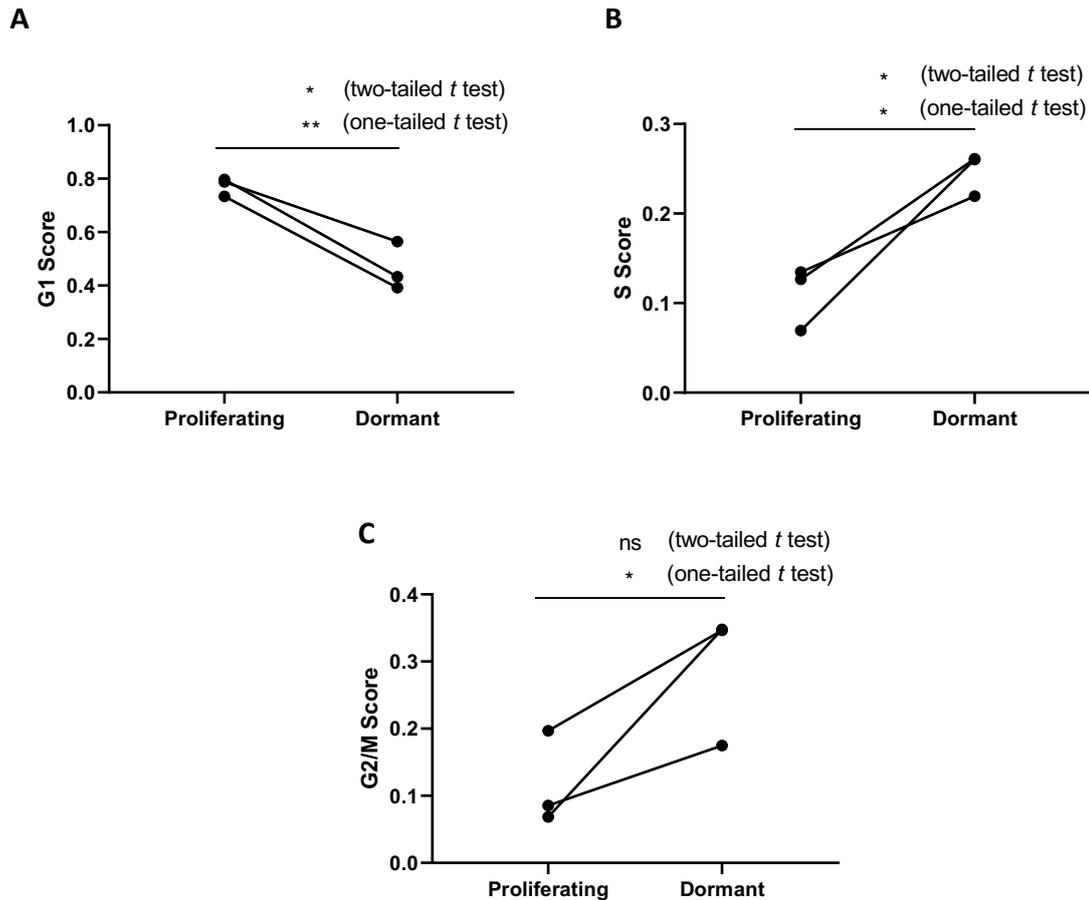
The top 30 differentially expressed genes between dormant and proliferating cell populations *in vivo* from processing stream I are presented in **Figure 3.15A** ( $FDR \leq 0.05$  cutoff). For processing stream II, these are presented in **Figure 3.15B**. The ECM protein biglycan (*BGN*) was the most significantly expressed gene in both sets of analyses, with fold changes of  $\sim 82$  (stream I) and  $\sim 472$  (stream II). Similarly, there was an almost complete absence of raw reads within the proliferating cell populations for biglycan (data not shown). Alignment to a concatenated human-mouse genome significantly increased the total number of differentially expressed genes in dormant and proliferating cells and this is reflected in the structure of the list of top 30 genes. For example, the gene with the largest fold change swapped from biglycan (*BGN*), which became the 4<sup>th</sup> largest, to phosphodiesterase 1A (*PDE1A*).

Given that *BGN* was identified as the most significantly differentially expressed gene in both datasets, and consistently ranked amongst the highest in terms of fold change, this gene was identified as a strong target for downstream analysis. In addition, cancer cells are highly responsive to ECM composition, and, as an ECM scaffold protein, biglycan may play an important signalling role in dormancy induction and maintenance [110, 117, 353].



**Figure 3.15.** Top 30 up- and down-regulated protein-coding genes in dormant versus proliferating MDA-MB-231 cells derived from mouse brain tumour xenografts. Genes were first ranked by significance, with a threshold of  $FDR \leq 0.05$ , then by fold change. Displayed are the top 30 genes with the largest fold changes in both directions of regulation. A) Genes identified via processing stream I. B) Genes identified via processing stream II. Statistical analysis was performed using R package *DESeq2*.

Finally, I performed statistical analyses of cell phase scoring using a two-tailed  $t$  test. In dormant cells, the phase with both the most significant and greatest numerical change in score, compared to proliferating cells, was G1 phase. In all three independent experiments, dormant cells were characterised by a significant reduction in G1 scoring (Figure 3.16A).

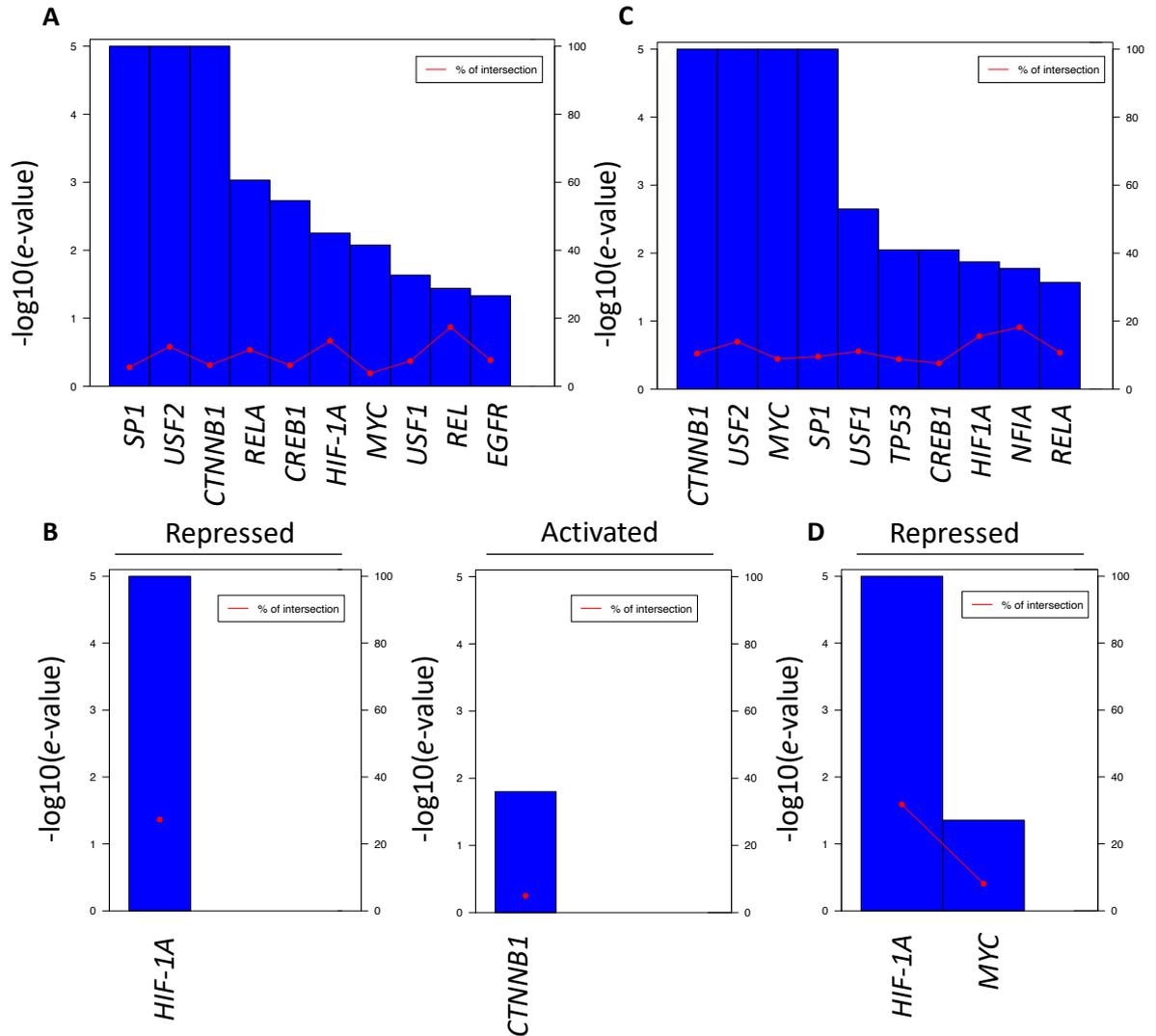


**Figure 3.16. Prediction of cell cycle phase in dormant breast adenocarcinoma cells *in vivo*.** Gene expression data of dormant and proliferating breast cancer cells *in vivo* was ranked and cut off according to significance ( $p \leq 0.05$ ). Shown is a prediction of cell cycle phase, A) G1, B) S, C) G2/M, for experiment-linked dormant and proliferating cells, according to an analysis of differentially expressed genes using the R package *Scran*. Statistical analyses were performed by unidirectional and bidirectional paired  $t$  tests.

In addition, dormant cells were also characterised by a significant increase in S phase (**Figure 3.16B**), whilst the increase in G2/M phase scoring was not significant (**Figure 3.16C**). However, when analysed by one-tailed *t* test, the increase in G2/M phase scoring was statistically significant. It is critical to consider that expression of gene pairs in MDA-MB-231 cells may not directly compare with expression of gene pairs in the *Scran* test set, given the huge range of biological variability between cell lines and species. It is also possible that at the time of arrest, gene pairs pertaining S phase may co-express with gene pairs pertaining to G2/M. However, given the significant increases in 2 out of 3 mid-cell cycle phase probabilities, as opposed to significant decreases in the dormant cell population in all 3, this indicates that the dormant cells I have isolated may not be arrested in G0 phase. One cannot draw any firm conclusions from this data; however, this opens up the possibility that dormant cells may arrest mid-cycle and warrants experimental investigation to support it.

### 3.6.1 Transcription Factor Analysis

Based on calculated fold changes of gene expression in dormant cells, potentially implicated transcription factors were predicted using TFactS software [244]. Both sign-less and sign-sensitive analysis were performed. Sign-less analyses were used to identify associations between inputted gene sets and the corresponding transcription factor. Sign-sensitive analyses provided an indication of the direction of regulation for a particular transcription factor; however, the reference database is manually annotated as data is published [244] and therefore may not be as complete as a sign-less analysis. All analyses were conducted using a threshold of  $FDR \leq 0.05$  for inputted gene lists as to provide the largest statistically significant gene pool for maximum transcription factor overlap. It was assumed that genes with relatively low fold changes in expression may still have significant functional effects, therefore no threshold was set for fold change.



**Figure 3.17. Predicted transcription factors implicated in breast cancer cell dormancy in the brain.** Genes differentially expressed between dormant and proliferating cancer cells were ranked according to statistical significance. A cut-off was made at  $FDR \leq 0.05$ . Up- and down-regulated genes were analysed by TFacts software [244]. A) Sign-less analysis (to identify implication only) of genes identified by alignment to the human genome. B) Sign-sensitive analysis, with statistically significant repressed and activated transcription factors. C) Sign-less analysis of genes identified by alignment to a merged human-mouse genome, with genes filtered for those exclusively aligning to human, with D) corresponding transcription factors identified as significantly repressed; no transcription factors were identified as activated. Gene expression data input was calculated from three independent experiments.  $e$ -value is given as the number of expected false positives with multiple testing, with a cut-off of  $\leq 0.05$ . Shown in red is the percentage at which a transcription factor was identified as significant within an  $e$ -value threshold of  $\leq 0.05$ .

Transcription factor	Associated downregulated genes in dormant cells	Associated upregulated genes in dormant cells
<b>c-Myc</b>	<p><i>LMNA, LTA4H, VDAC2, PCMT1, <b>LDHA</b>, ACTR3, DUSP1, <b>ENO1</b>, ETFDH, CCT7, AKR7A2</i> (Stream I)</p> <p><i>ACTR3, DUSP1, SLC25A3, <b>ENO2</b>, MCM3, PSMD8, MT1A, PPIF, CCT7, PCMT1, UNG, TYMS, AKR7A2, <b>PDHA1</b>, C1QBP, ACTG1, LGALS1, POLR1C, PTPN6, LTA4H, RPL15, RPL9, RPL8, RPL41, RPL7, HERPUD1, HSPE1, HS090AA1, <b>ENO1</b>, HLA-DPB1, SAP18, SCARB1, LMNA, FOSL1, VDAC2, <b>LDHA</b>, <b>GAPDH</b></i> (Stream II)</p>	<p><i>CDC16, ZFP36L2, MAN2A1, VHL, CYBA, AGPS, ATP1A2, TGFB2, SYMPK, SLC39A8</i> (Stream I)</p> <p><i>LAMP1, SLC36L2, VHL, ZFP36L2, AGPS, ZXDB, ATP1A2, SYMPK, TGFB2, CYBA, CDC16, MAN2A1</i> (Stream II)</p>
<b>HIF-1A</b>	<p><i><b>LDHA</b>, ITGB2, TIMP1, <b>ENO1</b>, <b>ALDOA</b>, <b>PGK1</b></i> (Stream I)</p> <p><i><b>GAPDH</b>, <b>LDHA</b>, <b>PGK1</b>, <b>ENO1</b>, <b>TIMP1</b>, <b>ALDOA</b>, <b>ITGB2</b></i> (Stream II)</p>	<p><i>No genes</i> (Stream I)</p> <p><i>No genes</i> (Stream II)</p>

**Table 3.4. Differentially expressed genes between dormant and proliferating breast cancer cells *in vivo* in the context of c-Myc and HIF-1A transcription factors.** Shown are the corresponding genes and their direction of regulation with respect to the repressed transcription factors c-Myc and HIF-1A. Genes pertaining to the aerobic glycolysis pathway are highlighted in red.

From the analysis of gene sets calculated from alignment to a single genome (processing stream I), 10 potentially implicated transcription factors were identified (**Figure 3.17A**). In comparison, 10 were also identified in gene sets calculated from alignment to a merged human-mouse genome (processing stream II) (**Figure 3.17B**). Overlapping between the two were specificity protein 1 (SP1), upstream stimulatory factor 2 (USF2),  $\beta$ -catenin (CTNNB1), transcription factor p65 (RELA), CAMP responsive element binding protein 1 (CREB1), hypoxia-inducible factor 1 subunit alpha (HIF-1A), c-Myc (MYC) and

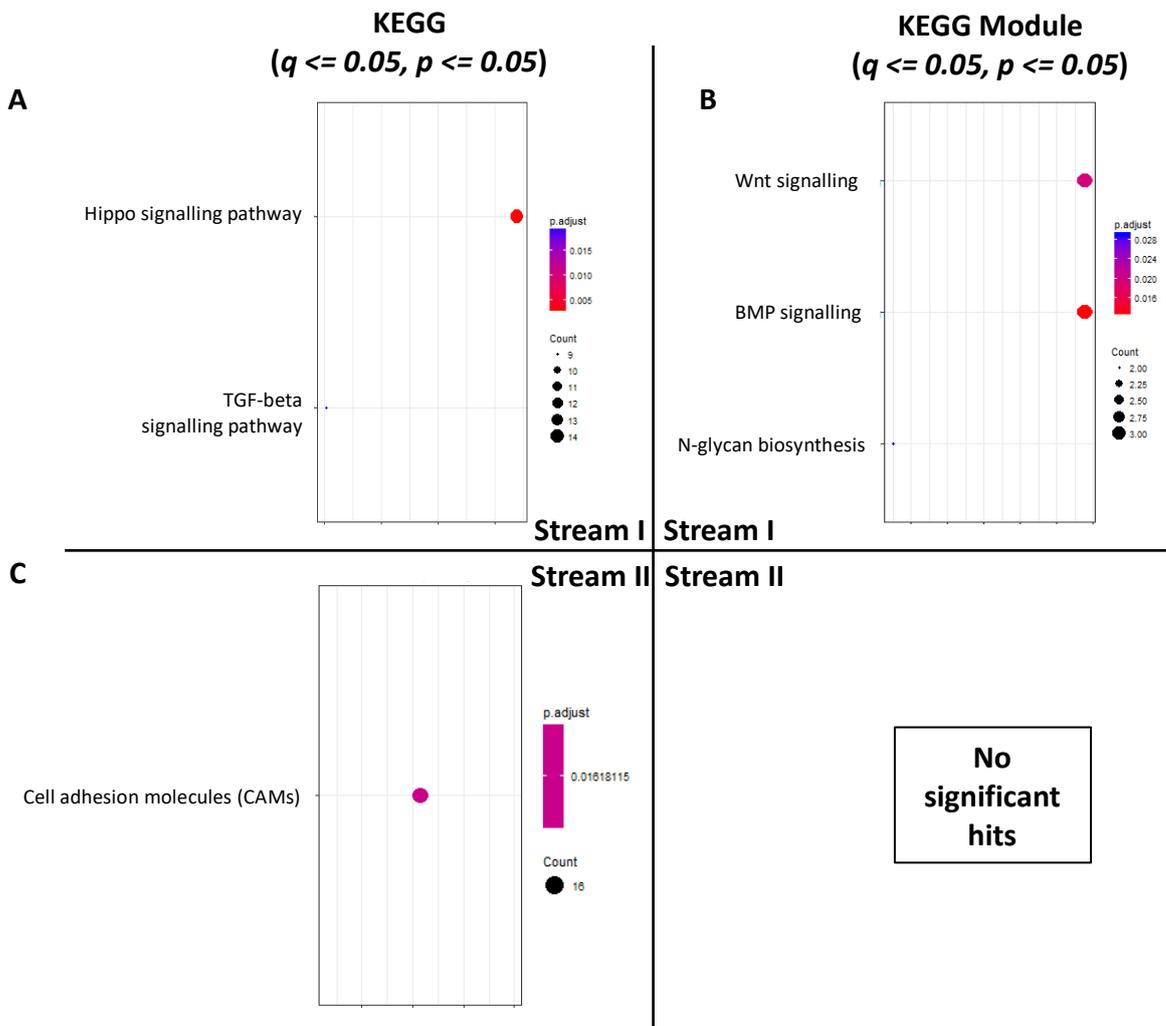
upstream stimulatory factor 1 (USF1). Non-overlapping transcription factors were early growth response protein 1 (EGR1) (stream I) and p53 (TP53) (stream II). Based on published data [244], it was determined that in stream 1 data, HIF-1A and CTNNB1 were identified as significantly repressed and activated, respectively (**Figure 3.17C**). In stream II data, there were no transcription factors identified as activated, however HIF-1A and MYC were identified as repressed in dormant cells *in vivo* (**Figure 3.17D**). Based on these observations, DEGs associated with MYC and HIF1A were analysed further (**Table 3.4**). Interestingly, many genes within the aerobic glycolysis pathway were all found to be downregulated in dormant breast cancer cells.

### 3.6.2 Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Enrichment Analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of DEGs between dormant and proliferating breast cancer cells was performed for both analysis streams. KEGG is a community-generated resource built upon genomic, chemical and functional data. It allows for a deeper analysis of sequencing outputs as it allows the user to map significant data in the form of functional networks and pathways [266]. For such analysis, genes were pre-filtered to a cutoff of FDR  $\leq 0.05$ , fold change (FC)  $\geq 2$ . This was to restrict the gene pool as too many genes may obscure legitimately enriched pathways which contain few genes [267].

For processing stream I, results were filtered for q and p values of 0.05. Between the KEGG (**Figure 3.18A**) and KEGG module (**Figure 3.18B**) analyses, the Hippo, TGF- $\beta$  (with the related BMP signalling pathway) and Wnt signalling were identified as enriched. When applied to genes identified by processing stream II, no enriched KEGG modules were observed, however cellular adhesion molecules (CAMs) were observed as a KEGG output (**Figure 3.18C**). The q and p value restrictions on output terms were then removed, to identify the top 20 of all possible KEGG pathways and KEGG modules for processing stream II analysis. Interestingly, as well as cell adhesion molecules, genes pertaining to Hippo, Wnt and TGF-beta signalling were all identified by KEGG (**Figure 3.19A**). When KEGG modules were analysed, however, such pathways were not

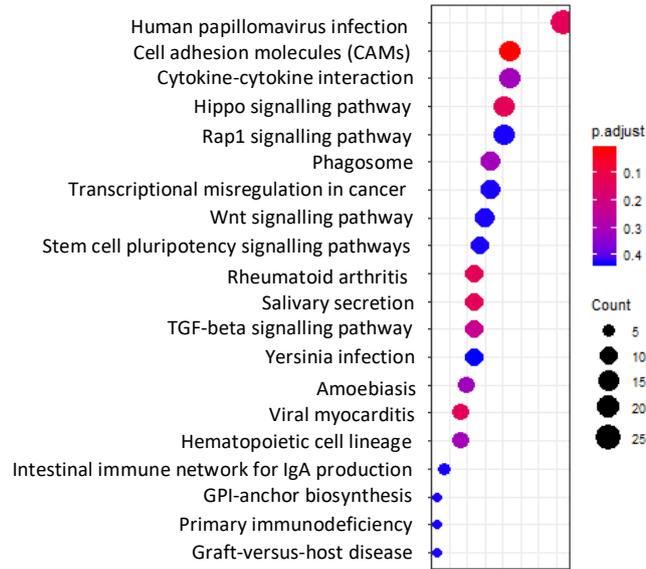
identified, however there were terms relating to cellular biosynthesis and metabolism (**Figure 3.19B**). Given such observations, a network map was constructed using gene expression data from processing stream II, to produce an overview of the potential crosstalk between TGF- $\beta$ , Wnt and Hippo pathways. As shown in **Figure 3.20**, all three pathways converge on the transcriptional regulator, YAP.



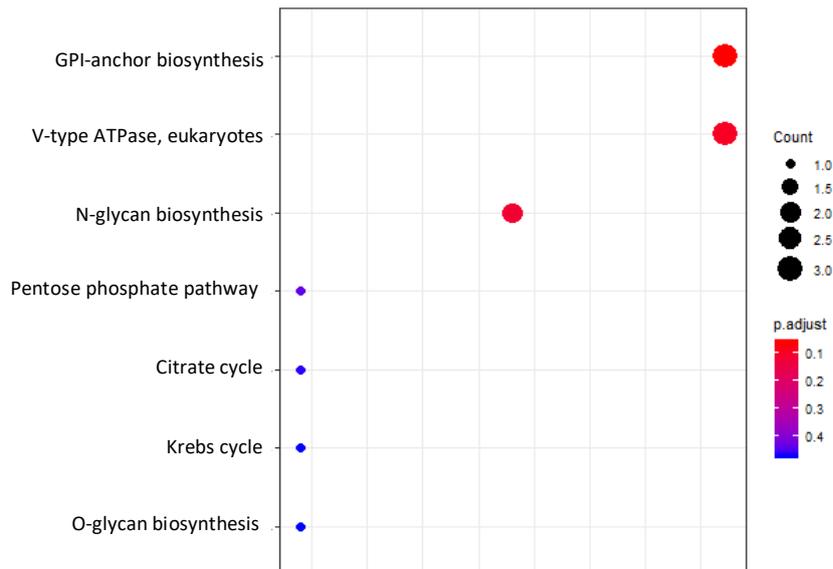
**Figure 3.18. KEGG and KEGG module enrichment analysis (processing streams I and II).** KEGG (A, C) and KEGG module (B) analyses were formed by R package ClusterProfiler, with a pre-set list of differentially expressed genes from both processing streams. A, B) Processing stream I ( $q$  and  $p$  value cutoff for term output was set at 0.05). C) Processing stream II ( $q$  and  $p$  value cutoff for term output was set at 0.05). Dormancy input genes were identified by  $FDR \leq 0.05$ ,  $fold\ change \geq 2$ . Functionally implicated pathways and networks are displayed. The size of the circle indicates the number of genes identified within the pathway or network, whilst the colour represents statistical significance.

## Stream II

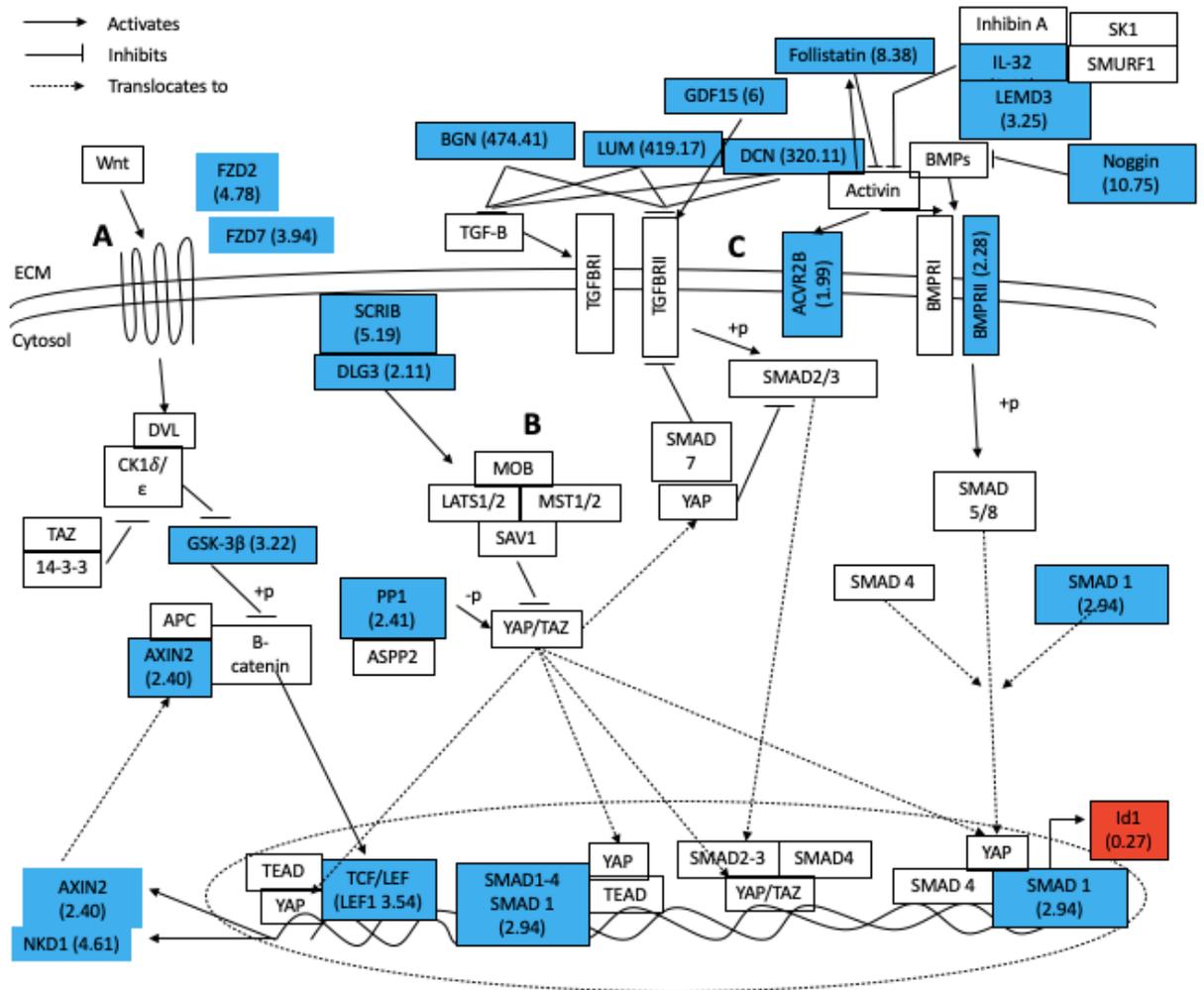
A



B



**Figure 3.19. KEGG and KEGG module enrichment analysis (processing stream II) with no statistical cutoff on output terms.** KEGG (A) and KEGG module (B) analyses were formed by R package ClusterProfiler, with a pre-set list of differentially expressed genes from both processing streams. No  $q$  and  $p$  value cutoff was employed for output terms. Dormancy input genes were identified by  $FDR \leq 0.05$ ,  $fold\ change \geq 2$ . Functionally implicated pathways and networks are displayed. The size of the circle indicates the number of genes identified within the pathway or network, whilst the colour represents statistical significance. Circles further to the right also correspond to circle size: they contain a higher number of associated genes.

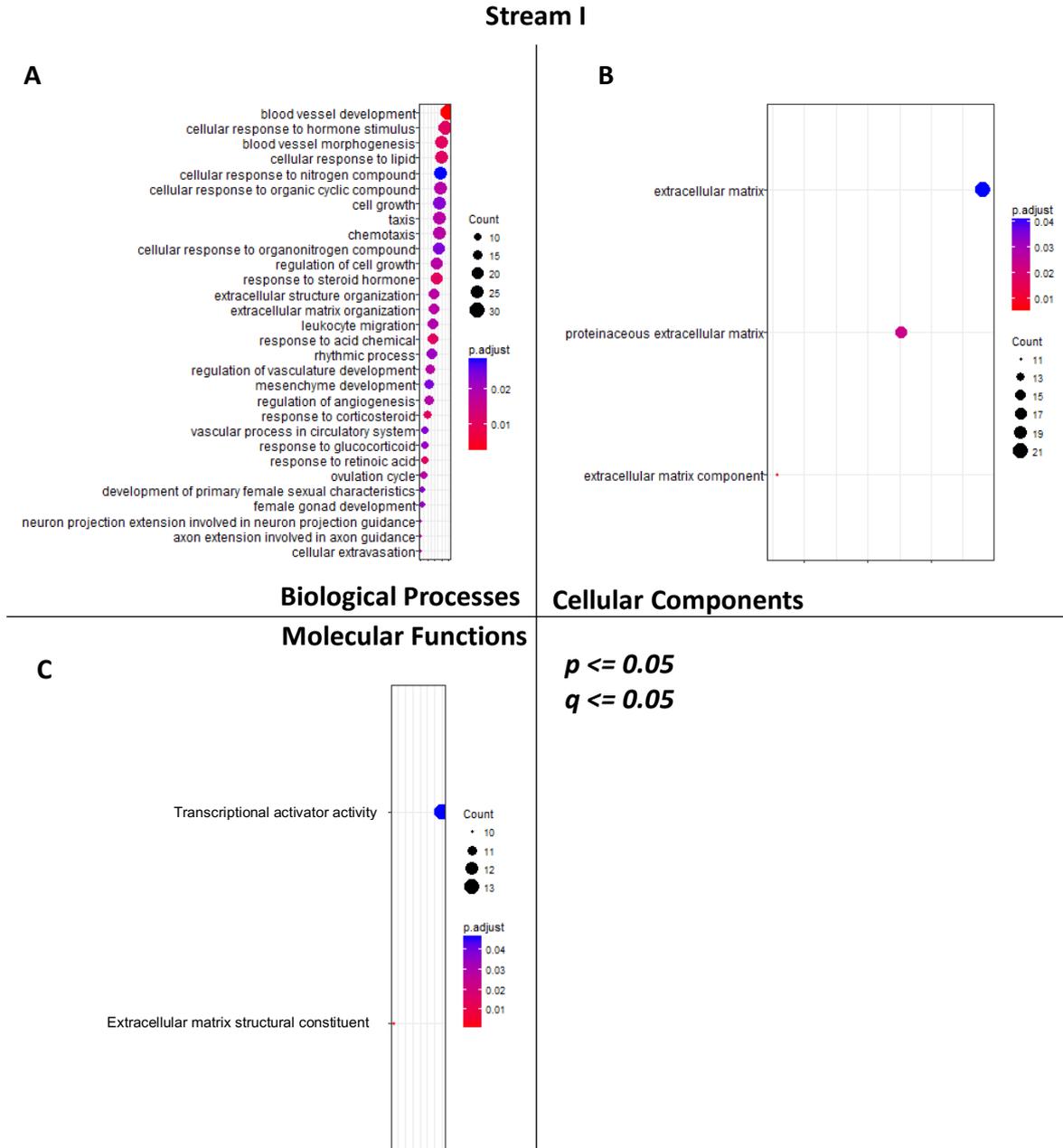


**Figure 3.20. Potential crosstalk between Wnt, HIPPO and TGF- $\beta$  signalling pathways in dormant MDA-MB-231 cells *in vivo*.** Differentially expressed genes in dormant MDA-MB-231 cells, identified by processing stream II (FDR  $\leq$  0.05, fold change  $\geq$  2) were analysed in the context of their respective pathways. A) Canonical Wnt signalling pathway. B) HIPPO signalling pathway. C) TGF- $\beta$  signalling pathway. Given are proteins with their respective gene expression fold change in dormancy, representative of three independent experiments. Blue = genes upregulated in dormant cells. Red = genes downregulated in dormant cells. Abbreviations: Wingless-related integration site (Wnt), frizzled (FZD), dishevelled (DVL), casein kinase (CK), tafazzin (TAZ), glycogen synthase kinase (GSK), adenomatous polyposis coli (APC), axis inhibition protein (AXIN), naked cuticle homolog (NKD), biglycan (BGN), lumican (LUM), decorin (DCN), growth/differentiation factor (GDF), interleukin (IL), sphingosine kinase (SK), SMAD specific E3 ubiquitin protein ligase (SMURF), LEM domain-containing protein (LEMD), bone morphogenic protein receptor (BMPR), activin A type IIB receptor (ACVR2B), yes-associated protein (YAP), transforming growth factor (TGF), scribble (SCRIB), discs large homolog (DLG), protein phosphatase (PP), transcription factor (TCF), lymphoid enhancer-binding factor (LEF), Salvador homolog (SAV), macrophage stimulating (MST), large tumour suppressor kinase (LATS), inhibitor of DNA binding (ID), TEA domain family member (TEAD).

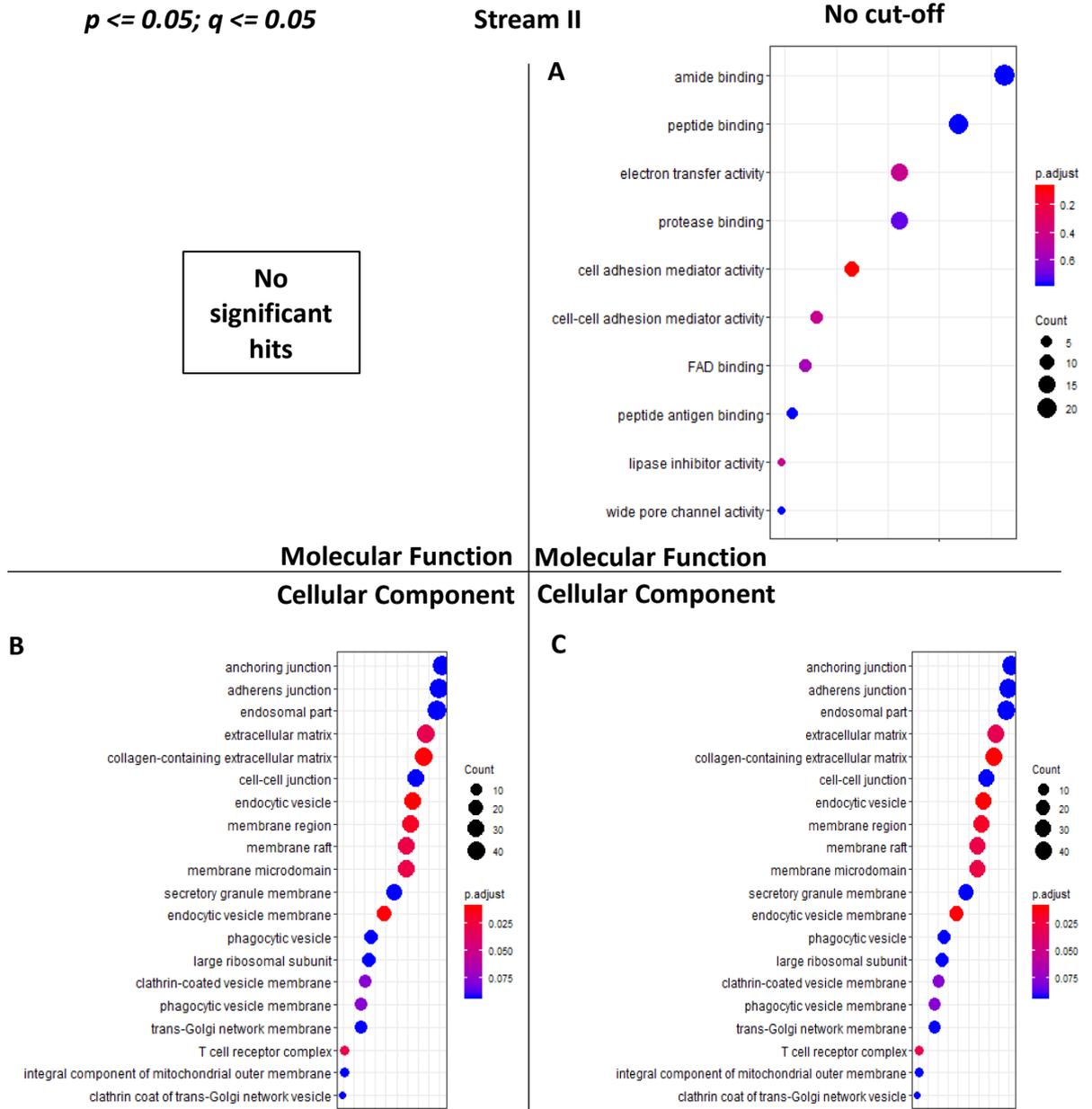
In addition, Gene Ontology (GO) terms were analysed for biological processes, cellular components and molecular functions, using output thresholds for each term set at  $p$  and  $q$  values of 0.05. The same threshold ( $FDR \leq 0.05$ , fold change  $\geq 2$ ) was applied prior to analysis for all input DEGs. Many of the enriched biological processes (processing stream I) could be grouped and it appears that the majority of over-expressed terms within the dormant population pertained to genes involved in regulation of angiogenesis, ECM organisation and hormone response to stimuli (**Figure 3.21A**). The ECM, in terms of enriched constituents, was also enriched within the cellular component (**Figure 3.21B**) and molecular functions (**Figure 3.21C**) categories, and thus was chosen as a focus in downstream analyses.

When repeated using terms identified by processing stream II, there were no enriched molecular functions (MFs) observed when using  $p$  and  $q$  value thresholds of 0.05 for term outputs. However, under the same output parameters, enriched cellular components pertained to the extracellular matrix, membrane components and cell junctions (**Figure 3.22B**). For biological processes, terms that were enriched pertained highly to metabolic and biosynthetic processes (**Figure 3.23A**). For cellular component and biological processes terms, removal of the  $q$  and  $p$  value threshold for output terms did not alter the top 20 terms identified (**Figures 3.22B, 3.23B**, respectively). However, removal of such parameters uncovered some molecular functions, which were generally pertaining to cell adhesion (**Figure 3.22A**).

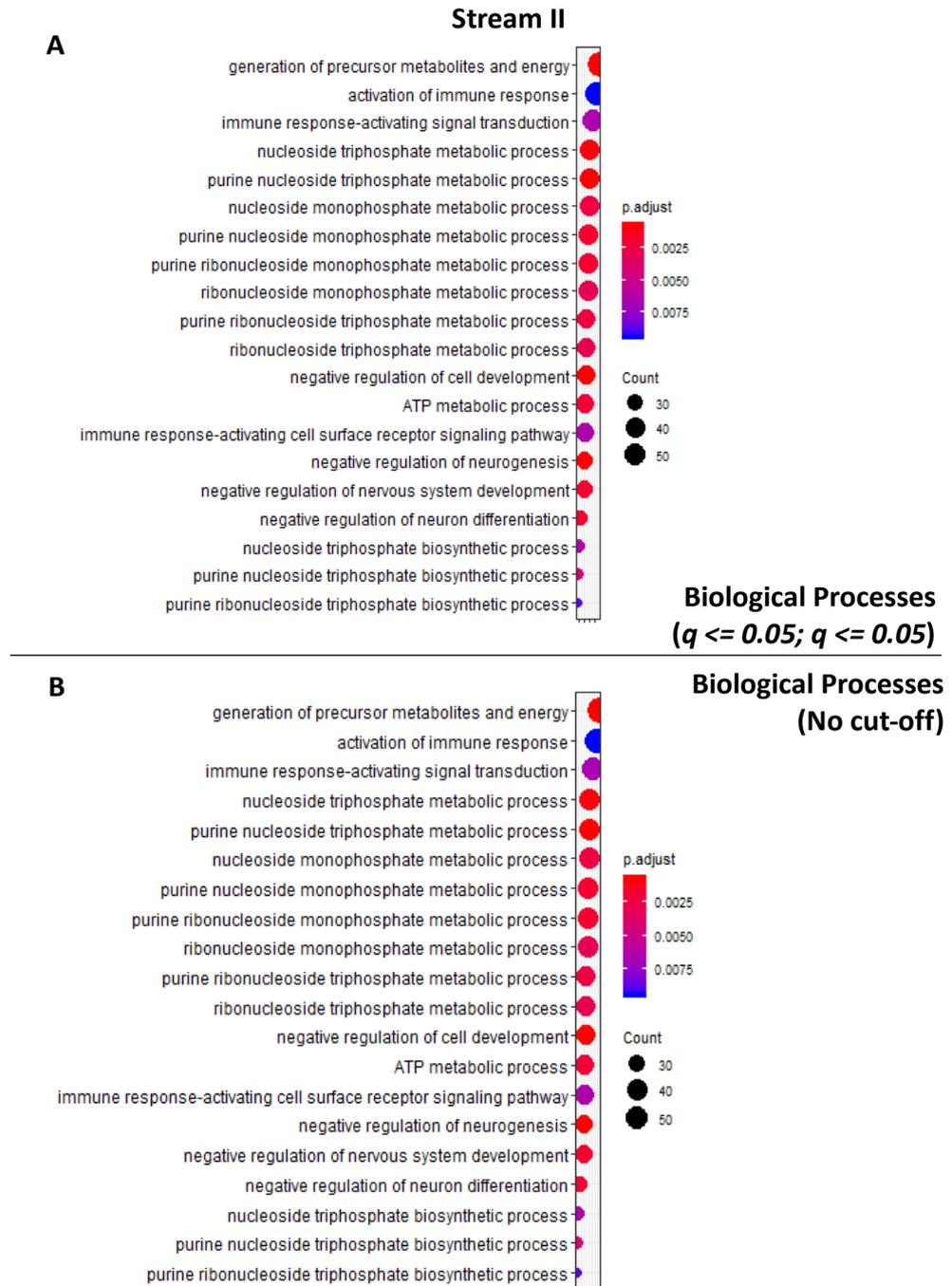
Given that many of the identified ontologies pertain to ECM components and signalling, *BGN*, an ECM component, remains a prominent target of interest due to its high level of significance of differential expression and large fold change in dormant cells. Another downstream target for further analysis is YAP, given that three top ontology hits in dormant cells, notably Wnt, HIPPO and TGF- $\beta$  signalling, all converge on this transcriptional activator. Finally, given a predicted repression of HIF-1A and MYC transcription factors, which are crucial for maintaining glycolysis in cancer cells [336], and an enrichment of terms pertaining to metabolic processes was identified in dormant cells, aerobic glycolysis will be explored further as a target of interest.



**Figure 3.21. Gene Ontology (GO) enrichment analysis (processing stream I).** DEGs from dormant MDA-MB-231 cells (processing stream I) were used as input to determine the top 20 GO terms in A) biological processes, and all terms enriched in B) cellular component, and C) molecular function categories. Analysis performed by R package *ClusterProfiler*. DEG cutoff at FDR  $\leq 0.05$ , fold change  $> 2$ . Output thresholds for each term set at p and q values of 0.05. Functionally implicated GO terms are displayed. The size of the circle indicates the number of genes identified within the GO category, whilst the colour represents statistical significance. Circles further to the right also correspond to circle size: they contain a higher number of associated genes.



**Figure 3.22. Gene Ontology (GO) enrichment analysis (processing stream II).** Molecular function (A) and cellular component (B, C) analyses were performed by R package ClusterProfiler, using a pre-set list of differentially expressed genes from processing stream II. B)  $q$  and  $p$  value thresholds for each term were set at 0.05. A, C: no  $q$  and  $p$  value threshold for term output. Dormancy input genes were identified by  $FDR \leq 0.05$ ,  $fold\ change \geq 2$ . Functionally implicated GO terms are displayed. The size of the circle indicates the number of genes identified within the GO category, whilst the colour represents statistical significance. Circles further to the right also correspond to circle size: they contain a higher number of associated genes.

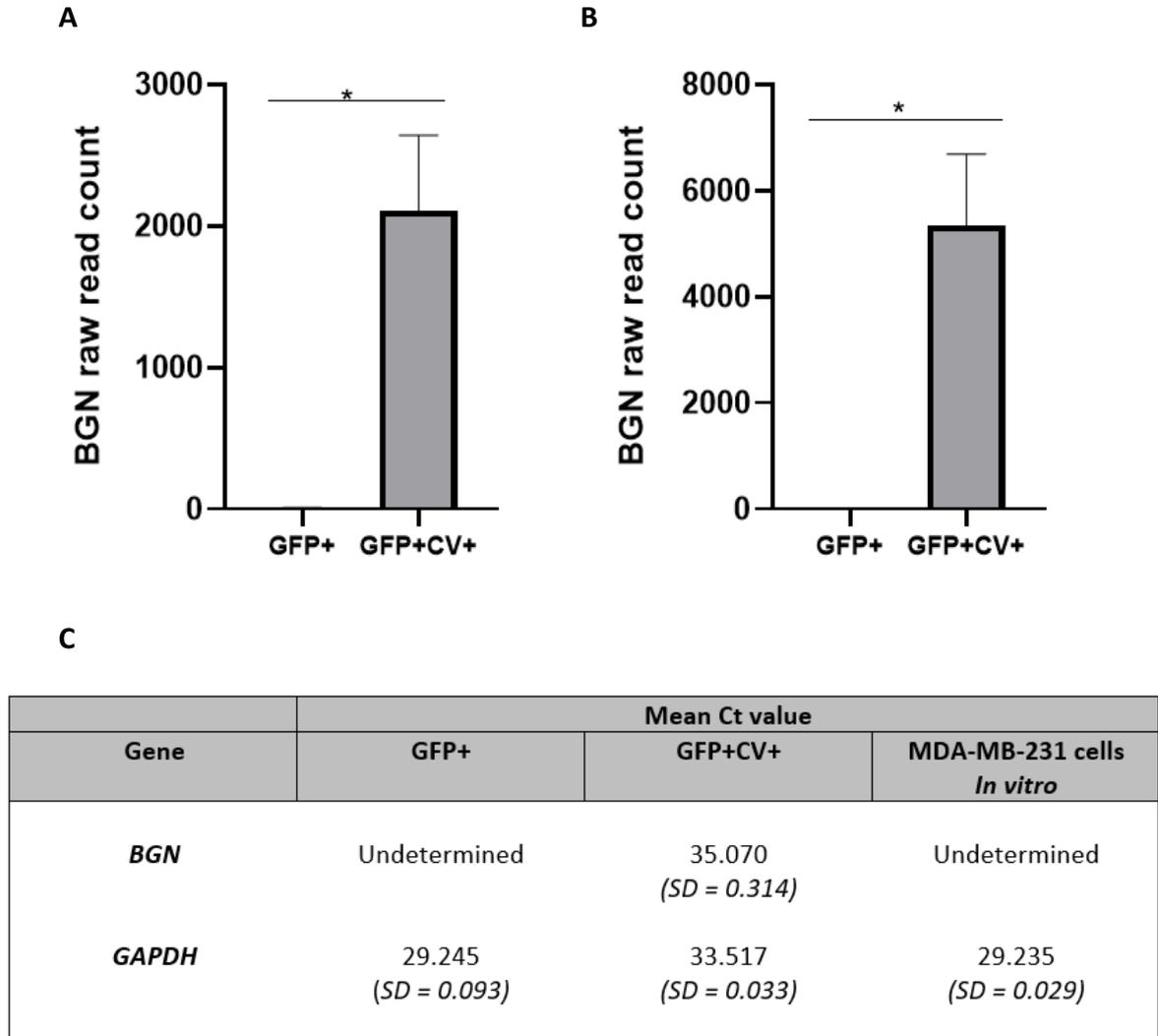


**Figure 3.23. Gene Ontology (GO) enrichment analysis (processing stream II) with no statistical cutoff on output terms.** Biological Processes analyses were performed by R package ClusterProfiler, using a pre-set list of differentially expressed genes from processing stream II. A)  $q$  and  $p$  value thresholds for each term were set at 0.05. B) no  $q$  and  $p$  value threshold for term output. Dormancy input genes were identified by  $FDR \leq 0.05$ ,  $fold\ change \geq 2$ . Functionally implicated GO terms are displayed. The size of the circle indicates the number of genes identified within the GO category, whilst the colour represents statistical significance. Circles further to the right also correspond to circle size: they contain a higher number of associated genes.

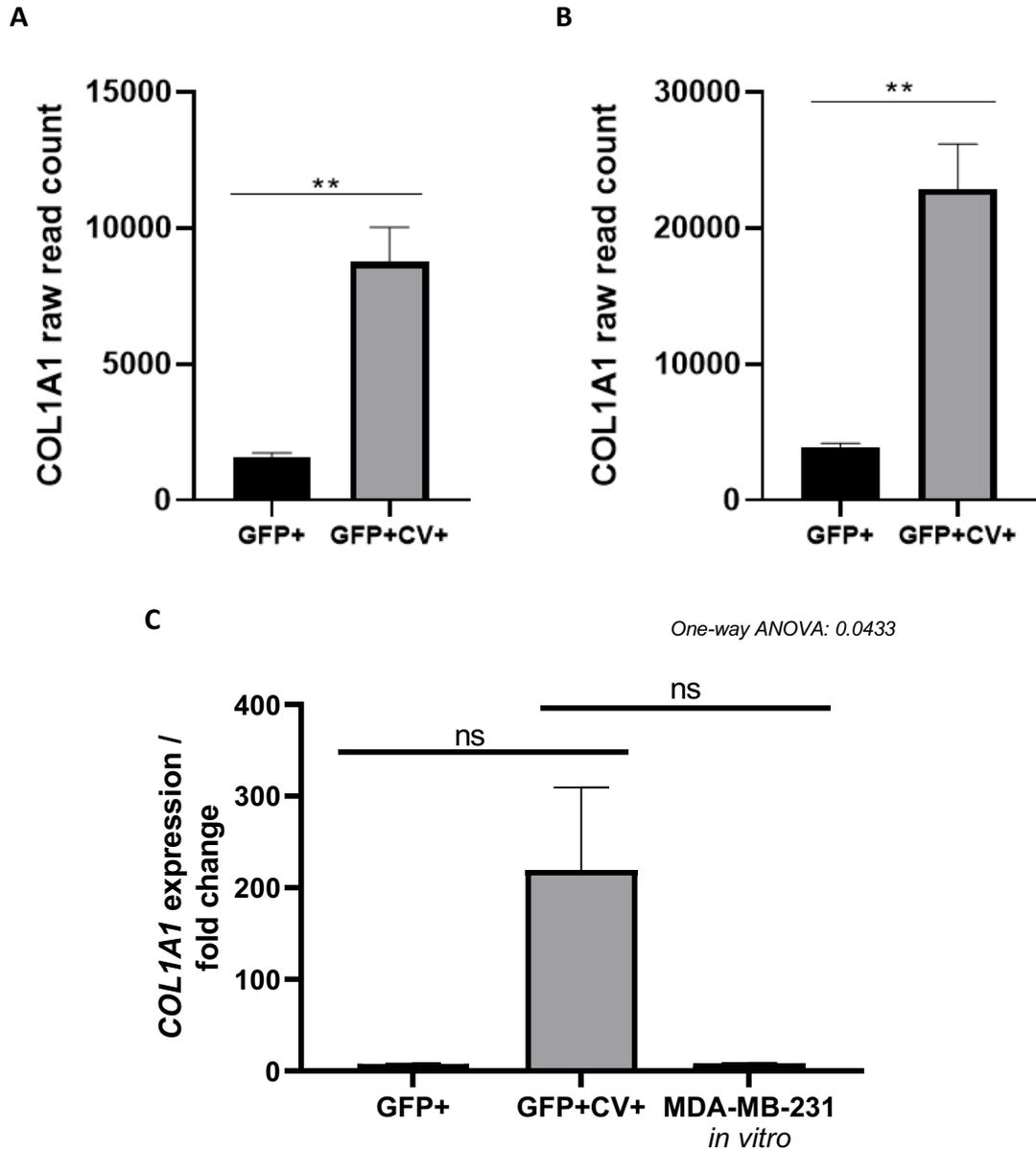
### 3.7 Validation of mRNAseq Data

Using cDNA derived from dormant and proliferating cancer cells, the gene expression of biglycan and collagen I (*COL1A1*) was determined by qPCR. In dormant cells *in vivo*, an 81.56- and 474.41-fold change in biglycan expression was observed, using analysis streams I and II, respectively. In terms of raw reads, there was an almost complete absence within the proliferating population, with a significantly higher number of reads within the dormant population for both genes (**Fig 3.24A**). When the expression was determined by qPCR, no biglycan was detected within the proliferating cell population, as expected (**Fig 3.24B**). It may be possible that a small amount of biglycan cDNA was present, however this would have been outside the sensitivity range of this assay. However, biglycan was detected within the dormant cell population (**Fig 3.24B**), confirming my differential gene expression analyses. Finally, analysis of biglycan expression in MDA-MB-231 cells cultured *in vitro* revealed an absence of biglycan (**Fig 3.24B**). Biglycan expression *in vivo* in proliferating cells mirrors the expression of cells cultured *in vitro*.

For *COL1A1* gene expression, a fold change in expression of 5.80- and 7.63-fold was observed in the dormant cell population *in vivo*, using analysis streams I and II, respectively. The pattern of raw read counts *in vivo* were such that a low number of reads for *COL1A1* were detected in the proliferating cell population, with significantly higher read count in dormant cells (**Fig 3.25A**). Similar patterns of expression were observed when validated by qPCR (**Fig 3.25B**). The fold change in dormant cells was 219.64. The expression of biglycan in MDA-MB-231 cells cultured *in vitro* was almost identical to that of proliferating cells *in vitro*. This, along with patterns of biglycan expression, validated my mRNAseq data as correct and supports the distinct gene expression profiles displayed by dormant versus proliferating breast cancer cells in this model.



**Figure 3.24. Biglycan expression in proliferating and dormant MDA-MB-231 cells *in vivo*, compared to cells cultured *in vitro*.** The raw read counts for biglycan expression in demonstration of mRNA abundance in proliferating (GFP+) and dormant (GFP+CV+) cells *in vivo* are shown. Biglycan read counts processing methods I and II (A and B, respectively) +/- SEM. C) Descriptive statistics of Ct values obtained from qPCR analysis of cDNA derived from proliferating and dormant cells. Measured alongside was the relative expression of biglycan in MDA-MB-231 cells cultured *in vitro*. The housekeeping reference gene was *GAPDH*.  $N = 3$ . Data calculated as a mean across biological triplicates from three independent experiments.



**Figure 3.25. Collagen I expression in proliferating and dormant MDA-MB-231 cells *in vivo*, compared to cells cultured *in vitro*.** The raw read counts for collagen I expression in demonstration of mRNA abundance in proliferating (GFP+) and dormant (GFP+CV+) cells *in vivo* are shown. Collagen I read counts processing streams I and II (A and B, respectively) +/- SEM. C) qPCR analysis of cDNA derived from proliferating and dormant cells. Measured alongside was the relative expression of collagen I in MDA-MB-231 cells cultured *in vitro*. The housekeeping reference gene was *GAPDH*.  $N = 3$ . Data calculated as a mean across biological triplicates from three independent experiments +/- SEM. Statistics is a one-way ordinary analysis of variance (ANOVA) with multiple comparisons between all groups.

### 3.8 Discussion

In this chapter, I determined the gene expression profiles of dormant and proliferating breast adenocarcinoma cells, derived from mouse brain tumour xenografts. The results suggested that dormant and proliferating cancer cell populations display distinct transcriptomes *in vivo*. GO and KEGG enrichment analysis between the two cancer cell populations identified terms pertaining to HIPPO, Wnt and TGF- $\beta$  signalling pathways, amongst a number of other terms. This was accompanied by an enrichment in terms pertaining to ECM constituents and metabolic processes. The most significantly upregulated gene in dormant cells was the proteoglycan, biglycan. Furthermore, the HIPPO, Wnt and TGF- $\beta$  signalling pathways all possess extracellular-derived activators, implicating the brain microenvironment in dormancy induction. Transcription factor analysis identified a suppression in c-Myc and HIF-1A transcription factors, which was associated with the downregulation of a number of enzymes which are critical for glycolysis in cancer cells. One of the prominent hallmarks of cancer is high levels of aerobic glycolysis, which strongly facilitates tumour growth [268], therefore a downregulation of genes encoding enzymes within the glycolysis pathway may inhibit metastatic outgrowth upon dissemination and induce a state of dormancy. As such, these findings may offer a novel insight into the pathways which underpin the induction and/or maintenance of breast cancer cell dormancy in the brain.

#### 3.8.1 Identification and Isolation of Dormant and Proliferating Breast Adenocarcinoma Cells

An approach was needed to reliably distinguish dormant and proliferating breast cancer cells *in vivo*. It was important to consider whether the identified dormant cells were indeed true dormant cells. I have shown *in vitro* that CV dye is lost rapidly through proliferation. Given that CV dye forms a stable covalent bond with cell surface proteins [269], cells which retained the dye *in vivo* were likely to be growth-arrested, and that those which lost the dye had done so via proliferation, and not significantly by diffusion of the dye into the extracellular space/other cells. A current problem in dormancy

research is that it is not clear at which point in the cell cycle dormancy occurs, for example G0/G1, mid-cycle, or actively cycling but happens to be absent for a particular expression marker during the time point cells were captured. Ki-67, a marker for proliferation, is absent in G1 phase, and may mis-identify cycling cells as dormant [270]. Similarly, proliferating cell nuclear antigen (PCNA) levels are very low in quiescence and fluctuate throughout the cell cycle, peaking in S phase [271]. Minichromosome maintenance (MCM) proteins are constitutively present within the nucleus throughout all phases. They bind chromatin in G1 phase and are displaced in S phase through to M phase, at which point they re-bind chromatin prior to telekinesis [271]. The advantage of my approach, in the absence of consensus markers for dormancy, is that it provided a broad overview of the proliferation status of individual cancer cells, from the point of injection to harvest, regardless of the phase in which they were arrested. The following discussion will determine the extent to which my data, and this model, align with current dormancy literature.

### **3.8.2 mRNAseq and Differential Gene Expression Analysis**

Despite alignment to a concatenated human-mouse genome, processing stream II still presented with some contaminating mouse-derived reads in the dormant population, as with processing stream I. The degree of contamination negatively correlated with the number of isolated dormant cells. STAR aligner is quoted to be highly accurate in mapping reads from single cells [272], however there is currently no published data which describes the most suitable method for the isolation and sequencing of a rare and small population of cells from xenograft tissue. It is already described that cross-species contamination is always an issue in which tumour cells represent a small RNA component, compared to that of host tissue [273]. In future studies, it may be possible to use my data to generate a higher number of dormant cells to overcome this problem. I believe mouse contamination in this study, at least in terms of known MDA-MB-231 SNPs, was mitigated sufficiently, given that similar SNP correlation coefficients were obtained between all samples, for both proliferating and dormant cells.

Agreement between SNPs was also reflected in the GO analyses of both processing streams. In stream I, ECM organisation, angiogenesis and steroid hormone signalling were amongst enriched terms. Under identical selection parameters, GO analysis identified terms pertaining to ECM constituents, membrane components and cell junctions were identified in stream II analyses. It was interesting that ECM constituents were identified in both streams. Adherence to the ECM has been shown to be critical for cancer cell growth and frequently implicated in dormancy induction [274]. Whilst in line with the literature, further study should consider extracellular influence on dormancy induction, in this model. The ECM is highly influential on cancer progression, interacting with cancer cells via cell-matrix adhesion complexes, and facilitates biophysical signalling [104]. A caveat for this analysis is that processing stream II created a much tighter cluster of transcriptomes in the proliferating cell populations, however this created a looser cluster for the dormant cell populations. This, and differences in GO terms, may simply be due to the size of gene input lists between processing streams, and as such the focus should be on where processing streams are convergent.

### 3.8.3 Comparison to Known Dormancy Mechanisms

Shown in **Table 3.5** is a 48-gene signature for tumour cell dormancy, based on combined datasets derived from tumour cell quiescence- or angiogenic failure-induced dormancy [196, 275]. Of the 22 previously defined genes thought to be upregulated in dormancy, 6 genes were in agreement with my data, whilst I observed another 2 genes which were significantly downregulated. Of the 26 previously defined genes thought to be downregulated in dormancy, 2 genes were in agreement with my data, whilst I observed another 6 genes which were significantly upregulated. It is important to recognise that whilst there may be little agreement, there are a number of caveats. There are no complete datasets using a similar model pertaining to breast cancer cell dormancy mechanisms within the brain, I must draw my comparisons from pooled data across a number of cancers, differential sites of metastasis and types of dormancy. However, dormancy induction and maintenance may be dependent on cell type, context, and a number of intersecting pathways. It is promising that whilst I have conflicting directions of expression, there is agreement between which genes are differentially expressed in

dormant cells. In my gene expression data, I have identified a number of genes and ontologies pertaining to the ECM and ECM remodelling. In the published dataset, alongside an upregulation of *COL1A1* and *COL4A5*, there is an upregulation of a number of ECM remodelling genes: *ADAM10*, *DDR1*, *GATA6*, *IGFBP5*, *MMP2*, *P4HA1* and *SOX9* [196, 275, 276]. In some cases, where I may not have directly observed the differential expression of a particular gene, I have evidence of the activity of its protein product. For example, *AMOT* has previously been shown to be upregulated in dormancy. AMOT proteins can bind YAP directly to negate its activity [196, 275, 276]. In my dormant cells, I predicted activation of the HIPPO pathway, which also negates YAP activity. Furthermore, I predicted enrichment of the TGF- $\beta$  signalling pathway. In addition to *TGFB2* previously being observed to be upregulated in dormancy, and being upregulated in the published dataset, *SREBF1* has previously been shown to be upregulated in dormancy. One of the splice variants upon transcription is SREBP1, which can bind to the promotor region of TGF- $\beta$  to induce its expression [196, 275, 276]. Alignment between my model and published datasets may not be immediately clear, however there is enough to provide some initial validity to my model.

Dormant cell status	Gene symbols
Upregulated	<i>ACVR1</i> , <i>ADAM10</i> , <i>AMOT</i> , <i>BHLHE41</i> , <i>COL1A1</i> , <i>COL4A5</i> , <i>CTSD</i> , <i>DDR1</i> , <i>EPHA5</i> , <i>GATA6</i> , <i>HIST12BK</i> , <i>IGFBP5</i> , <i>MMP2</i> , <i>NR2F1</i> , <i>P4HA1</i> , <i>SOX9</i> , <i>SREBF1</i> , <i>STAT3</i> , <i>TGFB2</i> , <i>THBS1</i> , <i>TP53</i> , <i>TPM1</i> [196, 275, 276]
Downregulated	<i>APEX1</i> , <i>ASNS</i> , <i>ATF3</i> , <i>ATF4</i> , <i>BUB1</i> , <i>BUB1B</i> , <i>CDKN3</i> , <i>CEBPG</i> , <i>CKS2</i> , <i>DNMT1</i> , <i>DTYMK</i> , <i>EGFR</i> , <i>EGR1</i> , <i>ESM1</i> , <i>FOSL1</i> , <i>FOXD1</i> , <i>FOXM1</i> , <i>IGF1R</i> , <i>IL8</i> , <i>JUN</i> , <i>MMP1</i> , <i>ODC1</i> , <i>PIK3CB</i> , <i>PLAT</i> , <i>TIMP3</i> , <i>TK1</i> [196, 275, 276]

**Table 3.5. Comparison to known dormancy signatures by gene expression.** Based on models of *in vivo* dormancy driven by tumour cell quiescence [196] or angiogenic failure [275], a 49-gene signature for tumour cell dormancy was constructed. Genes upregulated in dormant cells were considered as positive dormancy genes, whilst those downregulated in dormant cells were considered as negative dormancy genes [276]. Genes highlighted in green where the published dataset and my model are in alignment, whereas red indicates agreement in differential expression only, not direction of regulation i.e. up- or down-regulated.

### 3.8.4 Enriched Molecular Pathways in Dormant Breast Cancer Cells *in vivo*

In dormant breast cancer cells *in vivo*, KEGG analysis of differentially expressed genes suggested activation of the HIPPO pathway. Activation of the HIPPO pathway was evidenced in my model by upregulation of *SCRIB* in dormant cells. *SCRIB* is a scaffold protein which forms a complex with discs large 1-4 (*DLG1-4*) and is concentrated at cell junctions [277]. In a mouse model of epidermal carcinogenesis, *SCRIB* has been shown to be a potent tumour suppressor [277]. Similarly, *DLG3* was upregulated in the dormant population, thus the two protein may play a role in dormancy through HIPPO pathway activation. Pharmacological inhibition of YAP/TAZ has previously been suggested to be a potential anti-cancer strategy [278]. If I can confirm that the HIPPO pathway is indeed activated in my dormant cell population, by analysing the localisation and phosphorylation state of YAP/TAZ, then it would be worth exploring some available options to target YAP/TAZ. There are a number of small-molecule mediators which indirectly or directly target YAP/TAZ through the HIPPO pathway, however many of these lack specificity and long-term administration may result in unfavourable side-effects [279]. One example of Pazopanib, which activates the HIPPO pathway and induces proteasomal degradation of YAP/TAZ, sensitising those cells to anti-cancer drugs [280].

With respect to TGF- $\beta$  signalling, in the dormant cell population *in vivo*, a significant upregulation in inhibin beta A chain precursor (*INHBA*) expression was observed. *INHBA* encodes a subunit of both activin and inhibin [281]. Previous experimental data has demonstrated G1-phase cell cycle arrest in MCF-7 breast cancer cells through increased activin signalling [282]. Increased activin signalling induces a negative feedback loop, through upregulation of the activin inhibitor, follistatin [283], which was also upregulated in the dormant cell population in my model *in vivo*. It is thus possible that increased *INHBA* expression relates to increased activin activity and dormancy. Furthermore, a significant upregulation of Noggin (*NOG*) was observed within dormant breast cancer cells *in vivo*. Noggin is a secreted inhibitor of BMPs that prevents BMP binding to its receptors [284], suggesting an inhibited BMP network in dormant cells. Additionally, inhibitor of differentiation 1 (*ID1*) is a downstream transcriptional target of

BMP signalling, which functions as a regulator of cell differentiation. *ID1* over expression has been shown to induce cancer cell growth [285]. A significant downregulation of *ID1 in vivo* was observed within the dormant population, which may contribute further to the dormant phenotype. Finally, the most significantly upregulated gene in the dormant cell population *in vivo*, BGN, is of particular note. In addition to being a potent inhibitor of TGF- $\beta$  receptor activity, and synthesised downstream of an active TGF- $\beta$  pathway, exogenous biglycan has been shown to inhibit pancreatic and bladder cancer cell proliferation and induce breast cancer normalisation [286, 287]. Therefore, biglycan may be a significant contributor to the induction and maintenance of dormancy in my model, and its involvement with the TGF- $\beta$  pathway should be further investigated.

I have also shown evidence of an enrichment Wnt signalling components in the dormant population *in vivo*, but my data is conflicting and unclear. Firstly, I observed an upregulation of *FZD7* and *FZD2*, and, as receptors for Wnt ligands, may increase the number of Wnt binding sites [288]. Paradoxically, *GSK3B*, encoding the  $\beta$ -catenin destruction complex component, GSK3 $\beta$ , was significantly upregulated in my dormant cells *in vivo*, whilst the Wnt-activated transcription factor gene, *LEF1*, was also upregulated. My data does not indicate Wnt ligand availability in my model. In addition, *LEF1* upregulation at the gene expression level may not be indicative of its activity, given that as a transcription factor it must be bound to other proteins to function, such as TAZ [289]. Perhaps more indicative of Wnt pathway activity was the upregulation of two Wnt target genes, axis inhibition protein 2 (*AXIN2*) and naked cuticle homolog 1 (*NKD1*). Both genes are expressed to limit the intensity of Wnt-mediated signalling, as both *AXIN2* and *NKD1* function as Wnt antagonists [290, 291]. It is also important to consider that dormancy may require the co-operation of a number of pathways. For example, active HIPPO signalling is a known Wnt signalling inhibitor [292]. In addition, c-Myc and cyclin D require YAP-TEAD (HIPPO) and  $\beta$ -catenin-TCF/LEF (Wnt) co-operation [292]. It is not possible to reach any conclusions with regards to Wnt involvement in dormancy, however its enrichment and close co-operation with other enriched pathways in my model warrants further investigation.

Finally, I identified a number of potentially implicated transcription factors in dormancy: c-Myc, HIF-1A and p53, critical regulators of the Warburg effect. In the dormant cell population *in vivo*, a subtle downregulation of *ALDOA*, *GAPDH*, *PGK1*, *PGAM1*, *ENO1* and *LDHA* was observed, genes which encode critical enzymes within the glycolysis pathway [213]. It is currently unexplored as to whether an inhibition of aerobic glycolysis could be a driver of cancer dormancy. The glycolytic pathway, in addition to producing ATP, provides intermediates for critical cellular components such as amino acids and nucleotides [213]. Interestingly, a recent study has shown that an inhibition of aerobic glycolysis results in a reduction in YAP/YAZ transcriptional activity [293]. As YAP/TAZ are central to the crosstalk between HIPPO, TGF- $\beta$  and Wnt pathways, the role of aerobic glycolysis in the induction of dormancy warrants further investigation.

### 3.8.5 Limitations

A caveat and limitation to using a vital dye as a proliferation marker is that this approach does not necessarily identify cells which may have proliferated following extravasation, lost the dye, and which potentially became dormant, if any. However, no models currently exist in which identifying such cells is possible, given the current lack of consensus in dormancy markers. Furthermore, with respect to bioinformatic analysis and validation, a starting input of RNA for cDNA synthesis was a further limitation. Subsequent study may allow for the generation of a higher number of dormant cells; however, this study was restricted by low dormant cell yield per mouse brain. In addition, with as few as 14 dormant cells per mouse brain, any contamination from host cells is significant, therefore one must be particularly robust in data analysis quality control.

### 3.8.6 Summary

In summary, the results in this chapter suggest that dormancy, in disseminated breast adenocarcinoma cells to the brain, may be characterised by a downregulation of aerobic glycolysis at the gene expression level. Transcriptomic analysis revealed a downregulation of *ALDOA*, *GAPDH*, *PGK1*, *PGAM1*, *ENO1* and *LDHA*, genes which encode

critical and rate-limiting enzymes within the glycolytic pathway. Further investigation is therefore warranted in order to understand whether inhibited aerobic glycolysis is a driver of dormancy, rather than as an effect of the phenomenon. Further investigation is also required to understand further the role of the small leucine-rich proteoglycan, biglycan. As the most significantly upregulated gene in dormant cells, biglycan is synthesised downstream of an active TGF- $\beta$  signalling pathway, acting as a negative regulator of TGF- $\beta$  signalling. Currently, no such data exists regarding the potential influences on proteoglycan expression and metabolism. In addition, a comparison of dormant and proliferating cancer cells by GO and KEGG analysis identified an enrichment in TGF- $\beta$  signalling, as well as HIPPO and Wnt signalling pathways. These pathways converge on YAP, a transcriptional regulator which holds significant influence over cancer cell proliferation and growth.

## Chapter 4

### Characterisation of Biglycan as a Potential Mediator and/or Marker of Cellular Dormancy

#### 4.1 Introduction

Differential gene expression analysis of dormant and proliferating breast cancer cell populations within my brain tumour xenograft model identified a significant upregulation of the proteoglycan, biglycan. In proliferating cells, an almost absent read count within each sample was observed, yet biglycan read counts in dormant cells were orders of magnitude higher. By statistical significance, biglycan was the most differentially expressed gene in dormant cells. The addition of recombinant biglycan to bladder cancer cells *in vitro* has been associated with a reduction in cellular proliferation, whilst its knock-down *in vivo* has been associated with increased proliferation [294]. Biglycan overexpression in pancreatic cancer is associated with G1 cell cycle arrest [286] and ECM-bound biglycan was shown to induce breast cancer cell normalisation [287]. Biglycan is therefore of interest in the context of dormancy, and its potential role as a mediator and/or marker of breast cancer cell dormancy in the brain is to be explored within this chapter. It stands to reason that biglycan overexpression may be a potential driver of a reversible growth-arrest phenotype *in vivo*. Given the challenges in identifying dormant cancer cells in tissues, in particular the selection of suitable markers, biglycan itself may be a novel marker of dormancy. Furthermore, if biglycan overexpression is implicated in the induction and maintenance of dormancy, this may implicate other, targetable pathways, whose activation or inhibition may contribute to reversible growth arrest. This is to be explored within this chapter.

### 4.1.1 Hypothesis

In this chapter, there are two hypotheses which are to be explored. The first is that overexpression of biglycan in breast cancer cells in the brain induces and maintains a state of dormancy. The second hypothesis is that biglycan expression is a marker of breast cancer cell dormancy in the brain

#### 4.1.1.1 Aims and Objectives

To determine whether biglycan overexpression is a driver and/or marker of breast cancer dormancy in the brain, biglycan expression must be validated at the protein level in dormant cancer cells. Following this, the potential growth-inhibitory effects of biglycan on breast cancer cells *in vitro* are to be explored. This will complement an analysis of other human breast cancer models to determine whether breast cancer cells metastasised to the brain are potentially amenable to biglycan-induced growth arrest. To address these points, the following aims were devised:

**1) To identify biglycan-expressing breast cancer cells in mouse brain tumour xenografts**

To achieve this, sections of mouse brain tumour xenografts, from the same model employed in Chapter 3, will be scanned for biglycan-expressing cancer cells. If cells are identified, a secondary objective will be to document their localisation within the brain.

**2) To develop an *in vitro* model of biglycan over-expression**

This will be achieved by determining the mRNA sequence of biglycan in dormant breast cancer cells, via analysis of mRNAseq reads. This data will be used to synthesise a biglycan gene to be placed into a suitable vector for over-expression of the protein *in vitro*.

**3) To investigate the effects of recombinant human biglycan on breast cancer cells in culture**

This will be achieved by titrating commercially available human recombinant biglycan against breast cancer cells *in vitro*. In addition, a secondary objective will be to determine whether soluble biglycan, or biglycan-coated culture plastic, has the greatest growth-inhibitory effect, if any.

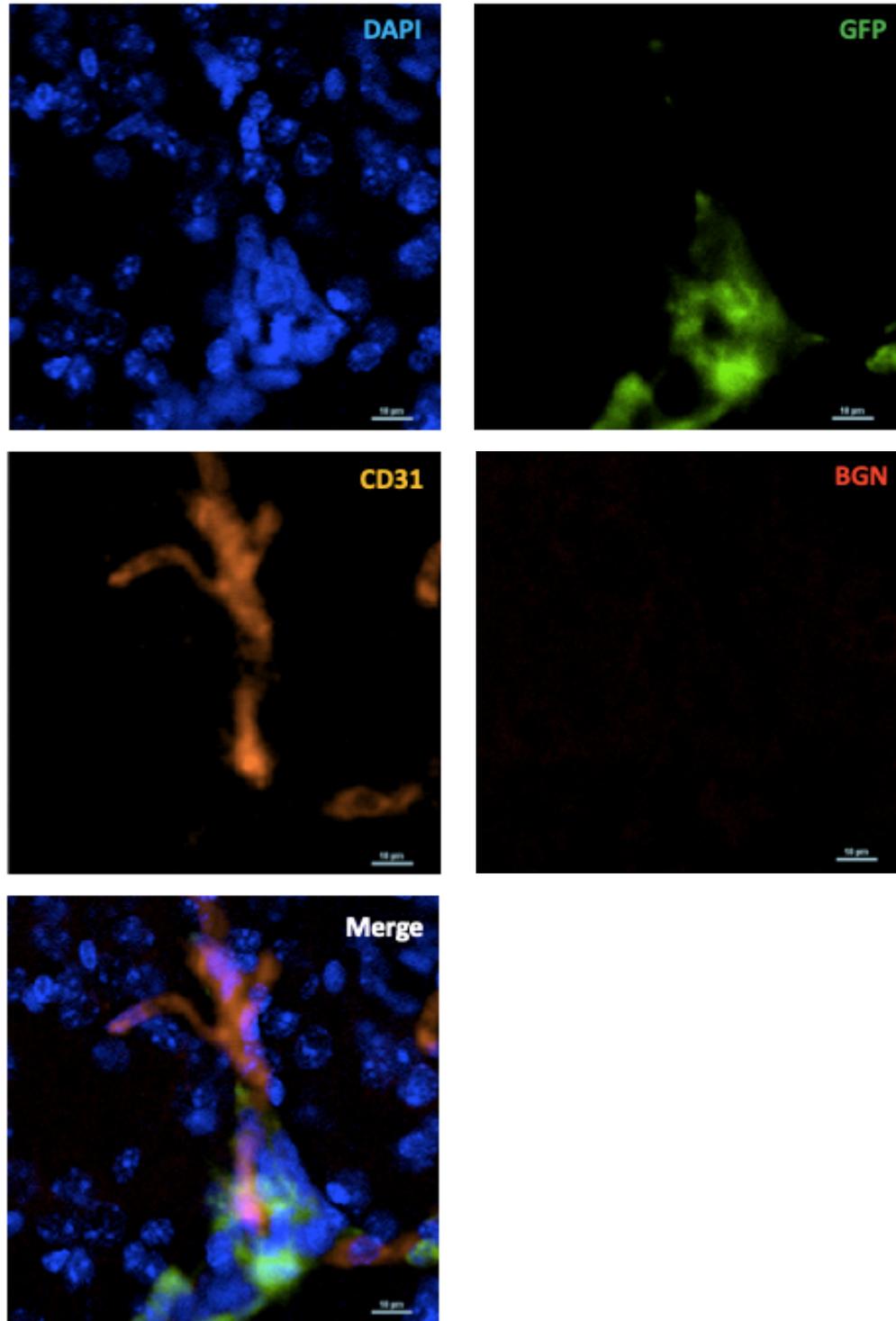
**4) To compare biglycan expression between the current xenograft model and in human patient samples**

To achieve this, biglycan protein expression will be determined in normal mouse brains, compared to patient-derived brain tumour xenografts. In addition, biglycan gene expression will be analysed in publicly available transcriptomic data in primary and metastatic breast cancers. This will provide some evidence to inform whether or not metastatic breast cancer cells are amenable to biglycan-induced growth inhibition.

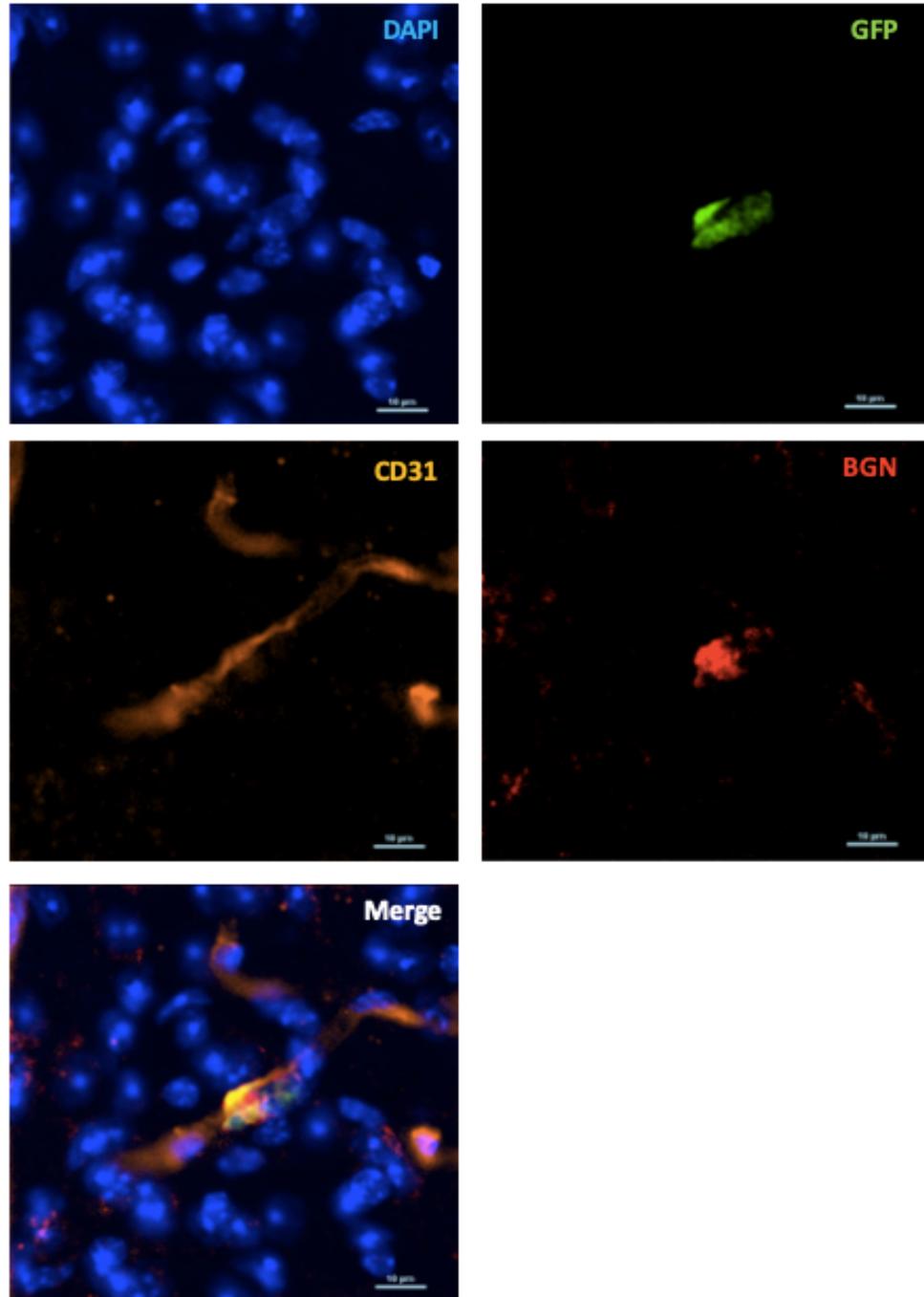
## 4.2 Determining the Localisation of Dormant MDA-MB-231 Cells in Mouse Brain Tumour Xenografts

Recent studies have suggested that the perivascular niche may regulate breast cancer dormancy and metastatic outgrowth [133, 136]. Therefore, the localisation of dormant cells within my model may uncover the microenvironment which is permissible for the induction of a dormant state. Gene expression analysis showed that biglycan was significantly upregulated in dormant cells. However, read counts in proliferating cells were very low and it was therefore reasoned that biglycan would only be detectable in dormant cells within the brain by immunofluorescence imaging. Consequently, sections of brain tumour xenografts in the carotid artery injection model were scanned for biglycan-expressing breast cancer cells.

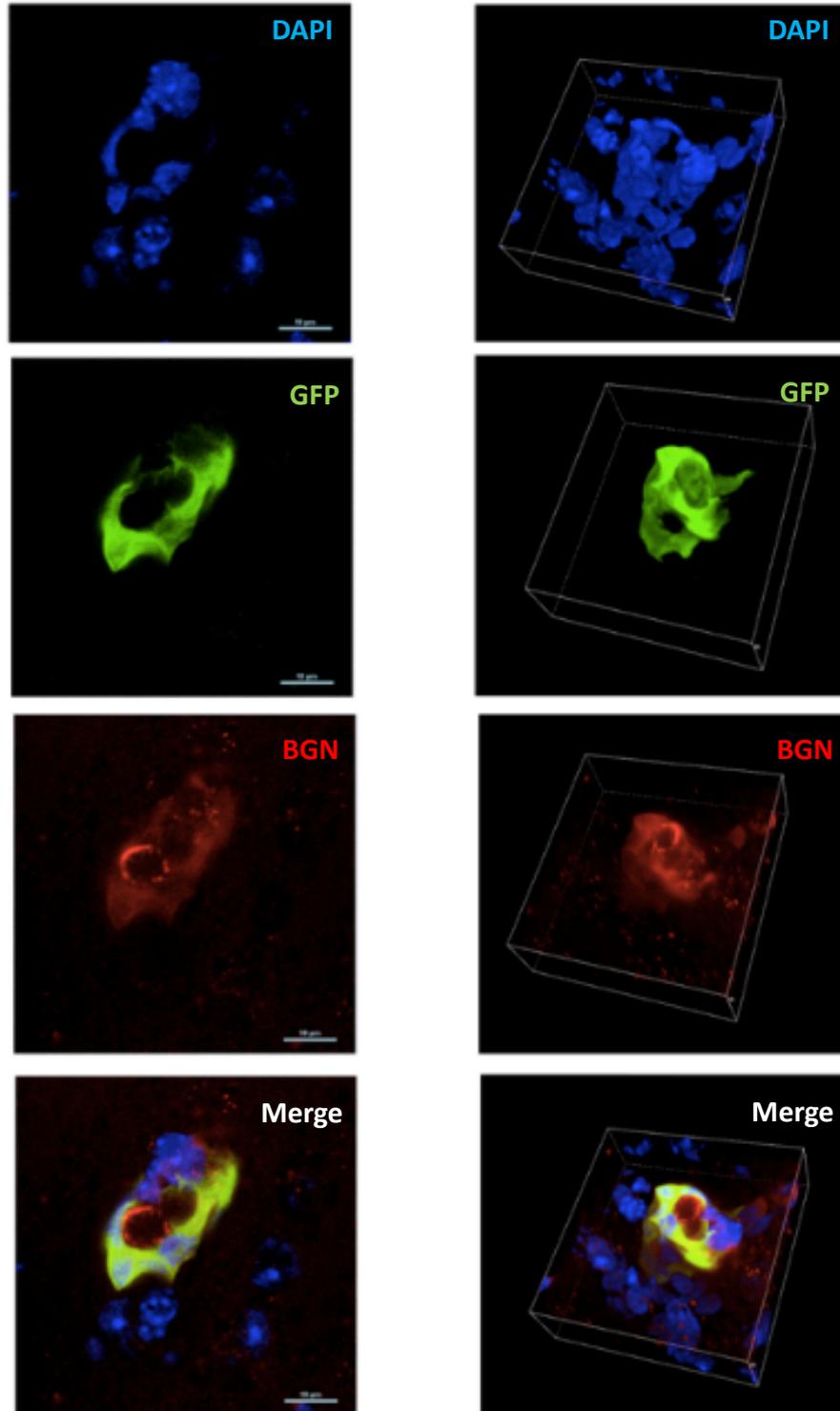
Initial staining of naïve mouse brains in my study, at the described magnification, showed that biglycan protein could not be detected (data not shown). As such, any biglycan observed is likely to have originated from cancer cells. The majority of breast cancer cells, 28 days post-intra-carotid injection, were negative for biglycan, throughout all regions of the brain (**Figure 4.1**). As reported in the previous chapter, as few as 14 dormant cells were isolated from each mouse brain. This is reflected in the number of positive identifications of BGN+ cancer cells within the brain, of which 2 events were observed. Of the first event, a solitary BGN+ cancer cell was observed within the brain. Interestingly, this cell was wrapped around brain microvascular endothelium, as identified by CD31 expression (**Figure 4.2**). Further investigation of a secondary event was conducted using 3D rendering, in the absence of CD31 staining. The BGN+ cancer cell, morphologically, appeared to be wrapped around another structure within the brain. In addition, biglycan appeared to be distributed throughout the local microenvironment. Of note, the architecture of biglycan staining displayed a dense ring of deposition, lining the structure around which the cancer cell was wrapped (**Figure 4.3**). Biglycan may therefore contribute to the modification of the local microenvironment in the induction of dormancy.



**Figure 4.1.** The majority of MDA-MB-231 cells which colonise mouse brains following intra-carotid injection are negative for biglycan. MDA-MB-231/GFP/CV cells were injected into the internal carotid artery of C.B-17 SCID mice and analysis was performed 28 days later by immunofluorescence. Brain sections were probed for green fluorescent protein (GFP), biglycan (BGN) and for the endothelial cell marker, cluster of differentiation 31 (CD31). Nuclei were counterstained with DAPI. Image representative of biglycan negative cancer cells displaying metastatic outgrowth in the brain. Scale bar = 10µm.



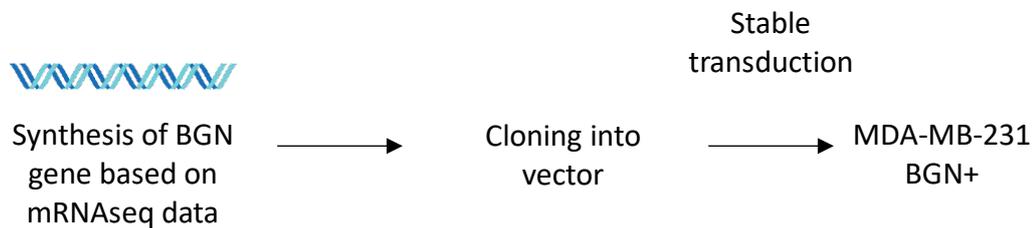
**Figure 4.2. Biglycan-expressing MDA-MB-231 cells in brain tumour xenografts localise around cerebral vasculature.** MDA-MB-231/GFP/CV cells were injected into the internal carotid artery of C.B-17 SCID mice and analysis was performed 28 days later by immunofluorescence. Brain sections were probed for green fluorescent protein (GFP), biglycan (BGN) and the endothelial cell marker, cluster of differentiation 31 (CD31). Nuclei were counterstained with DAPI. Image shows a breast cancer cell within the brain which display positive staining for biglycan. Scale bar = 10µm.



**Figure 4.3. Three-dimensional rendering of biglycan-secreting cancer cells *in vivo*.** MDA-MB-231/GFP/CV cells were injected into the internal carotid artery of C.B-17 SCID mice and analysed 28 days later by immunofluorescence. Brain sections were probed for green fluorescent protein (GFP) and biglycan (BGN). Nuclei were stained with DAPI. Scale bar = 10µm.

### 4.3 Analysis of Biglycan Overexpression in MDA-MB-231 Cell Line

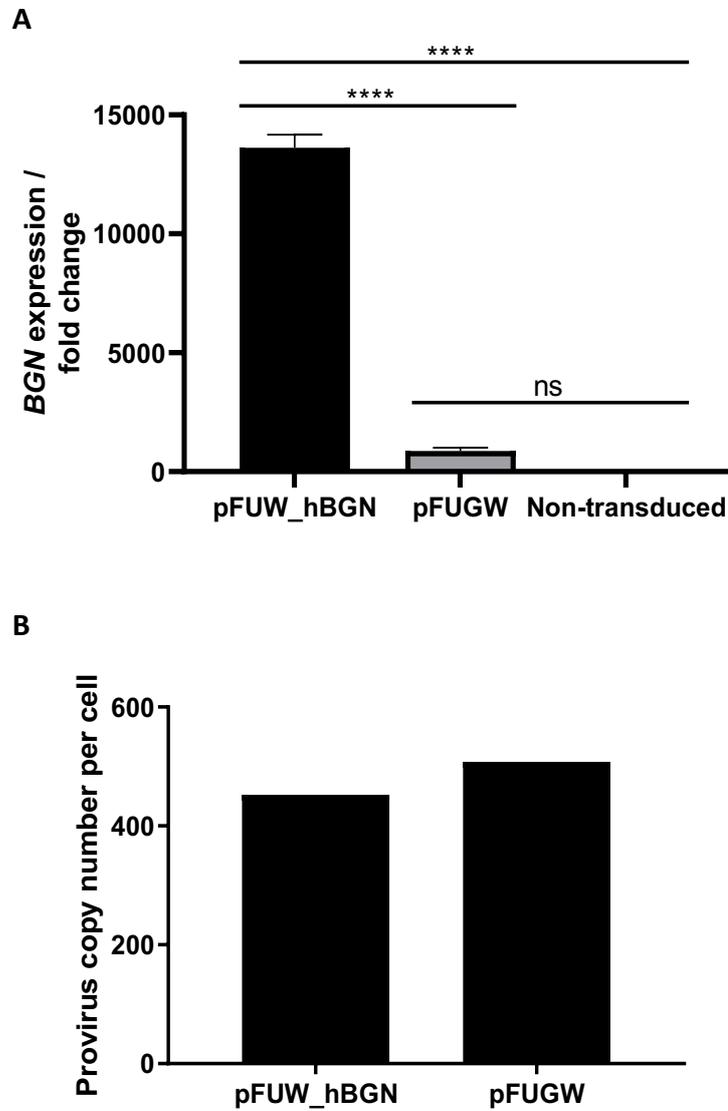
To establish the potential effect of biglycan on breast cancer proliferation, an *in vitro* biglycan over-expression model in MDA-MB-231 cells was developed. The consensus sequence of the biglycan mRNA transcript within dormant MDA-MB-231 cells *in vivo* was first determined. The representative sequence was obtained using Integrative Genomics Viewer software [235], with which BAM files containing human-aligned reads from dormant cells were inputted. The consensus human biglycan sequence differed from the published mRNA sequence [295] by one silent point mutation (**Supplementary Figure S5**). The human biglycan gene was artificially synthesised based on the sequence from dormant cells, and subsequently cloned into the pFUW vector, under the control of a constitutively active ubiquitin C promotor (pFUW\_hBGN). This vector was used to stably transduce MDA-MB-231 cells by lentiviral transduction (231 pFUW\_hBGN) (**Figure 4.4**).



**Figure 4.4. Generation of biglycan-expressing MDA-MB-231 cells.** Sequencing data for human biglycan within the dormant 231 cancer cell population was analysed by IGV software. The biglycan gene containing two silent point mutations was artificially synthesised and cloned into pFUW, prior to stable transduction of MDA-MB-231 cells.

pFUW\_hBGN and GFP-expressing vector pFUGW [257], respectively, were used to stably transduce MDA-MB-231 cells. As shown in **Figure 4.5A**, pFUW\_hBGN-transduced cell line displayed high levels of biglycan expression, while no biglycan expression was detected in pFUGW-transduced MDA-MB-231 cells. The transduction efficiency for the

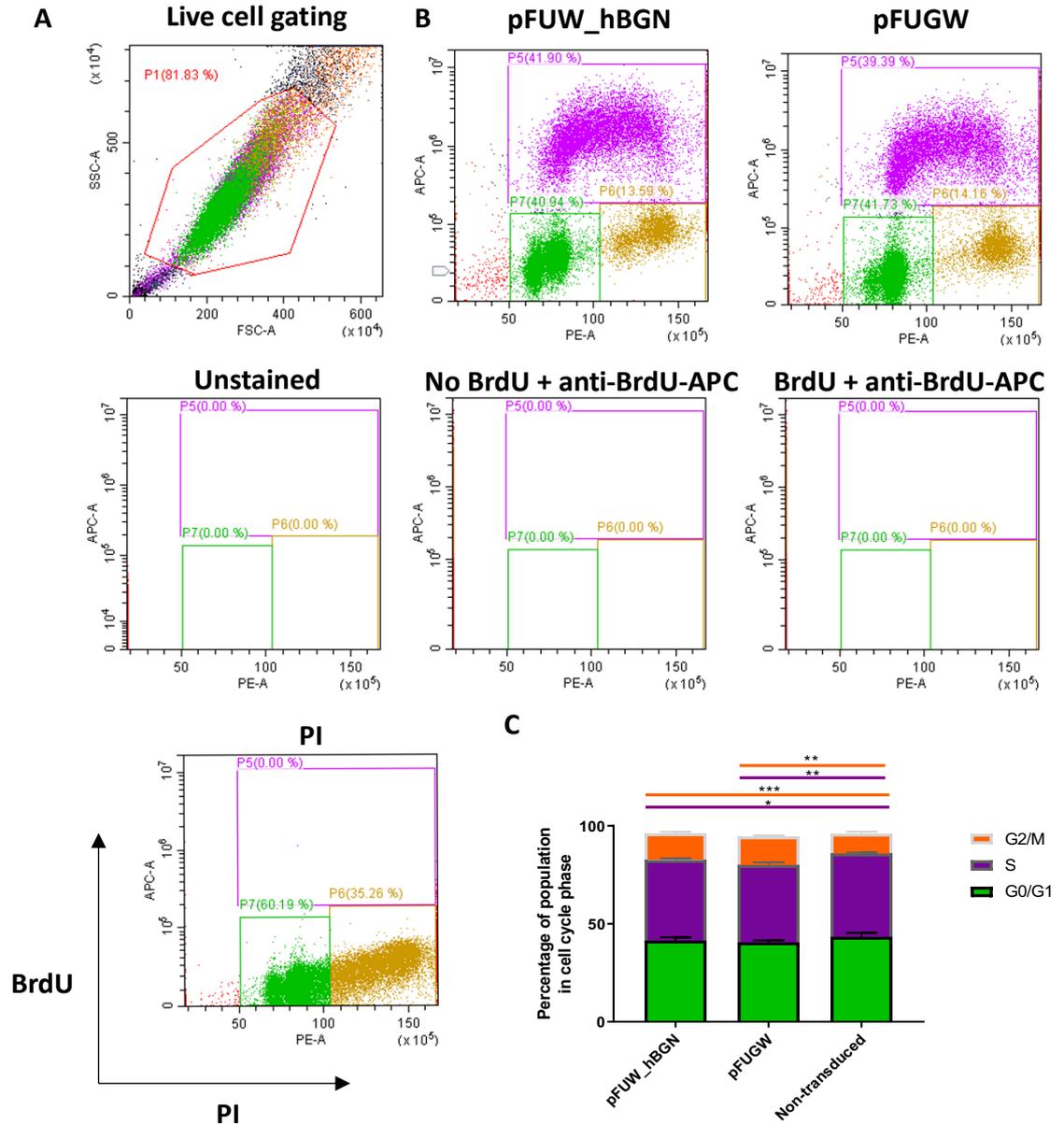
two transduced cell lines was comparable, as demonstrated by similar values of vector copy numbers (Figure 4.5B).



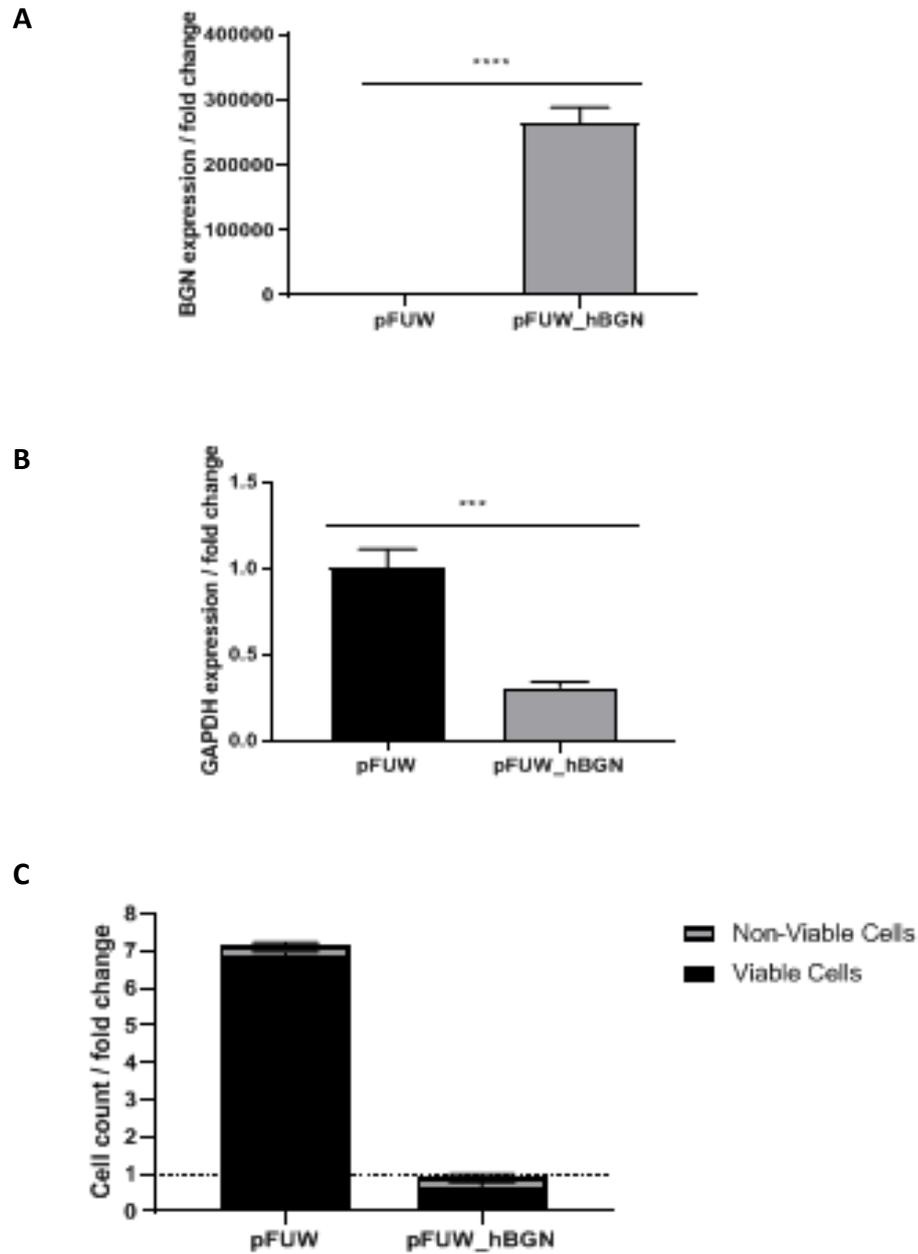
**Figure 4.5. Overexpression of human biglycan in MDA-MB-231 cells.** Biglycan and GFP-expressing constructs (pFUW\_hBGN and pFUGW, respectively) were used to stably transduce MDA-MB-231 cells. Biglycan and GFP were inserted into pFUW vehicle at the same site, downstream of a constitutively active ubiquitin C promoter. A) Quantitative PCR validation confirms biglycan expression in MDA-MB-231 cells. N=3. Statistical analysis performed was an ordinary one-way ANOVA with multiple comparisons test. B) Quantification of provirus copy number per cell of integrated pFUW\_hBGN and pFUGW lentiviral vectors.

There was no observable effect on MDA-MB-231 cell proliferation when transduced with pFUW\_hBGN or pFUGW at 1X virus concentration (data not shown). To confirm these findings, the potential growth-inhibitory effects of biglycan over-expression was determined by cell cycle FACS. MDA-MB-231 cells transduced with pFUW\_hBGN and pFUGW vectors were compared to non-transduced cells. Sub-confluent cells were labelled with BrdU and stained by anti-BrdU-APC antibody prior to analysis, as a marker of actively cycling cells. As shown in **Figure 4.6A**, live cells were gated by FSC and SSC. Brdu-APC-labelled cells were counter-stained with propidium iodide (PI) to enable gating around cell populations within G2/M, S and G0 phase. A small but significant reduction in cells in S phase, and a small but significant increase in cells in G2/M phase was observed when transduced with pFUW\_hBGN, compared to non-transduced cells (**Figure 4.6B**). However, the same observation was made in cells transduced with pFUGW (**Figure 4.6B**), compared to non-transduced cells. There was no significant difference of cells in any phase between cells transduced with pFUGW and cells transduced with pFUW\_hBGN. Therefore, at current levels of expression, biglycan was insufficient to induce changes to proliferation rate or induce cell cycle arrest.

It was reasoned that a higher level of biglycan expression may be required in order to induce cell cycle arrest in MDA-MB-231 cells. As a follow-up, pFUW\_hBGN and the control pFUW lentivirus were concentrated 40X prior to lentiviral transduction. Transduction with concentrated pFUW\_hBGN vector induced a significantly higher level of biglycan expression than the vehicle control (**Figure 4.7A**). Interestingly, this was associated with a significant downregulation of the glycolysis-associated enzyme, GAPDH (**Figure 4.7B**). GAPDH was observed to be downregulated within dormant MDA-MB-231 cells *in vivo*. High biglycan expression and a reduction in GAPDH expression were observed to be concurrent with complete cell cycle arrest. Viable, arrested cells transduced with pFUW\_hBGN were observed at 72 hours post-transduction (**Figure 4.7C**). However, repeat validation experiments could not be performed as the cell line was permanently arrested, and therefore could not be propagated for downstream study.



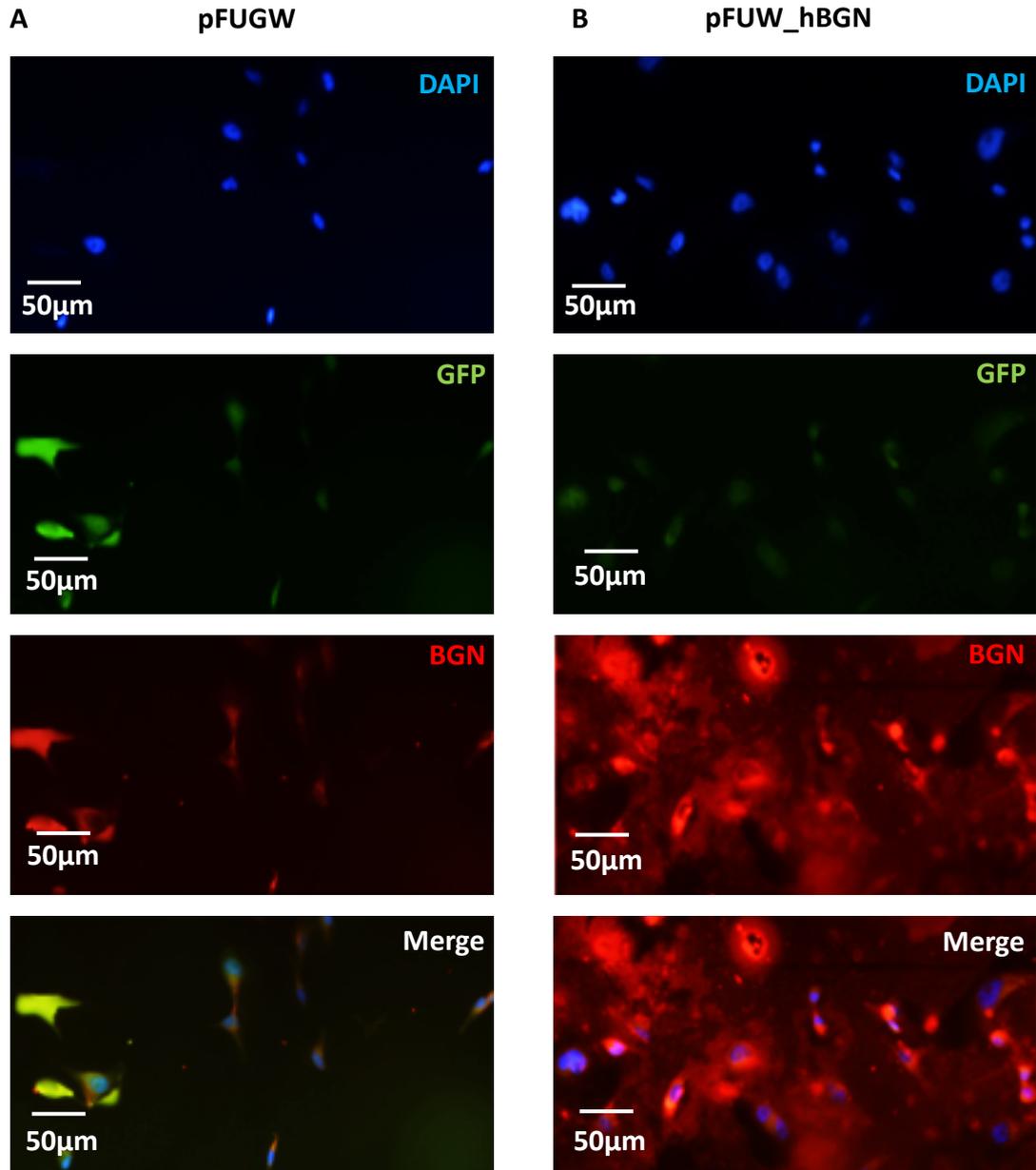
**Figure 4.6. Overexpression of biglycan in MDA-MB-231 cells at low level does not induce cell cycle arrest under high serum conditions *in vitro*.** MDA-MB-231 cells were stably transduced with constitutively active BGN- and GFP-expressing vectors, pFUW\_hBGN and pFUGW, respectively, at 1X MOI. Cell cycle phase was analysed by FACS. A) Gating strategy for live cells (P1). B) Cell cycle phase analysis by PI and BrdU fluorescence for transduced cells and FACS controls, as shown. Gating for cell cycle phase was as follows: G0/G1 phase (green – P7), S phase (purple – P5), G2/M phase (orange – P6). C) Bar chart of representative cell cycle phase. N=3. Statistical analysis performed was an ordinary one-way ANOVA with multiple comparisons test.



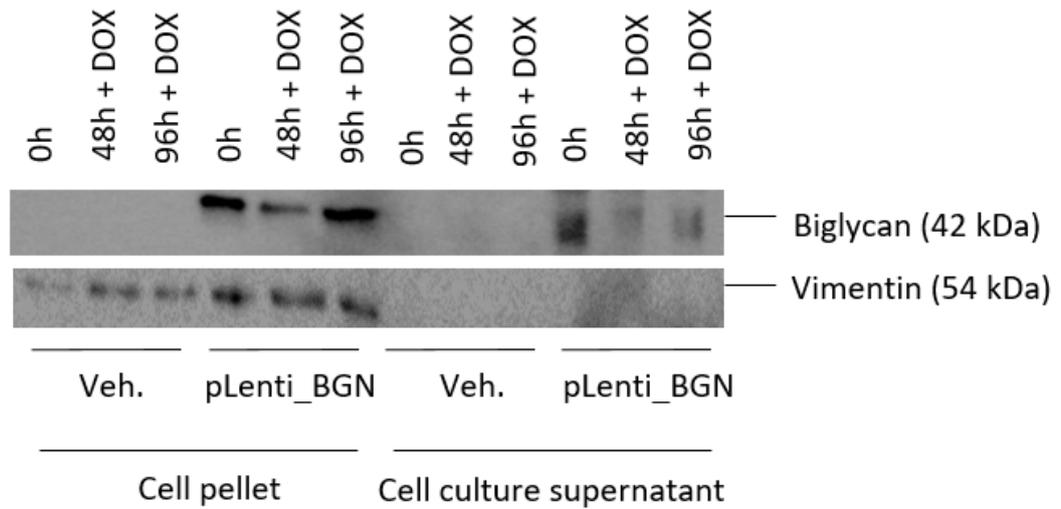
**Figure 4.7. Stronger biglycan overexpression induces downregulation of *GAPDH* and complete growth inhibition of MDA-MB-231 cells *in vitro*.** MDA-MB-231 cells were stably transduced with 40-fold concentrated pFUW and pFUW\_hBGN. RT-PCR analysis of A) *BGN* and, B) *GAPDH*, relative to *POLR2A*, at 72 hours post-transduction. C) Transduced cells were stained with trypan blue at harvest and assessed for viability. Given is the cell count as a proportion of cells compared to those seeded prior to transduction. N=3. Statistical analyses performed was a two-tailed student's *t* test.

To determine the cellular localization of biglycan following overexpression *in vitro*, MDA-MB-231 cells transduced with pFUW\_hBGN and pFUGW were analysed by immunofluorescence. When stained for biglycan, cells transduced with pFUGW demonstrated weak positivity in the channel used for the detection of BGN. (**Figure 4.8A**). However, this is likely due to fluorescence bleed-through from GFP. When cells transduced with pFUW\_hBGN were analysed, all cells were intensely positive for biglycan (**Figure 4.8B**). The extracellular space surrounding biglycan-positive cells also demonstrated a high level of fluorescence. This indicated that biglycan is secreted from MDA-MB-231 cells.

To further demonstrate that biglycan is secreted when overexpressed, MDA-MB-231 cells were stably transduced with Tet-OFF inducible pLenti Puro vehicle and pLenti\_BGN. Using pLenti\_BGN, biglycan was expressed in the absence of doxycycline, and detected in the total protein fraction and in the cell culture supernatant (**Figure 4.9**). Upon administration of doxycycline, the level of biglycan within the total protein fraction persisted after 96h. However, there was a reduction in the level biglycan detected in the cell culture supernatant, which was observed at 48h post-addition of doxycycline (**Figure 4.9**). These results confirm that upon its expression, biglycan is secreted from MDA-MB-231 cells. Upon inhibition of biglycan expression, this data also shows that MDA-MB-231 cells rapidly cease biglycan secretion, whilst intracellular biglycan protein persists.



**Figure 4.8. Biglycan is secreted when over-expressed in MDA-MB-231 cells *in vitro*.** MDA-MB-231 cells were stably transduced with the A) Green fluorescent protein (GFP)-expressing pFUGW, or, B) Biglycan (BGN)-expressing pFUGW\_hBGN lentiviral vectors. Cells were cultured through two passages and for at least 96 hours post-transduction prior to staining. Cells were fixed in PFA and cells were stained for human BGN and GFP.



**Figure 4.9. Biglycan secretion is inhibited using a Tet-OFF biglycan expression system in MDA-MB-231 cells *in vitro*.** MDA-MB-231 cells were stably transduced with Tet-OFF pLenti Puro vehicle (veh.) and pLenti\_hBGN. Shown is a Western blot analysis of biglycan core protein in the total protein fraction (cell pellet) and the cell culture supernatant. Doxycycline (DOX) was administered for 48 and 96h. Image representative of two independent experiments.

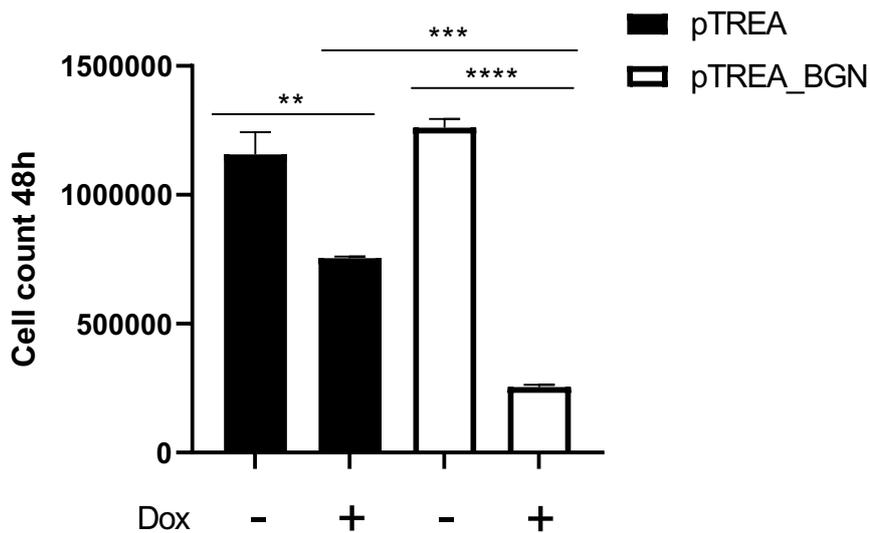
#### 4.4 Doxycycline-Inducible Model of Biglycan Overexpression

A novel model was required to determine the effects of biglycan-overexpression on MDA-MB-231 cell proliferation *in vitro*, that would allow for cell line propagation and controlled induction of biglycan expression.

The human biglycan sequence found in dormant cancer cells was cloned into the doxycycline-inducible pTREAuro3 lentiviral vector (pTREA\_BGN) [296]. The BGN sequence was inserted downstream of multiple Tet-ON regions and a minimal CMV promoter. A neomycin resistance gene, downstream of a constitutively active phosphoglycerate kinase (PGK) promoter, was designed, synthesised and also inserted into the vector to allow for selection. MDA-MB-231 cells were transduced with pTREA\_BGN and the GFP-expressing vehicle, pTREAuro3 (pTREA). A 40-fold concentration of viral vectors was employed. Preliminary experiments validating this

model did not show any growth-inhibitory effects at 1-fold concentration of viral vectors (data not shown).

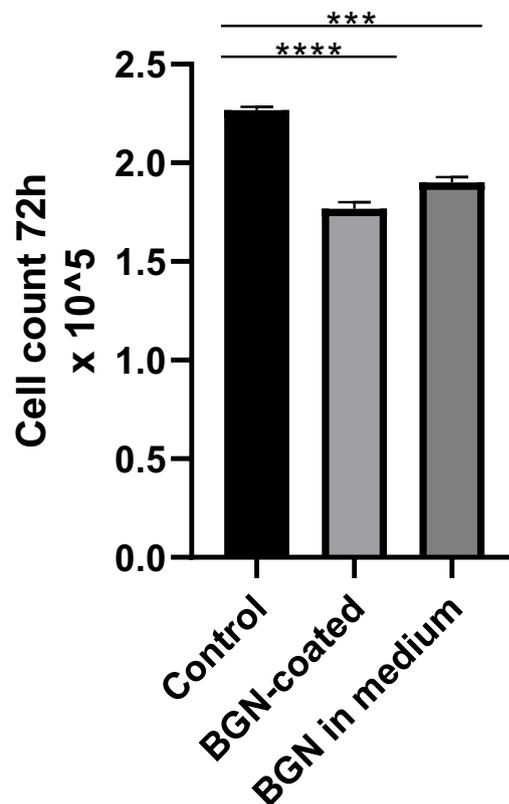
MDA-MB-231 cells, transduced with pTREA<sub>AutoR3</sub> and pTREA\_BGN, were cultured *in vitro* in the presence or absence of doxycycline, for 48h (**Figure 4.10**). In cells transduced with pTREA<sub>AutoR3</sub>, a small but significant reduction in cell count was observed following the addition of doxycycline. However, in cells transduced with pTREA\_BGN, addition of doxycycline was associated with a very large reduction in cell count over the same time period. At very high biglycan expression, biglycan therefore appears to inhibit cellular proliferation.



**Figure 4.10. Doxycycline-induced biglycan over-expression induces growth inhibition of MDA-MB-231 cells *in vitro*.** MDA-MB-231 cells were stably transduced with doxycycline-inducible green fluorescent protein (GFP)- and biglycan (BGN)-expressing vectors, pTREA and pTREA\_BGN, respectively. Shown is cell count after 48h of doxycycline exposure. Cells were transduced with lentiviral vectors at 40X concentration. N=3. Data is representative of 3 independent experiments.

#### 4.5 Determining the Anti-Proliferative Properties of Soluble Biglycan in MDA-MB-231 Cells

Analysis of biglycan over-expressing 231 cells *in vitro* determined that biglycan is being secreted into the media. Soluble biglycan was therefore explored further *in vitro*. MDA-MB-231 cells were cultured on plastic substrate in the absence/presence of recombinant human biglycan (rhBGN), as well as on rhBGN-coated plastic (**Figure 4.11**). Coated hrBGN and soluble rhBGN both induced a significant decrease in proliferation of MDA-MB-231 cells over 72h. Thus, biglycan secreted from dormant MDA-MB-231 cells may induce a dormant state in an autocrine fashion.



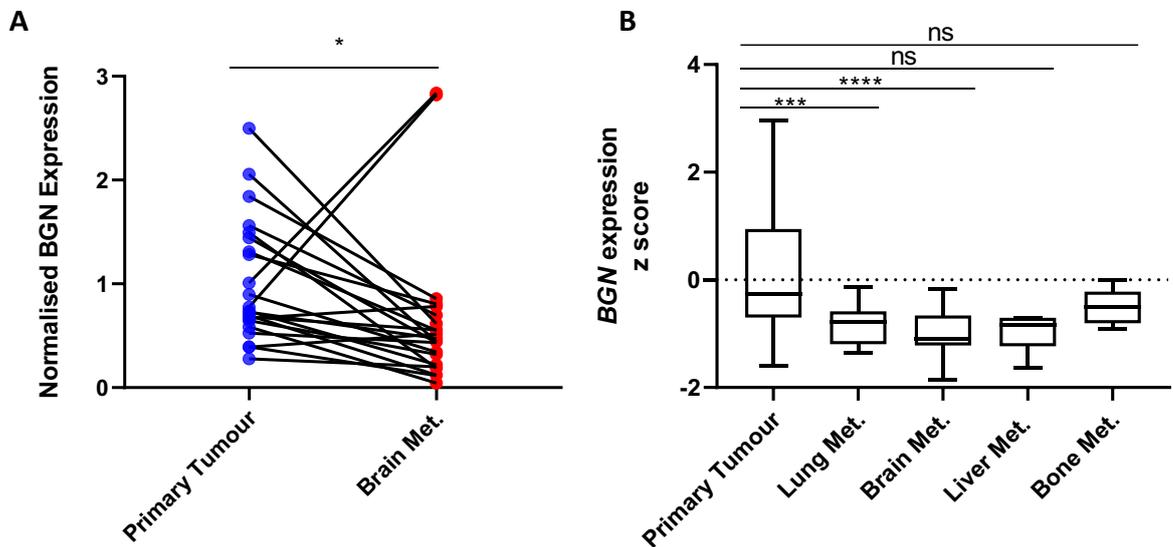
**Figure 4.11. Recombinant human biglycan induces growth arrest of MDA-MB-231 cells *in vitro*.** MDA-MB-231 cells were cultured on plastic substrate, rhBGN-coated plastic, and plastic substrate with 5 $\mu$ g per ml soluble rhBGN. Cells were cultured for 72h prior to counting. Culture medium constituted of high glucose DMEM/F12 with 10% FBS. N=3. Data representative of two independent experiments. Statistical analysis performed was an ordinary one-way ANOVA with multiple comparisons test.

#### 4.6 Analysis of Biglycan Expression in Human Primary Breast Cancer and Brain Metastases

I have shown in Chapter 3 that biglycan expression is minimal in proliferating MDA-MB-231 cells *in vivo*, with a very low biglycan read count detected. The expression status of biglycan in human breast cancer metastases was therefore investigated, to provide indications whether they would potentially be amenable to biglycan-induced growth inhibition. Initially, biglycan expression within primary tumours was compared to that of patient-matched brain metastases, using publicly available mRNAseq data (**Figure 4.12A**). In all but 2 out of 22 patients analysed, biglycan expression was found to be significantly reduced in brain metastases. In confirmation of these findings, a meta-analysis was performed on publicly available microarray datasets, comparing biglycan expression in primary breast tumours compared to lung, brain, liver and bone metastases (**Figure 4.12B**). Biglycan expression was significantly higher in primary tumours compared to lung and brain metastases. There was not a statistically significant reduction in biglycan expression in liver and bone metastases, however this may be attributed to small sample size and there was a tendency towards reduced expression. This is in itself not evidence of dormancy. Further experimental data in a controlled model would be required to compare cancer cell growth rates in primary vs metastatic breast cancer. However, this data suggests that a reduction in biglycan expression may be required in order for metastatic outgrowth to occur within the brain. As such, breast cancer cells within the brain may be amenable to biglycan-induced growth inhibition. It is also important to recognise that biglycan over-expression *in vitro* cannot be compared to this data. A crucial future study would therefore be to inject or implant biglycan-overexpressing cancer cells into mice brains and perform functional validations.

To determine the expression of BGN in brain metastases, total protein was isolated from two different HER2+ER+PR+ patient-derived brain tumour xenografts (PDXs) named breast cancer brain metastasis (BCBM) 2 and BCBM6. Analysis of BGN expression within PDXs may provide some evidence that breast cancer cells in the brain are amenable to BGN-induced growth inhibition, if expression of biglycan is either low or absent compared to biglycan-overexpressing cancer cells *in vitro*. In MDA-MB-231 cells stably

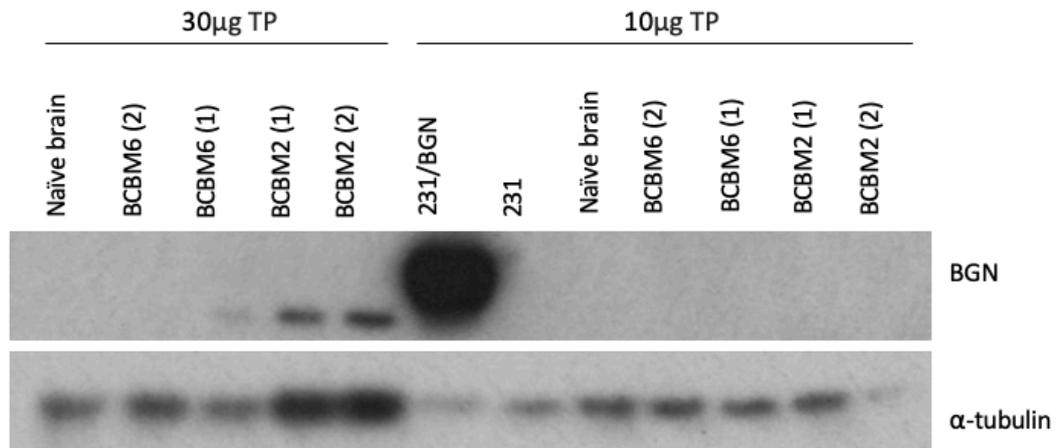
transduced with pFUW\_hBGN (see Chapter 4.3), for constitutive biglycan expression, high levels of biglycan were detected within 10 $\mu$ g of total protein lysate (**Figure 4.13**). In comparison, only BCBM2 demonstrated some biglycan expression in 30 $\mu$ g but not 10 $\mu$ g total protein (**Figure 4.13**). However, this was at much lower levels than that of the biglycan-expressing 231 cell line. BCBM6 demonstrated no biglycan expression, akin to that of the naïve mouse brain and non-transduced MDA-MB-231 cells (**Figure 4.13**).  $\alpha$ -tubulin loading controls varied between each sample, however this was more likely due to variations in tumour-to-stromal cells ratio, and between the MDA-MB-231 cell line and mouse brain, than loading errors. Human BCBrMs could therefore be amenable to growth inhibition by biglycan, given their innately low expression of the protein.



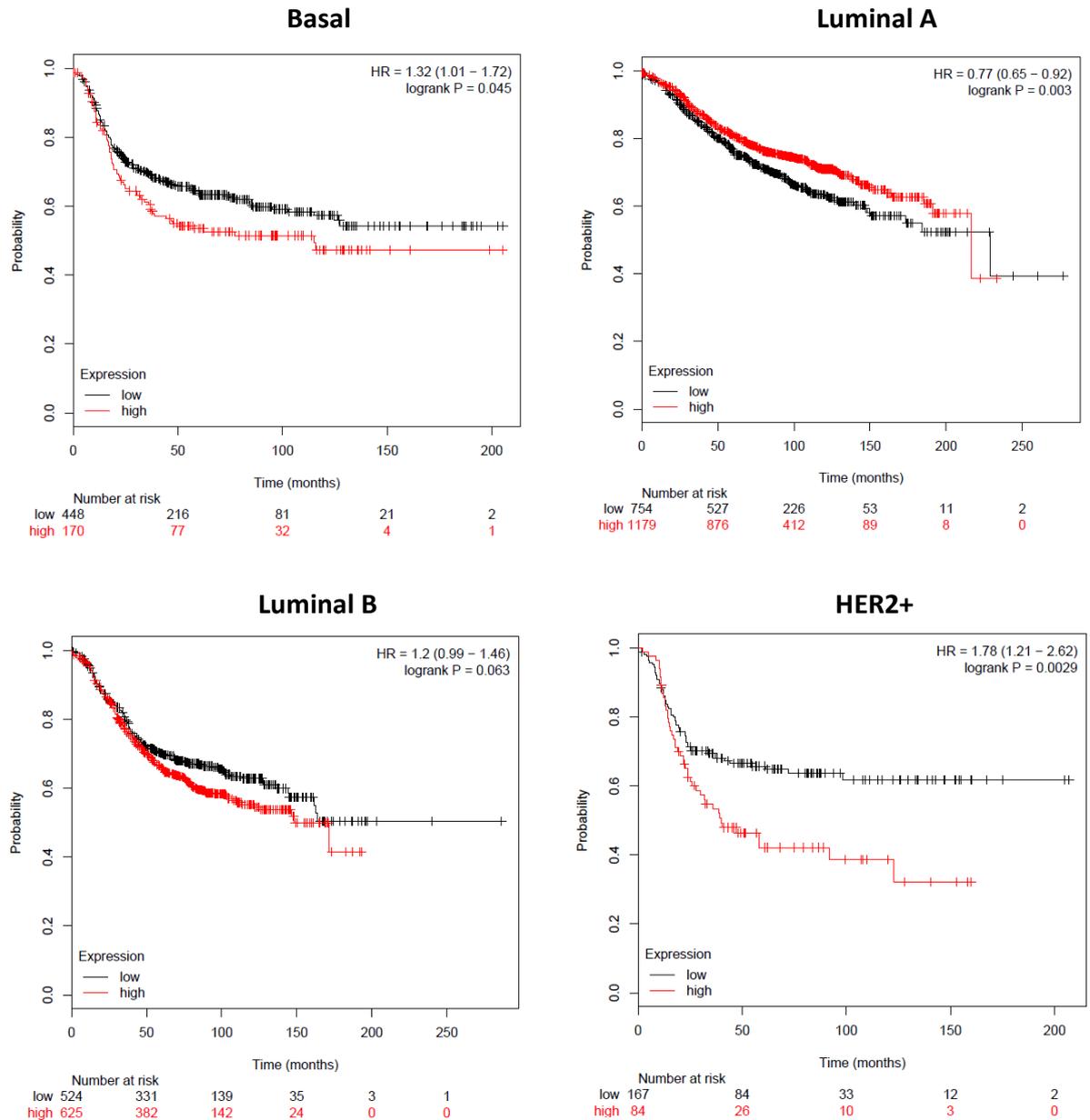
**Figure 4.12. Meta-analysis of biglycan expression in breast cancer metastases compared to primary tumours.** A) Analysis of biglycan expression in brain metastases compared to patient-matched primary breast tumours using publicly available mRNAseq data [247]. Metastasis = Met. N = 22. BGN expression was normalised to that of a suitable housekeeping gene, *POLR2A*. Performed was a paired two-tailed t test. B) Analysis of biglycan expression using publicly available microarray datasets of primary tumours compared to lung, brain, liver and bone metastases. Data accession numbers for each dataset were GSE2034, GSE5327, GSE12276, GSE14017, GSE14018, GSE43837 [249, 250, 251, 252, 253, 254]. Datasets were normalised independently and biglycan expression was normalised to a suitable housekeeping gene, *POLR2A*, prior to calculation of z scores. Performed was an ordinary one-way ANOVA with Tukey's multiple comparisons test. N = 548 (primary tumour), 19 (lung metastases), 31 (brain metastases), 5 (liver metastases), 10 (bone metastases).

Biglycan was then explored in the context of distant metastasis-free survival (DMFS) and relapse-free survival (RFS). To accomplish this, publicly available mRNA gene chip data for primary breast cancer and associated cancer survival statistics were analysed by Kaplan Meier plotter [255]. For basal and HER2+ breast cancers, high biglycan expression was correlated with a reduction in RFS (**Figure 4.14**). There was a further tendency for luminal B breast cancers to display a reduced RFS, at just outside the range of statistical significance (logrank  $p = 0.063$ , hazard ratio = 1.2). For luminal A breast cancers, however, high biglycan expression correlated with a significantly increased RFS. Interestingly, where biglycan expression was high within the primary breast tumour, DMFS was significantly reduced for luminal A, luminal B and HER2+ phenotypes (**Figure 4.15**). The effect was especially pronounced for HER2+ subtype. This trend was also observed for basal breast cancers, but did not reach statistical significance (logrank  $p = 0.054$ , hazard ratio = 1.74).

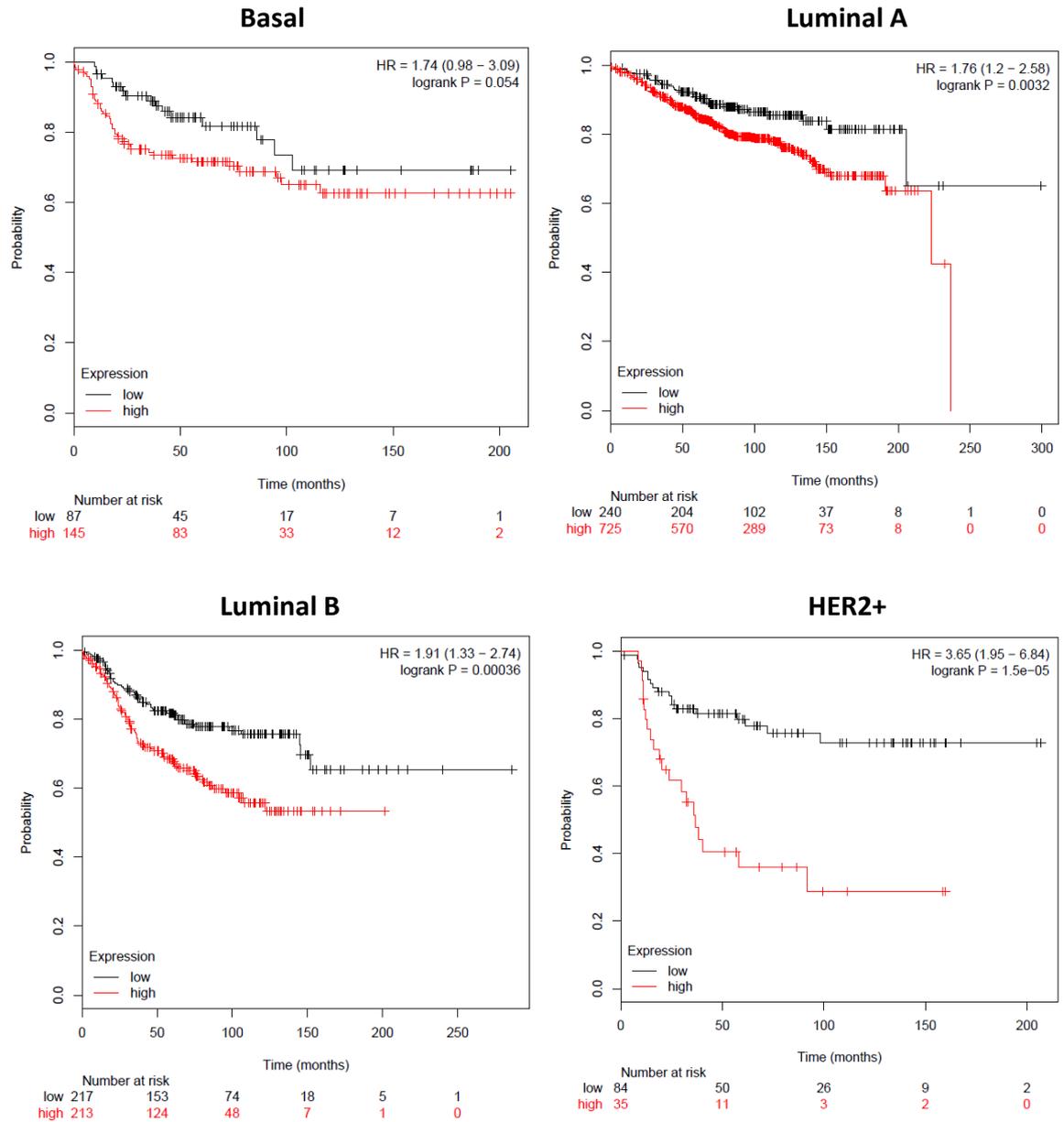
This data suggests a dual role of biglycan in cancer progression. Relatively high biglycan expression within the primary tumour may account for an increased likelihood of metastasis, whereas in the context of disseminated tissues, biglycan may possess anti-proliferative properties. It is unclear from current analyses of publicly available data whether biglycan is being synthesised by primary and metastatic breast cancer cells, or from the surrounding stroma. However, detection of biglycan in only one PDTX model, and not within naïve mouse brains. The antibody used in detection of biglycan recognises both human and mouse epitopes, therefore this indicated that biglycan in brain metastases is cancer-cell derived.



**Figure 4.13. Analysis of biglycan expression in patient-derived brain tumour xenografts.** Lysates prepared from the PDTX models of brain tumour xenografts, BCBM2 and BCBM6, were blotted for the expression of human biglycan. Total protein loaded was 10 and 30µg per sample. BCBM samples were analysed in technical duplicates. Biglycan expression was compared to that of naïve MDA-MB-231 and constitutively biglycan-expressing MDA-MB-231 (231/BGN) cells, at 10µg total protein.



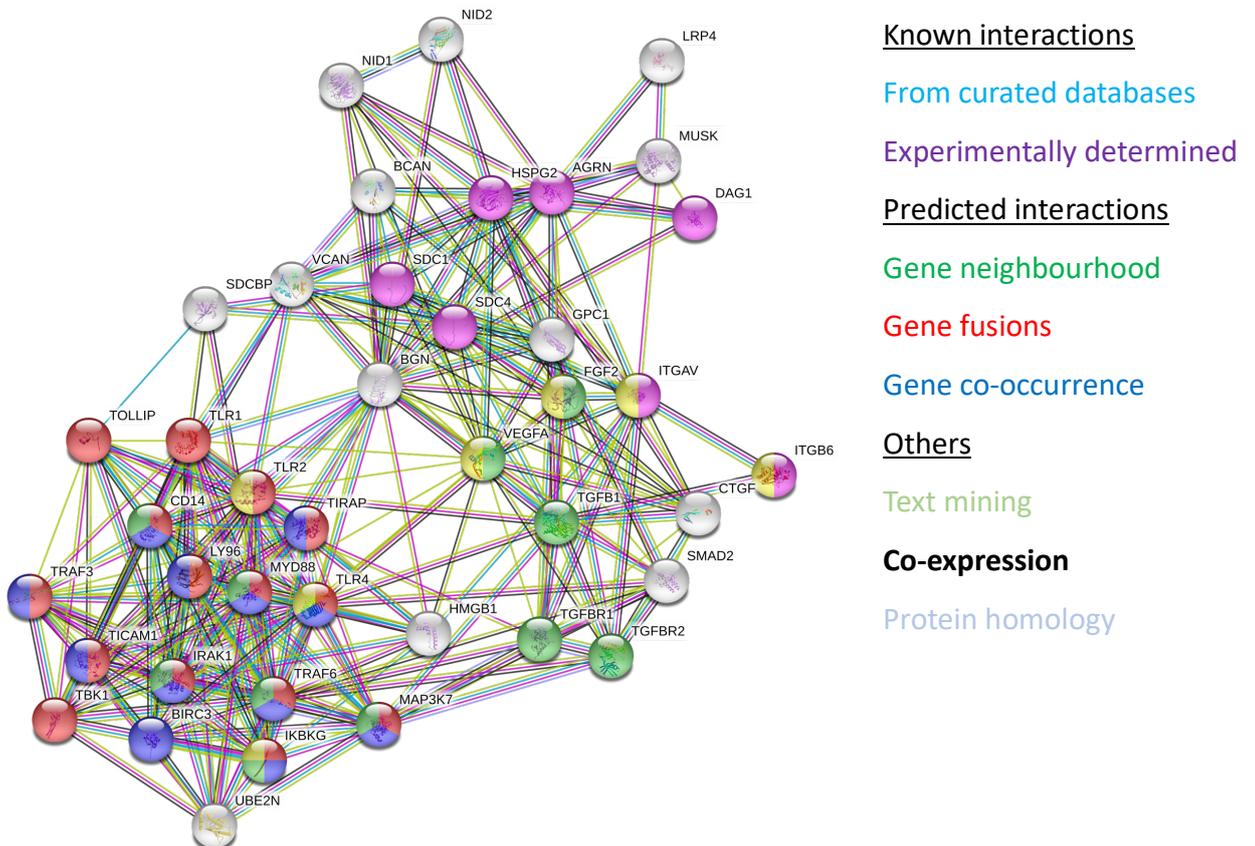
**Figure 4.14. Correlation of biglycan expression in primary breast tumours and relapse-free survival, by subtype.** Breast cancer relapse-free survival data was compared against publicly available mRNA gene chip data from Gene Expression Omnibus (GEO), European Genome-Phenome Archive (EGA) and The Cancer Genome Atlas (TCGA) databases. Kaplan Meier (KM) plotter was used to produce KM survival plots. BGN gene was identified by Affymetrix ID 201261\_x\_at. Patients were split according to median survival. Shown is data for primary breast cancers according to intrinsic subtype. HR = hazard ratio.



**Figure 4.15. Correlation of biglycan expression in primary breast tumours and distant metastasis-free survival, by subtype.** Breast cancer distant metastasis-free survival data was compared against publicly available mRNA gene chip data from Gene Expression Omnibus (GEO), European Genome-Phenome Archive (EGA) and The Cancer Genome Atlas (TCGA) databases. Kaplan Meier (KM) plotter was used to produce KM survival plots. BGN gene was identified by Affymetrix ID 201261\_x\_at. Patients were split according to median survival. Shown is data for primary breast cancers according to intrinsic subtype. HR = hazard ratio.

#### 4.7 Analysis of Biglycan Protein Interaction Network

To gain further insight into the potential roles of biglycan in the context of cancer and dormancy, a protein interaction network of human biglycan was constructed using STRING software [245]. KEGG enrichment analysis of the resulting network identified the Toll-like receptor (TLR), NF- $\kappa$ B, MAPK and PI3K-Akt signalling pathways, in addition to ECM receptor interactions, as enriched interacting pathways (**Figure 4.16**). Considering that biglycan is secreted, this may suggest that the induction of growth arrest by biglycan over-expression may occur in an autocrine fashion via the activation and/or inhibition of intrinsic pathways.



**Figure 4.16. Human biglycan protein interaction network.** The protein network of human biglycan was determined by STRING software [245]. Number of nodes = 41. Number of edges = 276. Minimum required confidence score = 0.400. FDR < 0.05 for all KEGG pathways. The colour of nodes is as follows: Toll-like receptor signalling pathway (red); NF- $\kappa$ B signalling pathway (dark blue); MAPK signalling pathway (green), ECM-receptor interaction (purple); PI3K-Akt signalling pathway (yellow).

## 4.8 Discussion

In this chapter, biglycan was investigated as a potential mediator and/or marker of cellular dormancy in breast cancer. The results suggested a dual role for this small leucine-rich proteoglycan (SLRP). Analysis of publicly available gene expression data revealed that high biglycan expression in primary breast tumours is associated with reduced metastasis-free survival, relapse-free survival and post-progression survival. However, biglycan is significantly downregulated in metastases, including BrM, as compared to the primary breast tumours. To support this, overexpression of biglycan in, and addition of recombinant biglycan to, MDA-MB-231 cells *in vitro* in my study demonstrated anti-proliferative properties. Given that the mRNA sequence of biglycan in dormant breast cancer cells *in vivo* only differs from the known sequence for human biglycan by two silent point mutations, the protein sequence of dormant cell biglycan must be identical to that of endogenous biglycan. Therefore, alternative mechanisms must exist which determine the functional role(s) of biglycan at sites of dissemination, compared to that of primary tumours, either through its expression levels, post-translational modifications, or interactions with the organ-specific microenvironment.

### 4.8.1 Localisation of Dormant Breast Cancer Cells to Brain Parenchymal Vascular Endothelium

Knowledge of the localisation of dormant cells may enable further understanding as to how dormancy is induced and maintained. The difficulty in detecting dormant cells to observe their localisation lies in their scarcity within the brain in my model, with as few as 14 to 15 cells per mouse brain. In my model, a solitary cancer cell, positive for biglycan expression, was observed to be localised adjacent to the vascular endothelium in brain parenchyma. A second detected event identified biglycan secretion from a solitary cell, or a small cluster of cancer cell or cells. Despite no counter-staining for endothelial cells in the second event, morphologically the cell or cells appeared to be wrapped around a secondary structure. A number of published experiments indicates that such a structure could be that of vascular endothelium and should be investigated in future studies. As a secreted protein, as shown when biglycan was over-expressed in MDA-MB-231 cells *in*

*vitro*, biglycan has previously been observed in the pericellular matrix (PCM) surrounding endothelial tubes [297]. Lining the structure of which the cancer cell was wrapped in my study was a ring-like dense deposition of biglycan. It would be worth exploring whether this is as a result of binding to the PCM. It was important to recognise that with so few events, what I have observed may be a chance finding, and the search for more cells should be continued. If these observations are legitimate, vascular localisation of biglycan-positive cancer cells would suggest that dormancy could be induced within the early stages of brain/secondary organ dissemination. Dormancy models have also shown that stable microvasculature is supportive of dormancy induction, whereas sprouting neovascular tips are supportive of micrometastatic outgrowth [133]. More information is therefore required as to the location of biglycan-expressing cancer cells.

#### **4.8.2 Regulation of Biglycan Expression in Dormancy**

In highly metastatic melanomas, tumour endothelial cells with high biglycan expression present with low methylation at the biglycan promotor [297]. Biglycan expression was increased when cells were treated with the demethylating agent 5-aza-dC, indicating that biglycan expression can be regulated by epigenetic modification [297]. In my study, the TGF- $\beta$  signalling pathway was predicted to be enriched in dormant cells. An early study of biglycan expression suggested that a TGF- $\beta$ -responsive element is present within the biglycan promotor, upstream of the transcriptional start site and also containing a number of binding sites for the specificity protein 1 (Sp1) transcription factor [298]. A further study defined that TGF- $\beta$  induces biglycan expression through the Smad-activating role of ALK5 [299]. In addition to Sp1, v-ets erythroblastosis virus E26 oncogene homolog (ETS) family members and activator protein 1 (AP-1) transcription factors have also been shown to differentially interact with the biglycan promotor [300]. ETS transcription factors, critical nuclear regulators of ubiquitous signalling cascades, are targets of the MAPK signalling pathway [301]. This is significant as MAPK signalling has previously been strongly associated dormancy induction [54]. Ultimately, the mechanisms of biglycan overexpression in dormant cancer cells are unknown, however the BGN promotor region presents with binding sites for transcription factors

downstream of a number of pathways. Further study is required in order to understand BGN regulation in dormancy, and to determine whether it is TGF- $\beta$ -dependent.

#### 4.8.3 A Dual Role of Biglycan?

It is interesting that I found that high biglycan expression within primary breast tumours correlated with a reduction in distant metastasis-free survival, relapse-free survival and progression post-survival. The effect of this was dependent on receptor status, with HER2+ breast cancers associated with the poorest clinical outcome. Conversely, analysis of publicly available datasets identified that biglycan expression was reduced significantly in breast cancer brain, liver, bone and lung metastases. The anti-proliferative properties of biglycan demonstrated in my study *in vitro* support this observation, using the metastatic MDA-MB-231 cell line. As such, there is clear contrasting evidence of biglycan-associated mechanisms between primary and metastatic breast cancers, and it will be critical to understand whether there is a dual role of biglycan.

#### 4.8.4 Biglycan Mechanisms of Action

As shown by STRING analysis, biglycan sits at the centre of a complex network of molecular pathways. Perhaps one of the most significant, and enriched within my dormant cell population *in vivo*, is TGF- $\beta$  signalling. In the development of breast cancer metastases, active TGF- $\beta$ 1 has been identified as a tumour-promoting ligand at the neovascular tips [133]. Extracellular biglycan, along with decorin and lumican, other SLRPs which were upregulated in dormant cells in my model, can sequester TGF- $\beta$  ligands, limiting TGF- $\beta$  bioavailability [302]. Further work should determine whether biglycan was suppressing TGF- $\beta$ -mediated metastatic outgrowth *in vivo*. For example, in other models, mice lacking biglycan are observed to display increased TGF- $\beta$  activity [303], whilst I observed that biglycan expression is reduced in brain metastases compared to primary breast cancers. However, the mechanism of biglycan action on TGF- $\beta$  activity, if any, may be contextually dependent. Importantly, biglycan overexpression, and addition of recombinant biglycan to, MDA-MB-231 cells *in vitro*,

demonstrated growth-inhibitory properties. Whilst MDA-MB-231 cells were cultured *in vitro* for decades, potentially losing some of their original phenotype, they were derived from a lung pleural effusion and were metastatic at the point of harvest [304]. One should therefore be careful when making comparisons to data from primary breast cancers. Furthermore, although it is unknown as to whether biglycan can induce dormancy via inhibition of TGF- $\beta$  signalling, there is evidence to suggest high and low affinity binding sites for TGF- $\beta$  sequestration by biglycan, with  $K_d$  values of 1 – 20nM and 20 – 200nM [302]. This may explain why MDA-MB-231 cell growth-arrest could only be achieved *in vitro* using high concentrations of lentivirus, and far greater inhibition than adding recombinant biglycan. It would be useful to compare the levels of biglycan in supernatant from over-expressing cells, vs those with added recombinant biglycan.

In primary breast cancers, a significant source of TGF- $\beta$  is from infiltrating macrophages, which induces lysyl oxidase (LOX)-mediated increase in stiffness, promoting proliferation [106, 288]. In normal brain tissues, TGF- $\beta$  levels are very low or not detectable at all [305]. TGF- $\beta$  signalling was enriched in dormant breast cancer cells in my model, therefore it should be explored whether this is due to a source of TGF- $\beta$  ligands from neovascular tips, or independent sources. In my *in vitro* studies, there were no extracellular sources of TGF- $\beta$  ligands other than those secreted by other MDA-MB-231 cells. *In vitro*, extracellular TGF- $\beta$  was not required to sustain proliferation, whereas growth inhibition was observed when biglycan was overexpressed, or added. It is not clear from my study as to how biglycan is expressed, or what its mechanism of action is. However, c-Myc transcription factor was repressed in its activity in dormant cells in my *in vivo* model. In proliferating cells, TGF- $\beta$  signalling leads to c-Myc repression and a cytostatic response, mediated by p21 [306]. This is consistent with my KEGG analysis that TGF- $\beta$  signalling is activated in dormant cells. There is also evidence to suggest that biglycan directly binds BMP-2 and its receptors, positively modulating its activity [307]. In breast cancer, activation of BMP-2 significantly inhibits the proliferation of MDA-MB-231 cells, marked by a p21-mediated G1 arrest [308]. In my dormant cells *in vivo*, the promoter of breast cancer proliferation, *ID1* gene, was significantly downregulated downstream of BMP signalling, therefore it would be useful to explore biglycan-induced growth arrest in the context of the wider TGF- $\beta$  superfamily. In addition to TGF- $\beta$ ,

biglycan may also act as an autocrine ligand for both TLR2 and TLR4, leading to rapid downstream activation of p38, ERK and NF- $\kappa$ B signalling pathways [309]. It has been shown that TLR4 knockdown in MDA-MB-231 cells inhibits growth and proliferation [310]. As well as being in contradiction of the data presented in my study, p38 is associated with the dormancy phenomenon, and not with metastatic progression [196, 197]. In terms of NF- $\kappa$ B, some studies have indicated that whether NF- $\kappa$ B promotes quiescence or proliferation is context-dependent. Constitutive activation of inhibitor of nuclear factor kappa-B (IKK $\beta$ ) in ER+ breast cancer inhibits E2-induced proliferation *in vitro* and suppresses tumour growth *in vivo* [311]. Importantly, growth inhibition was reversible and therefore a hallmark of a dormant phenotype [311]. It was finally noted that in co-operation with ER, NF- $\kappa$ B was also capable of promoting invasion and migration *in vitro*, whilst being a potent promotor of metastasis *in vivo* [311]. Overall, this suggests that biglycan could be a promotor of metastasis within the primary tumour via NF- $\kappa$ B activation, while it may induce dormancy in the context of brain metastases.

#### 4.8.5 Summary

I have demonstrated that biglycan-expressing breast cancer cells within the brain associate with vascular endothelial cells, however more events should be captured. This should be carried out whilst exploring the potential role of the perivascular niche in dormancy induction. Interestingly, biglycan expression within primary tumours was positively correlated with a reduction in relapse-free- and distant metastasis-free survival. In breast cancer metastases, by analysis of publicly available datasets, biglycan was reduced in all analysed secondary sites compared to that of primary tumours. This was corroborated in my *in vitro* studies in which biglycan overexpression and addition of recombinant biglycan to MDA-MB-231 cells demonstrated growth inhibitory properties. Biglycan may represent a novel and exploitable target for the treatment of breast cancer brain metastases, however significant work is required to deduce its mechanism of expression in dormant cancer cells, and how biglycan induces growth arrest.

## Chapter 5

### Characterisation of Aerobic Glycolysis in the Induction and/or Maintenance of Cellular Dormancy

#### 5.1 Introduction

Transcriptomic data analysis of dormant and proliferating MDA-MB-231 cells *in vivo* identified an enrichment of GO terms pertaining to cellular metabolism. Of particular interest was aerobic glycolysis. Dormant cells which are reversibly arrested often display a lowered metabolic state [312]. However, there is little evidence to suggest that a lowered metabolic state may actually be a driver of dormancy, rather than its consequence.

Aerobic glycolysis, otherwise referred to as the “Warburg effect”, is characterised by high glucose uptake and high rates of inefficient glycolysis, preferred over oxidative phosphorylation as a result of metabolic reprogramming, even at normal oxygen concentrations [313]. As the main fuel source for energy metabolism in the brain, D-glucose is distributed throughout the organ using brain vasculature [314]. D-glucose then enters the brain rapidly across the blood-brain-barrier, mediated by facilitated diffusion [314]. In chapter 4, a biglycan-positive cancer cell, and therefore predicted to be dormant, was identified to be localised adjacent to the brain microvasculature. As such, dormant cells are likely to receive a plentiful supply of glucose and its supply is unlikely to be a rate limiting factor. Differential gene expression analysis identified a significant downregulation in many enzymes involved in aerobic glycolysis in dormant cells *in vivo*. I also predicted that this may be the result of a repression of the transcription factors c-Myc and HIF-1A. It may be that transcriptional downregulation of aerobic glycolysis is a driver of dormancy, as opposed to being an effect of low energy requirements in growth-arrested cells.

### 5.1.1 Hypothesis

The hypothesis for this chapter is that a downregulation of aerobic glycolysis at the gene expression level can induce long-term and reversible growth arrest in breast cancer cells *in vitro*. Secondly, I hypothesised that such growth-arrested breast cancer cells *in vitro* display a comparable phenotype to dormant cancer cells isolated *in vivo*.

### 5.1.2 Aims and Objectives

This chapter aims to explore the effects of inhibited aerobic glycolysis on breast adenocarcinoma cells *in vitro*, compared to the mRNAseq profiling data of dormant and proliferating breast cancer cells *in vivo*, as explored in Chapter 3. To achieve this, this chapter aims to:

**1) Identify whether long-term and reversible growth arrest of breast cancer cells *in vitro* can be induced by inhibiting aerobic glycolysis**

This will be achieved by determining which genes are downregulated in dormant breast cancer cells *in vivo*, which will aid selection of an appropriate aerobic glycolysis inhibitor. The inhibitor will be titrated against breast cancer cells *in vitro*, to determine whether there is an optimal concentration at which survivable growth-arrest is achieved, and whether this effect is reversible, as in dormancy.

**2) On successful completion of aim (1), determine functional effects of inhibiting aerobic glycolysis *in vitro* and compare to the *in vivo* model employed in Chapter 3**

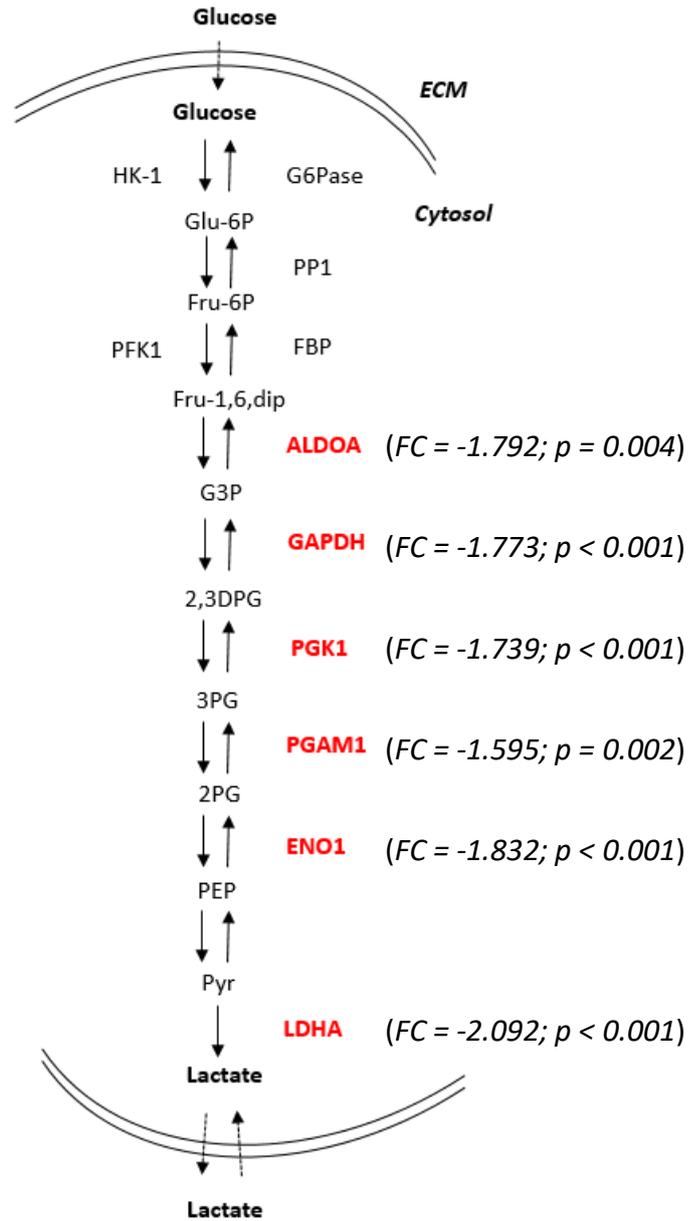
To achieve this, a number of functional *in vitro* assays, notably for gene expression, cell cycle phase analysis, protein localisation and protein phosphorylation, will be employed. Suitable candidates will be determined based on previously identified *in vivo* data regarding dormant cancer cells in the brain. These experiments will represent preliminary data on the effects of inhibited aerobic glycolysis, to inform the construction of suitable *in vivo* models in future studies in the exploration of novel therapeutic targets.

## 5.2 Inhibition of Aerobic Glycolysis Induces Reversible Cancer Cell Growth Arrest *In Vitro*, Mimicking Cancer Cell Dormancy

Genes contributing to the GO term “Enrichment in aerobic glycolysis” that were differentially regulated between dormant and proliferating cancer cells *in vivo* (gene expression data derived from processing streams I and II) were overlaid with a published schematic of the glycolysis pathway (**Figure 5.1**) [213].

All differentially expressed genes associated with the glycolysis pathway were downregulated in dormant cells, with fold changes ranging from -1.659 to -2.092, and encoded for the enzymes involved in the pathway downstream of the fructose 1, 6-bisphosphate (Fru-1,6-dip) intermediate. As it was unclear which of the enzymes may contribute to a dormant phenotype, I chose to inhibit aerobic glycolysis upstream of all differentially expressed genes. The glucose analogue, 2-deoxy-D-glucose (2DG), is metabolised to 2DG-6-phosphate (2DG-6P) by hexokinase (HK), and competitively inhibits glucose transporters (GLUTs) [316]. The 2DG-6P intermediate cannot be metabolised further and it non-competitively inhibits HK and competitively inhibits glucose-6-isomerase (PGI) [316]. The use of 2DG has previously been explored as a cancer therapeutic, given its ability to preferentially be captured by tumours and induce cell cycle arrest and apoptosis [317]. I therefore sought to determine whether the inhibition of aerobic glycolysis using 2DG can induce growth arrest in breast cancer cells *in vitro* whilst maintaining cell viability.

MDA-MB-231 cells were cultured under high glucose conditions, mimicking the high glucose availability within the brain [314]. 2DG was added to the growth medium at 0, 5, 10, 20 and 50mM concentrations. Cell counts were determined every 3 days, quantifying the number of viable and non-viable adherent and floating cells. As shown in **Figures 5.2A, B, C and D**, cell counts at 5mM 2DG were significantly lower than untreated cells ( $p \leq 0.001$  at days 3, 6 and 9 post-2DG addition), however the cells retained some capacity to proliferate.

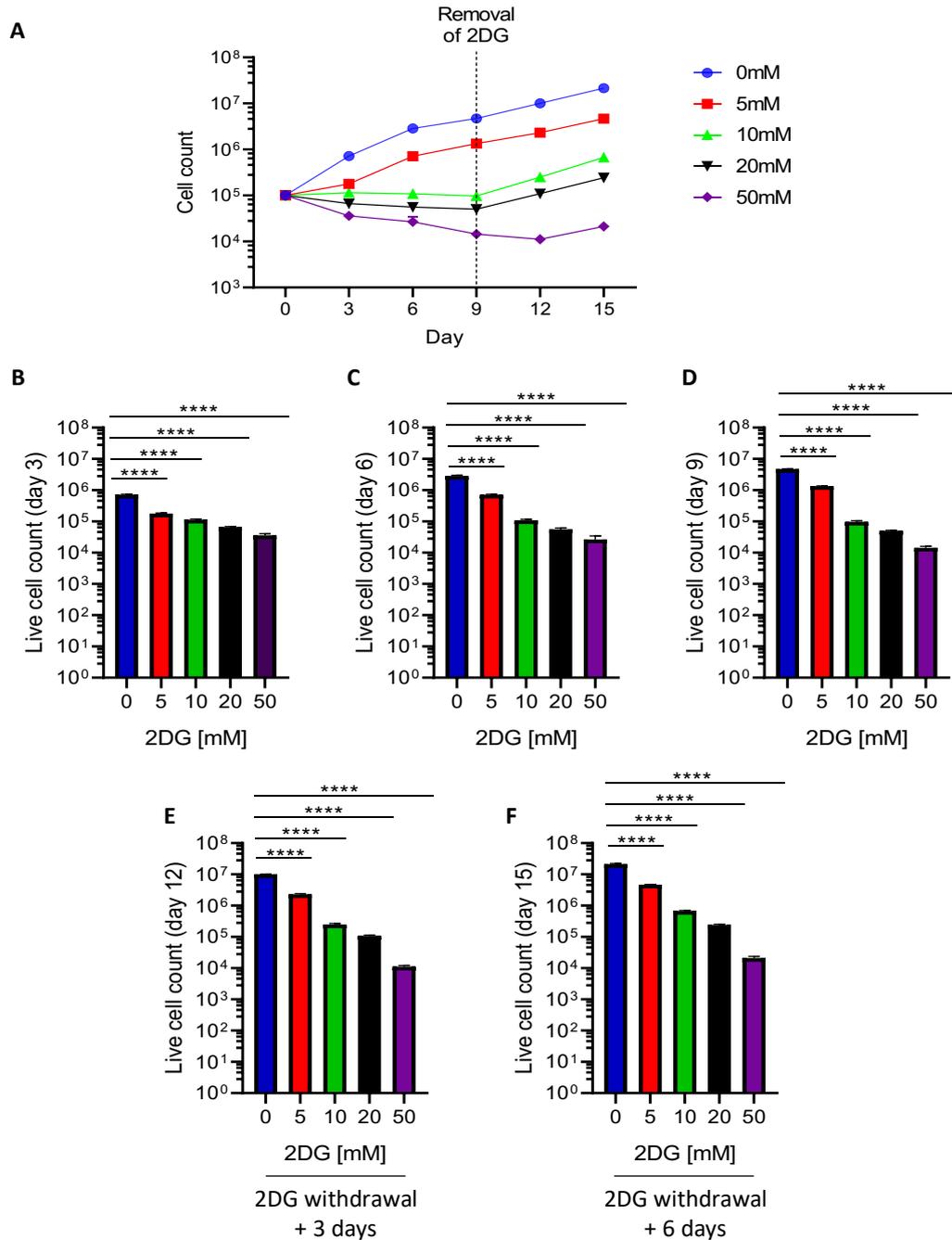


**Figure 5.1. Core enzymes in the glycolysis pathway are downregulated in dormant cells *in vivo*.** Simplified schematic of the conversion of glucose to lactate in glycolysis. Enzymes shown in red were downregulated in dormant cells *in vivo*, as determined by differential gene expression analysis. All enzymes were observed to be downregulated in both processing streams I and II. Hexokinase-1 (HK-1); glucose-6-phosphatase (G6Pase); glucose-6-phosphate (Glu-6P); protein phosphatase 1 (PP1); fructose-6-phosphate (Fru-6P); fructose-1,6-diphosphate (Fru-1,6-dip); aldolase (ALDOA); glyceraldehyde-3-phosphate (G3P); glyceraldehyde-3-phosphate dehydrogenase (GAPDH); 2,3-diphosphoglycerate (2,3DPG); phosphoglycerate kinase 1 (PGK1); 3-phosphoglyceric acid (3PG); phosphoglycerate mutase 1 (PGAM1); 2-phosphoglyceric acid (2PG), enolase 1 (ENO1); phosphoenolpyruvate carboxylase (PEP); pyruvate (Pyr). Fold change = FC. Adapted from [213].

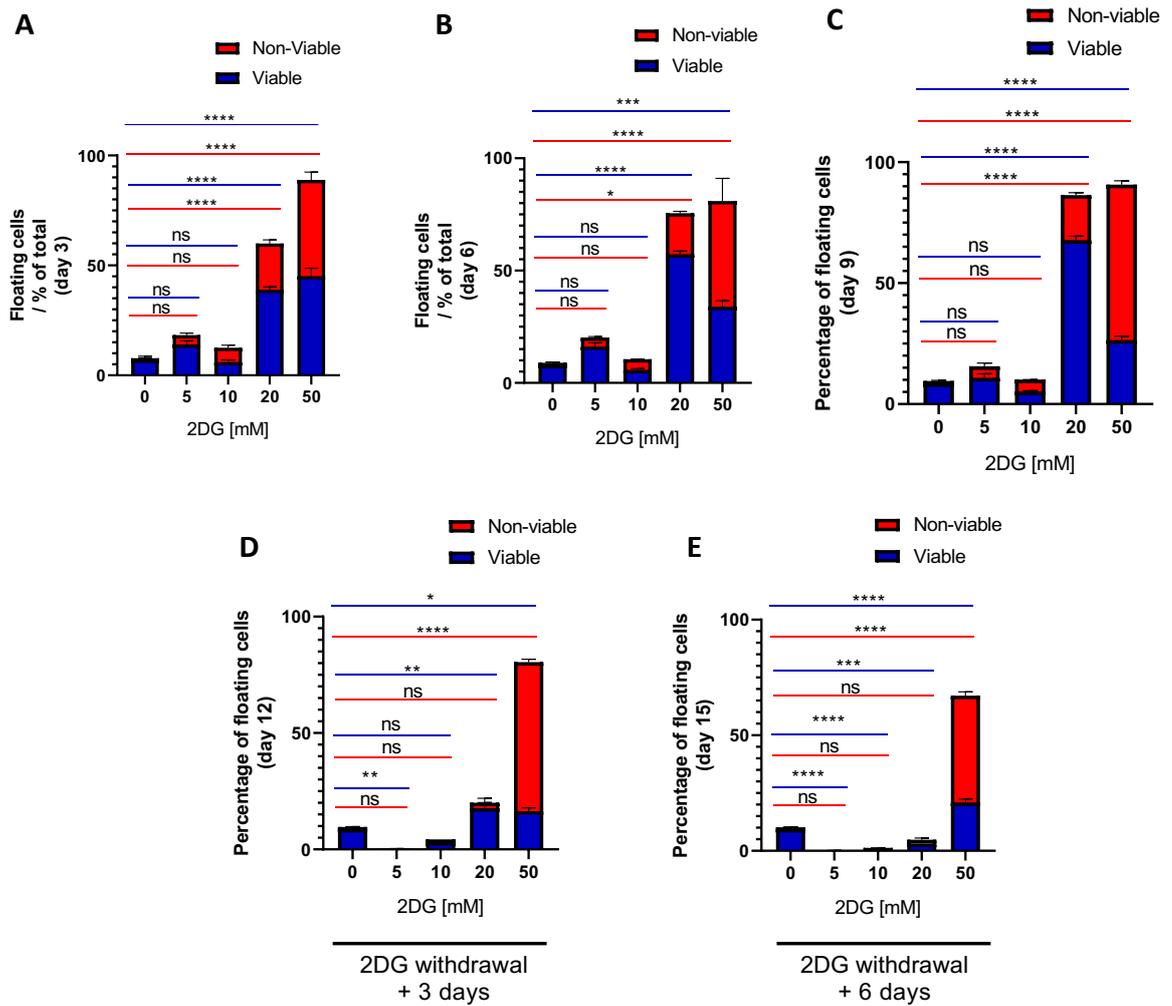
At 10mM 2DG, cell counts were also significantly lower than that of untreated cells and presented with complete growth arrest whilst retaining viability ( $p \leq 0.001$  at days 3, 6 and 9 post-2DG addition) (**Figures 5.2A, B C and D**). At concentrations greater than 10mM 2DG, there was an inversely proportional concentration-dependent decrease in cell count ( $p \leq 0.001$  at days 3, 6 and 9 post-2DG addition) (**Figures 5.2A, B C and D**). Interestingly, and pertinent to the dormancy phenomenon, growth arrest was reversible, as cellular proliferation resumed upon withdrawal of 2DG at all concentrations (**Figures 5.2A, E and F**). This suggested that inhibition of glycolysis, at a specific level, may be sufficient to induce reversible growth arrest in MDA-MB-231 cells *in vivo*.

Furthermore, as shown in **Figures 5.3A, B, C**, 5mM 2DG displayed a similar number of floating cells to that of untreated cells. At 10mM 2DG, at which cell growth was reversibly inhibited, I observed a small increase in the number of detached viable cells, however this was statistically insignificant (**Figures 5.3A, B, C**). At increasing concentrations of 2DG over 10mM, in which I observed concentration-dependent cell death, there was an increase in the numbers of detached non-viable cells (**Figures 5.3A, B, C**). Although there was no significant detachment of cells at 10mM 2DG or lower, withdrawal of 2DG, marked by recovery in cell count (**Figure 5.2**), was concurrent with a reduction in the proportion of detached viable and non-viable cells (**Figures 5.3D and E**). This demonstrates that inhibition of aerobic glycolysis, whilst also inducing growth arrest, may also be reversibly inhibiting the ability of cancer cells to adhere to the extracellular matrix. However, this cannot be concluded without further investigation, as the composition of the brain and the brain's ECM is vastly more complex than the plastic used to culture cells *in vitro*.

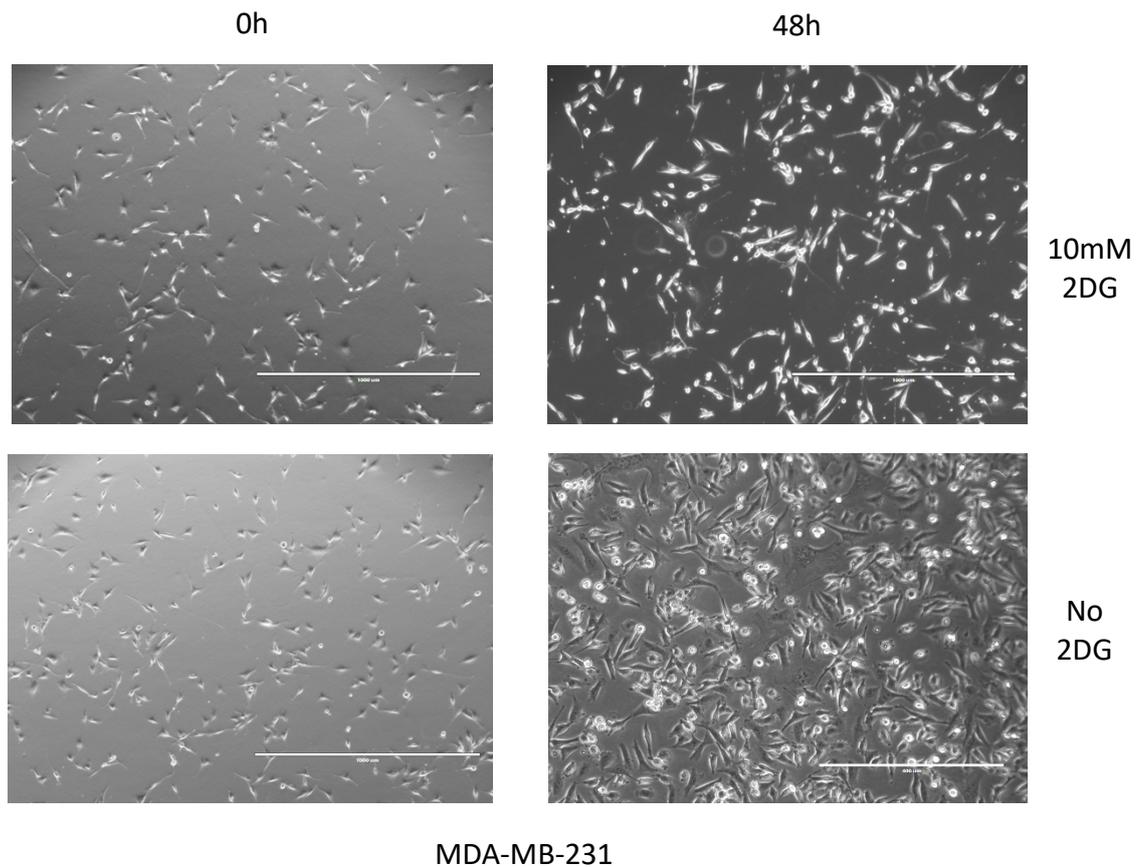
To observe MDA-MB-231 cells under 2DG-induced growth inhibition, 10mM 2DG was added to the cells at 24 hours post-seeding and incubated for 48 hours. As expected, cell growth was inhibited upon addition of 2DG, whilst cells retained normal size and morphology, comparable to that of untreated cells (**Figure 5.4**).



**Figure 5.2. Inhibition of aerobic glycolysis with the glycolysis inhibitor, 2DG, induces reversible growth arrest in MDA-MB-231 cells *in vitro* (I).** MDA-MB-231 cells were cultured *in vitro* with the glycolysis inhibitor, 2DG. The amount of 2DG required to induce growth arrest was determined by titration, using 0 (negative control), 5, 10, 20 and 50mM 2DG. After 9 days, 2DG was withdrawn and cells were cultured for a further 6 days. A) Live cell count of both adherent and detached cells at each time point shown. Withdrawal of 2DG is shown by the indicated dotted line. B-F) Quantitative comparison of the total number of live cells (adherent + floating) between each concentration of 2DG, at the time point shown.  $N=3$ . Shown is mean count  $\pm$  SEM. One representative experiment out of two is shown.



**Figure 5.3. Inhibition of aerobic glycolysis with the glycolysis inhibitor, 2DG, induces reversible growth arrest of MDA-MB-231 cells *in vitro* (II).** MDA-MB-231 cells were cultured *in vitro* with the glycolysis inhibitor, 2DG. The amount of 2DG required to induce detachment from plastic substrate was determined by titration, using 0 (negative control), 5, 10, 20 and 50mM 2DG. At day 3 (A), day 6 (B) and day 9, the number of floating cells, taken as a percentage of all cells at that time point, is shown. At day 9, 2DG was withdrawn and the percentage of floating cells was recorded at day 12 (D) and day 15 (E). Also shown is the ration of viable and non-viable floating cells.  $N=3$ . Shown is the mean count  $\pm$  SEM. One representative experiment out of two is shown.

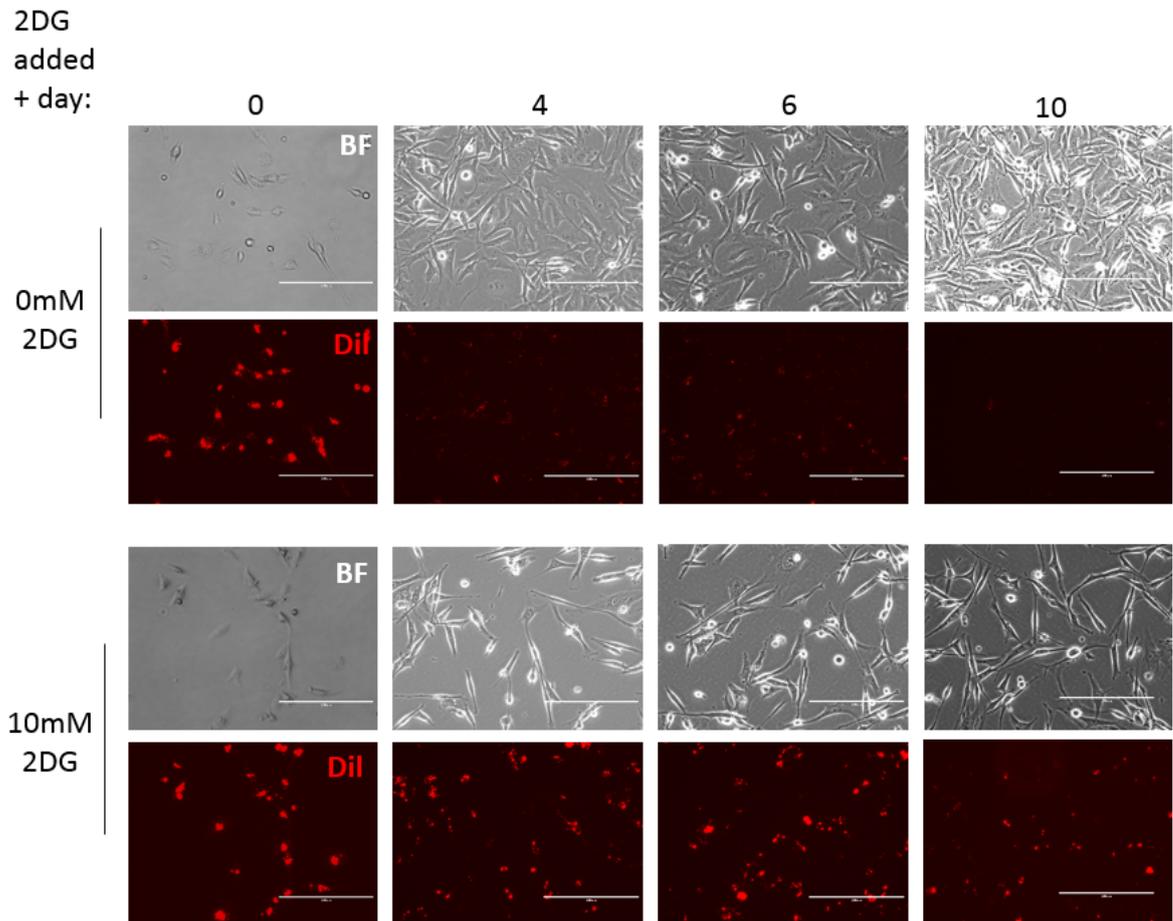


**Figure 5.4. Inhibition of aerobic glycolysis with 10mM 2DG inhibits growth of breast adenocarcinoma cells whilst retaining normal cell size and morphology.** MDA-MB-231 cells were cultured in the presence or absence (negative control) of 10mM 2DG. Representative images were obtained at 0h and 48h post-addition of 2DG using brightfield microscopy. Scale bar = 1mm. One representative experiment out of three is shown.

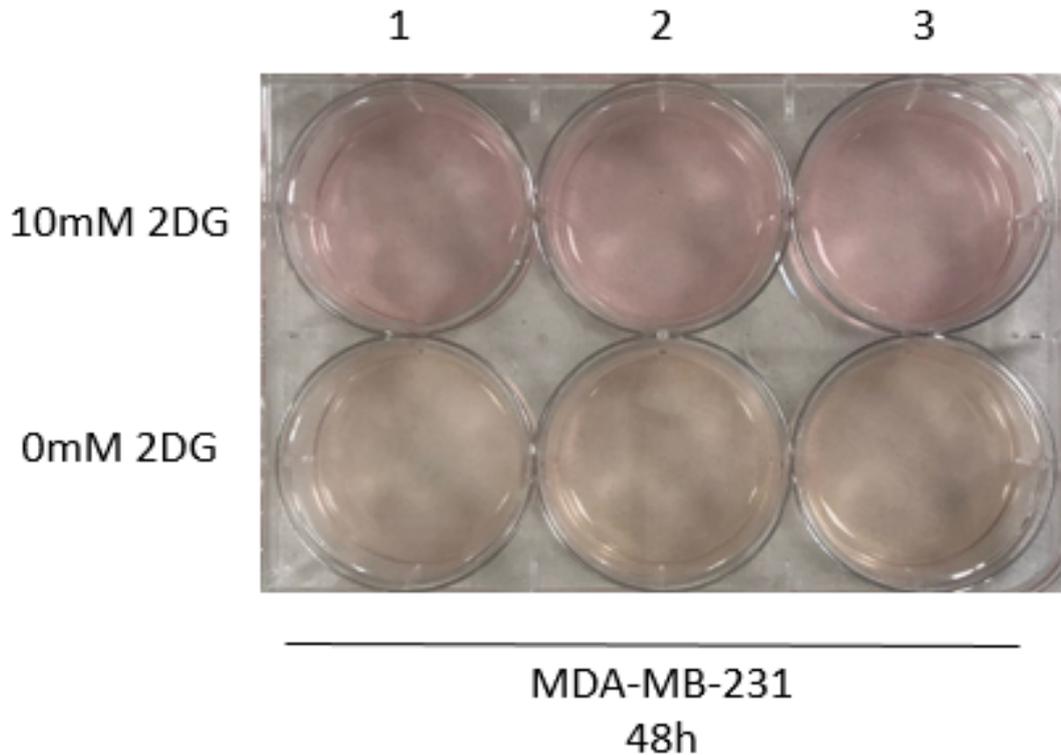
Next, MDA-MB-231 cells were labelled with Vybrant™ Dil lipophilic membrane dye and cultured in the presence or absence of 10mM 2DG (**Figure 5.5**). In untreated cells, the membrane dye was lost through proliferation within 10 days. This result is in line with the data reported in Chapter 3, showing that CV dye is lost rapidly through proliferation *in vitro*. When Dil-labelled cells were incubated with 2DG, over the same 10-day time period, cells established themselves as a viable monolayer, retaining normal morphological characteristics. Notably, the membrane dye was retained after 10 days under growth-arrested conditions. Whilst this shows that Dil can be used to track proliferation *in vitro*, there was a slight loss of intensity at day 10. This may be due to dye diffusion into the growth medium. Overall, this demonstrates long-term survivability of breast cancer cells that have been growth-arrested through inhibition of aerobic glycolysis. This infers that MDA-MB-231 cells are able to maintain, at least to some extent, physiological functions pertinent to cellular survival, whilst undergoing a reduced metabolic state. However, this may only be the case *in vitro* and there may be many other factors which influence cancer cell fate *in vivo*. Future study should aim to assess cancer cell growth and proliferation in the brain, whilst aerobic glycolysis is inhibited.

Cancer cells are known to increase the expression of glycolysis enzymes, consistent with increased rates of glycolysis [316]. This leads to an increase in extracellular acidity, which is supportive of pro-metastatic signalling cascades and chemoresistance [318]. I hypothesised that a reduction in extracellular acidity may therefore support the growth arrest phenotype of dormant cells [319]. Here, MDA-MB-231 cells were cultured *in vitro* in the presence and absence of 10mM 2DG, in medium containing phenol red pH indicator (**Figure 5.6**). As expected, untreated cells displayed yellowing of the medium, consistent with pH indicator-mediated colour change as a result of extracellular acidification. Those treated with 10mM 2DG did not undergo the same colour change, indicating a comparatively alkaline environment compared to untreated cells. This demonstrates that 2DG limits the secretion of lactate and protons into the extracellular space. No conclusions can be drawn as to whether this translates to any functional effects on proliferation, however. In my *in vivo* model, cancer cells exist amongst local brain populations which may also modulate extracellular pH. This experiment simply

demonstrates a reduced metabolic output in cancer cells treated with 2DG *in vitro*. Whether dormant cells can generate a relatively more alkaline microenvironment *in vivo* is for further investigation.



**Figure 5.5 – Tracking the growth-inhibitory effects of 2DG on MDA-MB-231 cells by membrane staining.** MDA-MB-231 cells were labelled with the lipophilic cell membrane stain, Vybrant™ Dil. Labelled cells were cultured under high glucose conditions, in the presence or absence (negative control) of 10mM 2DG. Images were obtained via brightfield (BF) microscopy at the times indicated. Fluorescence intensity of Dil membrane dye, obtained by epifluorescence microscopy, was obtained in parallel. Images are representative of three independent experiments. Scale bar = 120µm.



**Figure 5.6. Inhibition of aerobic glycolysis in MDA-MB-231 cells prevents acidification of growth medium *in vitro*.** MDA-MB-231 cells were cultured in the presence or absence (negative control) of 10mM 2DG, for 48h. The number of cells plated per well, of a 6-well culture plate, was  $1 \times 10^5$  at the time of seeding. Growth medium contained phenol red, which changes from red, to orange, to yellow, in response to increasing acidity. Shown are representative images of growth medium, 48h post-addition of 2DG. Shown are technical triplicates, denoted by 1, 2, 3. One representative experiment out of three is shown.

### 5.3 2DG-Induced Growth Inhibition does not Influence Glycolysis Enzymes or Biglycan Expression *In Vitro*

Transcriptomic data analysis of MDA-MB-231 cells *in vivo* identified a number of downregulated genes pertaining to aerobic glycolysis. I therefore determined whether the addition of 2DG to MDA-MB-231 cells *in vitro* also induces similar downregulation in gene expression. MDA-MB-231 cells were incubated with 10mM 2DG and the gene expression of *PGK1* and *GAPDH* was determined at 48h. Both genes were observed to be downregulated in dormant cells *in vivo* at the gene expression level. Despite a tendency towards downregulation in 2DG-treated cells, there was no significant change

in expression of *PGK1* nor *GAPDH* (Table 5.1) in cells treated with 2DG, compared to those cultured in the absence of 2DG. This result is expected and confirms that the mechanism of action of 2DG is through a restriction of available glucose and does not induce a downregulation of glycolysis genes. Thus, mechanisms of glycolysis downregulation *in vivo* and in my *in vitro* model using 2DG may be different.

	<i>PGK1</i>		<i>GAPDH</i>	
	10mM 2DG	0mM 2DG	10mM 2DG	0mM 2DG
<b>Fold change (SE)</b>	1 (0.35)	1.25 (0.41)	1 (0.33)	1 (0.41)
<b>Significant difference?</b>	No		No	

**Table 5.1. Inhibition of aerobic glycolysis using the glycolysis inhibitor, 2DG, does not inhibit the expression of core glycolysis genes.** MDA-MB-231 were cultured in the presence or absence (0mM/negative control) of 10mM 2DG. Cells were harvested after 48h. Shown is qPCR analysis of *PGK1* and *GAPDH* expression, relative to a housekeeping gene, *POLR2A*. N=3. Shown is mean fold change in expression +/- SEM. Statistical test performed was an unpaired two-tailed *t* test. One representative experiment out of three is shown.

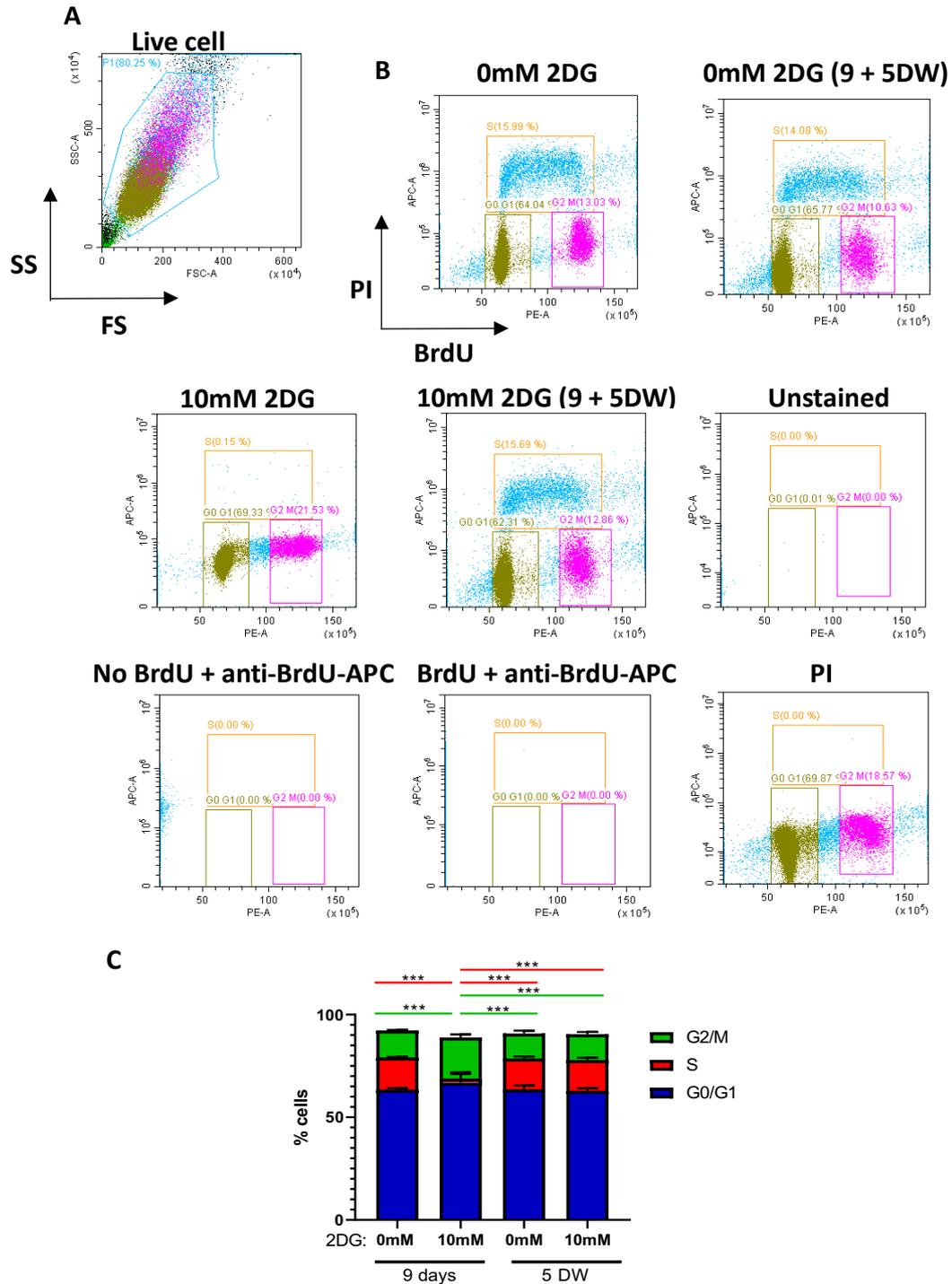
Furthermore, I then sought to determine whether *BGN* expression could be influenced through the inhibition of aerobic glycolysis. *BGN* expression was determined at 0, 24 and 48h post-addition of 10 mM 2DG (Table 5.2). There was no significant difference in *BGN* expression in 2DG-treated cells, as compared to untreated cells. This suggested that biglycan may be upstream of aerobic glycolysis, that it inhibits growth of MDA-MB-231 cells in a glycolysis independent manner, or that it isn't involved. Like the other *in vitro* experiments presented in chapter 5, no conclusions regarding *BGN* expression can be made, given that there may be many factors *in vivo* which influence its expression.

	<b>BGN</b>					
	<b>0mM</b>	<b>10mM</b>	<b>0mM</b>	<b>10mM</b>	<b>0mM</b>	<b>10mM</b>
	<b>2DG</b>	<b>2DG</b>	<b>2DG</b>	<b>2DG</b>	<b>2DG</b>	<b>2DG</b>
	<b>(0h)</b>	<b>(0h)</b>	<b>(24h)</b>	<b>(24h)</b>	<b>(48h)</b>	<b>(48h)</b>
<b>Fold change (SE)</b>	0.71 (0.29)	1.16 (0.12)	0.74 (0.16)	0.45 (0.64)	2.97 (0.68)	2.40 (0.75)
<b>Significant difference?</b>	No		No		No	

**Table 5.2. Inhibition of aerobic glycolysis does not affect the expression of biglycan *in vitro*.** MDA-MB-231 cells were cultured *in vitro* in the presence or absence (0mM/negative control) of 10mM 2DG for 0, 24 and 48h. Cells were harvested at the indicated time points and analysed by qPCR for the expression of *BGN*. *BGN* expression was normalised to that of a housekeeping gene, *POLR2A*. N=3. Statistical analysis performed was a one-way ANOVA with multiple comparisons test. One representative experiment out of three is shown.

#### 5.4 Inhibition of Aerobic Glycolysis Induces Reversible G2/M Phase Cell Cycle Arrest

The defining phenotype of dormant cancer cells is their ability to undergo a reversible cell cycle arrest at disseminated sites [320]. Therefore, I investigated the effects of 2DG-induced inhibition of aerobic glycolysis on the cell cycle at 9 days post-2DG (10 mM) addition, and at 5 days following 2DG withdrawal from these cultures. Live cells were gated according to forward scatter and side scatter (**Figure 5.7A**). Within the live cell gate, the proportion of cells in G0/G1, S and G2/M phases were determined according to relative BrdU and PI intensity (**Figure 5.7B**). Inhibition of aerobic glycolysis with 2DG induced a significant reduction in percentage of cells within S phase, with a significantly increased proportion of cells in G2/M phase. 2DG treatment did not alter the proportion of cells in G0/G1 phase (**Figures 5.7B, C**). Following withdrawal of 2DG, this effect was reversed, with the 2DG-treatment group displaying cell cycle phase phenotype comparable to that of the control group (**Figures 5.7B, C**).



**Figure 5.7. Inhibition of aerobic glycolysis induces reversible G2/M phase arrest in MDA-MB-231 cells *in vitro*.** MDA-MB-231 cells were cultured for 9 days in the presence or absence (negative control) of 10mM 2DG, under high glucose conditions. After 9 days, 2DG was withdrawn and cells were cultured for a further 5 days (5 DW). A) Live cells were gated by forward scatter (FSC) and side scatter (SSC). B) Bromodeoxyuridine (BrdU) and propidium iodide (PI) were used to gate cells according to cell cycle phase: G0/G1 (green gate), S (orange gate), G2/M (purple gate). Also shown are appropriate controls, as named. C) Quantification of the proportion of live cells within each cell cycle phase. N=3. Error bars represent +/- SEM. Number of independent experiments = 3.

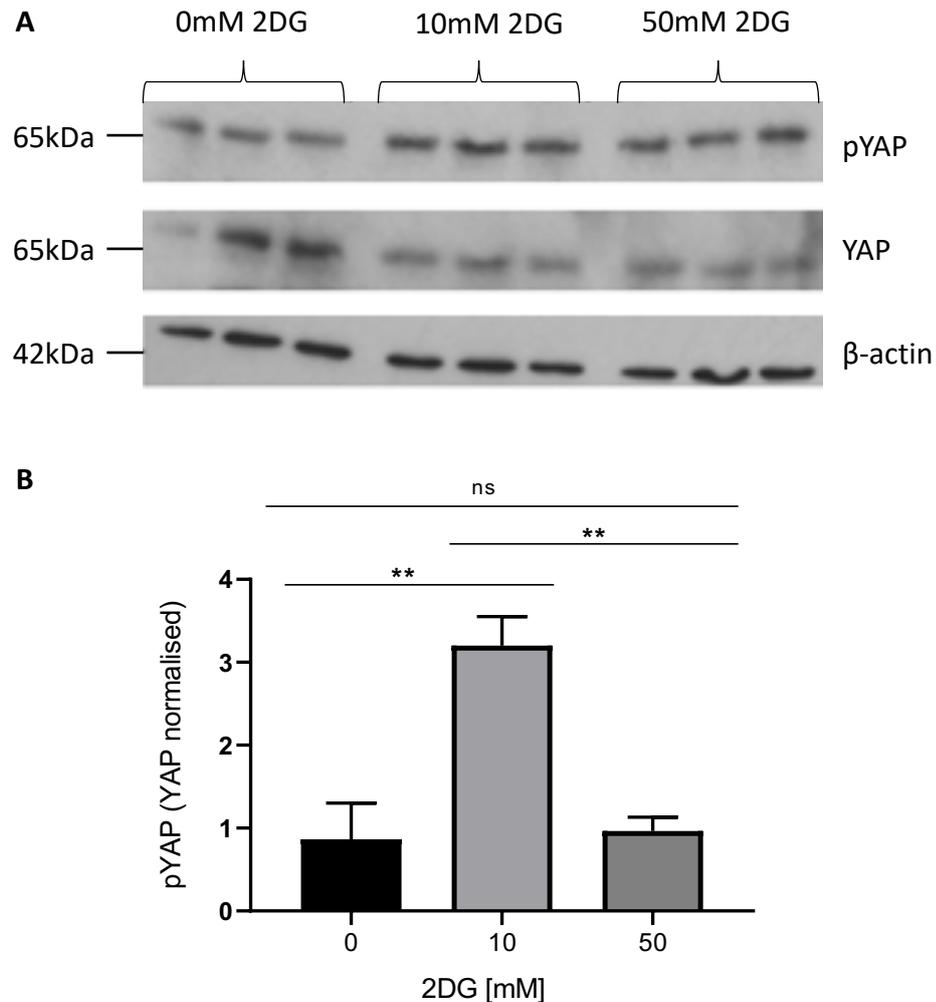
## 5.5 Inhibition of Aerobic Glycolysis may Induce Dormancy in Disseminated Breast Cancer Cells by Modulating YAP Activity

At the centre of the HIPPO signalling pathway, enriched in dormant cells *in vivo*, is the core transcriptional activator, YAP [321]. Dephosphorylated YAP is available to undergo nuclear shuttling, which activates transcription via interaction with a number of co-factors, for example TEADs, SMADs and  $\beta$ -catenin [322, 323, 324]. In cancer, YAP plays a role in tumorigenesis, metastasis and proliferation [321].

Inhibition of YAP transcriptional ability blocks cancer proliferation and perivascular niche-mediated metastatic outgrowth [136]. It has previously been shown that an inhibition of aerobic glycolysis in MDA-MB-231 cells *in vitro* decreases YAP transcriptional activity and impairs colony forming ability [293]. This was proposed to occur by blocking the ability of PFK1, an enzyme in the first step of glycolysis, to bind the YAP co-factors, TEADs [293]. This study was however performed using 2DG at 50 and 100mM concentrations [293]. In my study, at concentrations above 10mM, it was observed that 2DG *in vitro* induces cellular detachment and a concentration-dependent reduction in cell viability. I therefore here investigated the effects of glycolysis inhibition on YAP under conditions that sustain cell viability and enable reversible growth arrest, hallmarks of dormancy. MDA-MB-231 cells were cultured for 48h in the presence of 0, 10 and 50mM 2DG. The proportions of cellular YAP and phosphorylated YAP (pYAP) were determined by Western blot (**Figure 5.8A**). The analysed site of phosphorylation on YAP was Ser127, a highly conserved residue which is phosphorylated by LATS1/2 and its phosphorylation is strongly associated with cytoplasmic retention [324]. Both 10 and 50mM 2DG induced a reduction in YAP protein expression. Interestingly, however, the levels of pYAP increased upon addition of 10mM, but not 50 mM 2DG (**Figure 5.8B**).

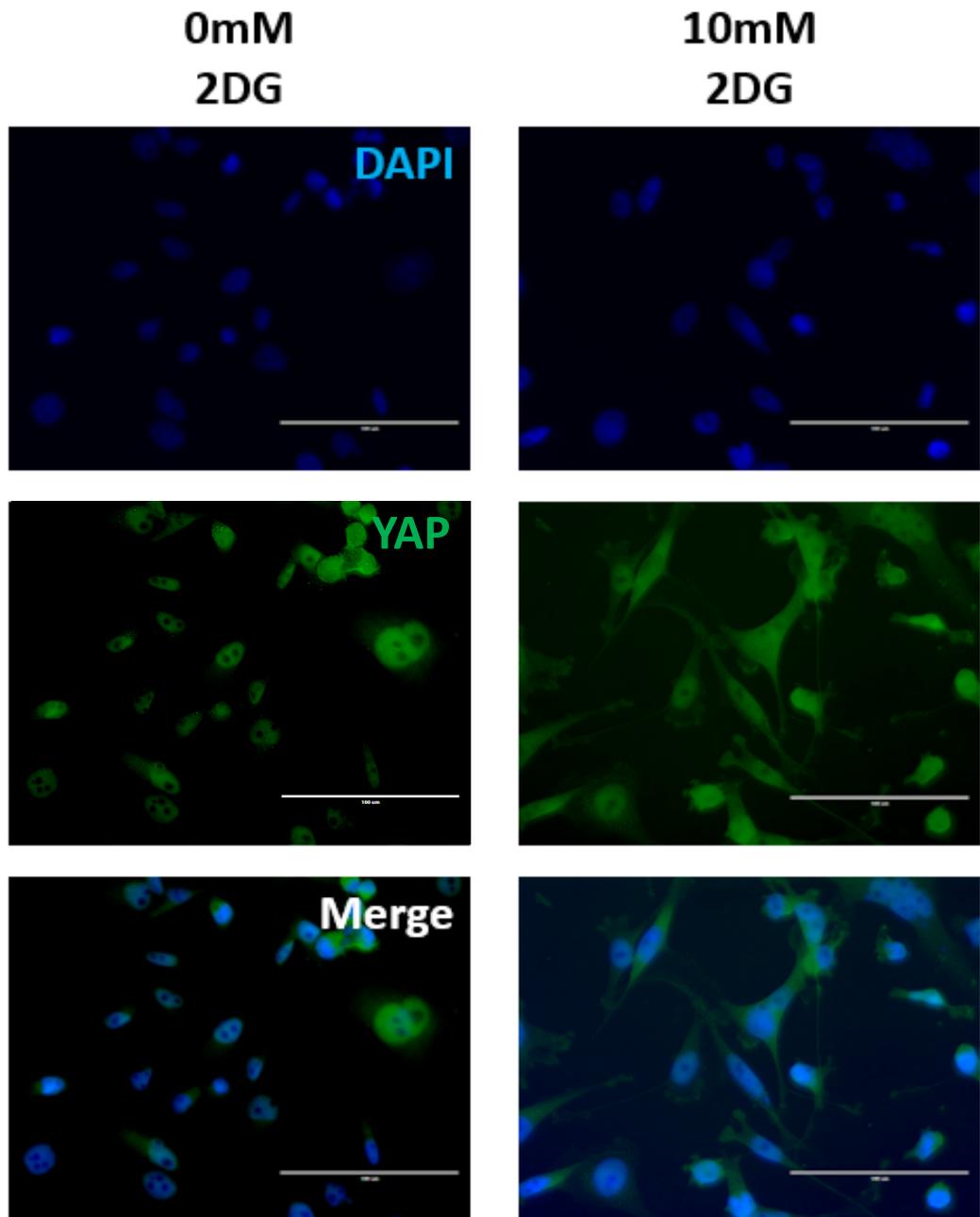
Taken together, this supports the theory that a particular level of glycolysis inhibition is required for dormancy, which may be related to YAP phosphorylation levels. At a higher concentration of 2DG, total YAP was reduced, yet without obvious changes in phosphorylation. It is not clear if higher concentrations of 2DG induces a stronger inhibition of glycolysis, or whether there are concentration-dependent side effects of 2DG, independent of glycolysis, that are responsible for cell death. Furthermore, given that YAP residue Ser127 is phosphorylated exclusively by LATS1/2 [324], these results

suggest that aerobic glycolysis is an activator of the HIPPO signalling pathway *in vitro*. Further study would be required to observe whether these effects can be replicated *in vitro*, as there may be alternative mechanisms which inhibit YAP phosphorylation, even when glycolysis is inhibited.



**Figure 5.8. Inhibition of aerobic glycolysis induces phosphorylation of the transcriptional activator, YAP, *in vitro*.** MDA-MB-231 cells were cultured in the presence of 0 (negative control), 10 and 50mM 2DG for 48h, prior to isolation of total protein. A) Cellular proportions of YAP and pYAP (Ser127) were determined by Western blot. The loading control employed was beta-actin, as shown. B) Cellular pYAP was normalised to YAP and relative abundance displayed for each concentration of 2DG. Statistical analysis performed was a one way ANOVA with multiple comparisons test. One representative experiment out of three is shown.

Given the increase in phosphorylated YAP in 2DG growth arrested MDA-MB-231 cells, I then sought to determine whether this is associated with a reduction in YAP localisation with the nucleus. MDA-MB-231 cells were cultured in the presence or absence of 10mM 2DG. Plastic substrate was selected over glass. The choice of substrate in modelling YAP localisation was critical. Previous reports observed that poly-L-lysine-coated glass strongly inhibits focal adhesion (FA) assembly of adipose-derived mesenchymal stem cells [325]. Where FAs could not form, cell area was significantly reduced, and YAP was excluded from the nucleus [325]. In optimisation experiments of YAP immunofluorescence staining in MDA-MB-231 cells, cells morphologically appeared rounder when cultured on poly-L-lysine-coated glass (data not shown). In addition, there was no discernible difference in YAP localisation between 2DG-treated and untreated cells. Resultantly, plastic substrate was selected which provided anchor points for cell spreading. At 48h post-addition of 2DG, cells were fixed and immunofluorescence staining for YAP localisation was performed (**Figure 5.9**). In the absence of 2DG, proliferating cells exhibited nuclear localisation of YAP. This is consistent with current data that YAP increases in nuclear abundance in cancer, activating transcription of proliferation-promoting factors [320]. In contrast, the addition of 10mM 2DG induced marked cytoplasmic retention of YAP, rendering it unavailable to partake in pro-metastatic nuclear transcriptional activity. Taken together, this suggests that breast cancer cell dormancy in the brain could be mediated via glycolysis-dependent inhibition of YAP nuclear translocation and should be investigated in future studies.



**Figure 5.9. Inhibition of aerobic glycolysis induces cytoplasmic retention of the transcriptional activator YAP.** MDA-MB-231 cells were cultured in the presence or absence (negative control) of 10mM 2DG for 48h, on plastic substrate *in vitro*. Immunofluorescence staining of YAP was performed, with DAPI nuclear counterstain. Images are representative of 3 independent experiments. Scale bar = 100  $\mu$ m.

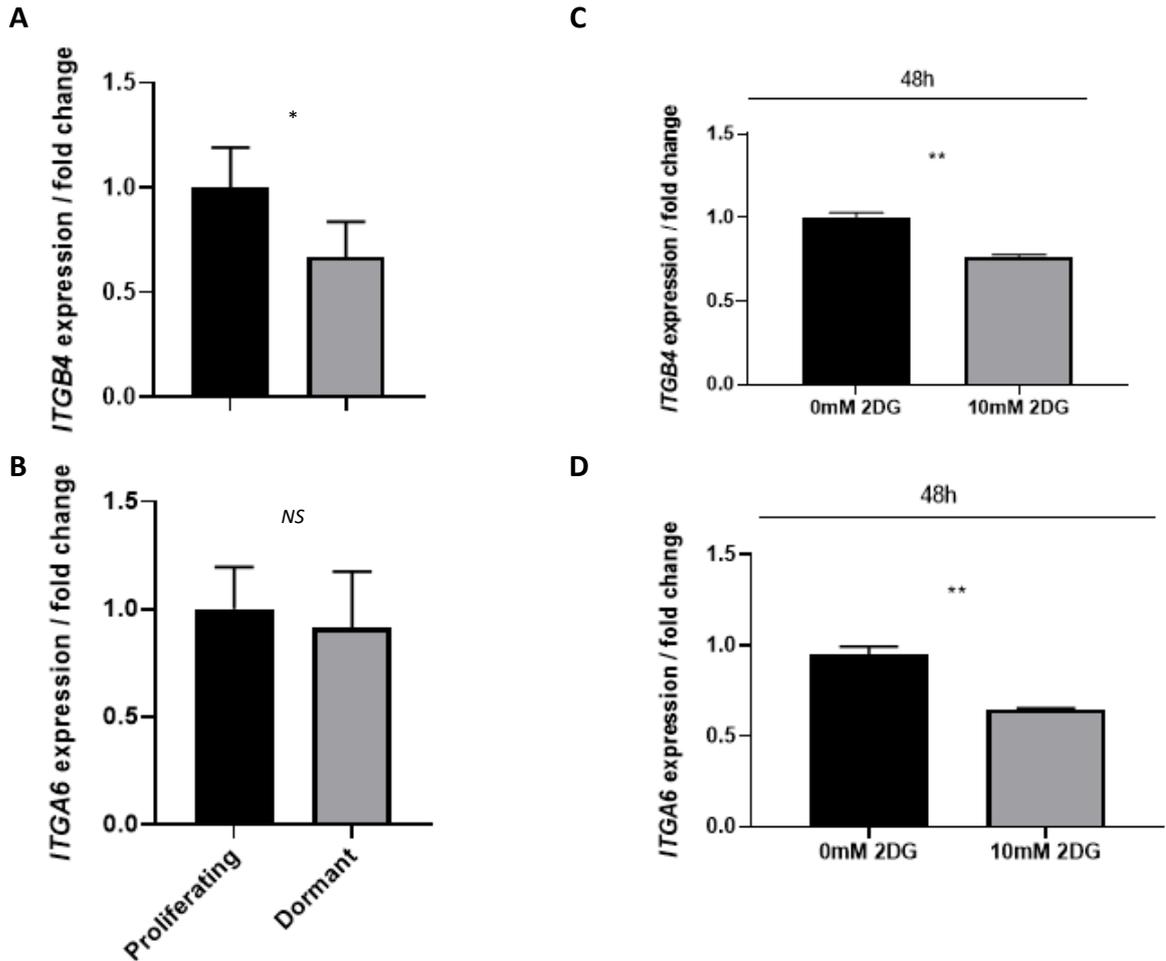
In a recent study, it was shown that cancer cell spreading on blood capillaries at disseminated sites is a critical step for metastatic colonisation [146]. Disseminated cancer cells displace pericytes and spread on capillaries via L1 cell adhesion molecule (L1CAM), which also activates YAP during the critical steps of metastatic outgrowth through an engagement of  $\beta$ 1 integrin [136]. In chapter 4, I have shown that biglycan-expressing cancer cells, presumed to be dormant, localise to brain vasculature, therefore *L1CAM* expression was of interest. MDA-MB-231 cells were cultured in the presence or absence of 10mM 2DG *in vitro* and qPCR was performed at 0, 24 and 48h (**Table 5.3**). No differential expression of *L1CAM* upon addition of 2DG could be detected at any time point. These results suggest that *L1CAM* expression is not affected downstream of 2DG-induced cytoplasmic retention of YAP *in vitro*. Subsequent analysis of transcriptomic profiling of dormant breast cancer cells isolated from the *in vivo* model, from both data processing streams, did not identify differential expression of *L1CAM* (data not shown). Therefore, the effects of *L1CAM* in the context of dormancy, if any, may not be dependent on its level of mRNA expression, and perhaps future work would benefit from studies involving its localisation and expression.

	<i>L1CAM</i>					
	0mM	10mM	0mM	10mM	0mM	10mM
	2DG (0h)	2DG (0h)	2DG (24h)	2DG (24h)	2DG (48h)	2DG (48h)
<b>Fold change (SE)</b>	0.99 (0.06)	1.03 (0.08)	1.01 (0.07)	1.01 (0.10)	2.31 (0.20)	1.70 (0.17)
<b>Significant difference?</b>	No		No		No	

**Table 5.3. Inhibition of aerobic glycolysis does not affect the expression of the YAP activator *L1CAM*.** MDA-MB-231 cells were cultured in the presence or absence (negative control) of 10mM 2DG for 0, 24 and 48h. Cells were harvest for 48h. Shown is qPCR analysis of *L1CAM* expression, relative to a housekeeping gene, *POLR2A*. N=3. Shown is mean fold change in expression +/- SEM. Statistical test performed was an unpaired two-tailed *t* test. Number of experimental repeats = 3.

It has previously been observed that YAP controls cellular biophysical properties through the formation of focal adhesions, which stabilise anchorage of the actin cytoskeleton to the cell membrane [325]. YAP is also involved in mediating cell interactions with the ECM through mediating the expression of various integrins [325]. Integrins are transmembrane receptors which consist of  $\alpha$  and  $\beta$  subunits, which combine as heterodimers [326]. Depletion of YAP not only induces alterations in expression patterns of a number of integrins, but may also modulate protein modifications, affecting cell stiffness, adhesion, migration and proliferation [325]. In dormant breast cancer cells *in vivo*, processing stream I, there was a significant downregulation in *ITGB4* at the gene expression level (**Figure 5.10A**). Furthermore, there was a tendency towards the downregulation of *ITGA6* expression, but this was not statistically significant ( $N = 3$ ,  $p = 0.0632$ ) (**Figure 5.10B**). Together, these genes encode subunits of the  $\alpha6\beta4$  integrin, which serves as the laminin receptor, binding laminin-332 and laminin-511 ligands [326]. The laminin receptor is involved in adhesion to the basement membrane and is known to be involved in the formation of micrometastases and cancer cell proliferation [327]. Upon addition of 10mM 2DG to MDA-MB-231 cells *in vitro*, there was a significant downregulation in *ITGB4* gene expression after 48h (**Figure 5.10C**). Similarly, *ITGA6* gene expression was also significantly downregulated (**Figure 5.10D**).

However, further study is required to address whether a downregulation of *ITGB4* and *ITGA6* at the gene expression level is biologically meaningful. I have previously shown that addition of 2DG to MDA-MB-231 cells at concentrations above 10mM induces detachment of viable cells from plastic substrate. It would first be useful to explore whether addition of 2DG induces a downregulation of laminin receptor subunit expression at the protein level and whether this effect is downstream of YAP inhibition, followed by *in vitro* functional studies of cellular adhesion. An important question to then ask, if a downregulation of aerobic glycolysis does interfere with laminin receptor adhesion, is whether this mechanism rapidly induces dormancy, following cancer cell extravasation to the brain. This could be explored using the same *in vivo* model of metastasis employed in this study, with the addition of 2DG treatment.

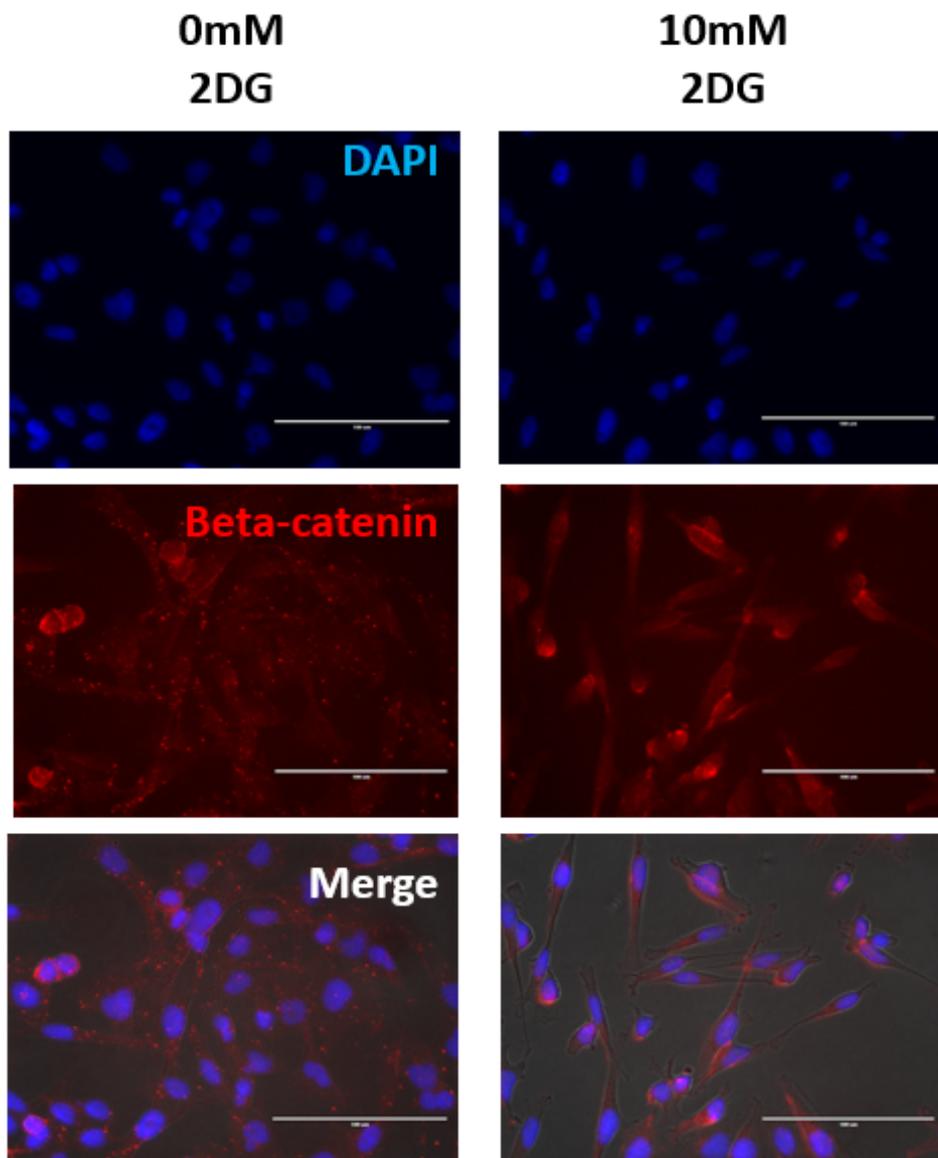


**Figure 5.10. Effect of inhibited aerobic glycolysis on the expression of laminin receptor subunits.** Expression of A) *ITGB4*, and, B) *ITGA6*, in dormant (GFP+CV+) and proliferating (GFP+) MDA-MB-231 cells isolated from the *in vivo* model. Fold changes are representative of three independent experiments and calculated downstream of size-factor normalisation by DESeq2. Quantitative PCR analysis of C) *ITGB4*, and, D) *ITGA6* expression in cultured MDA-MB-231 cells in the presence or absence (negative control) of 10mM 2DG, for 48h. Gene expression was normalised to that of a housekeeping gene, *POLR2A* +/- SEM. Statistical analysis performed was a two-tailed *t* test. One representative experiment out of three is shown. NS = not significant.

## 5.6 Inhibition of Aerobic Glycolysis does not Inhibit nor Promote Canonical Wnt Signalling in MDA-MB-231 Cells *In Vitro*

As a predicted enriched pathway in dormant MDA-MB-231 cells *in vivo*, Wnt signalling was probed in the context of aerobic glycolysis. As with YAP immunofluorescence staining, I probed  $\beta$ -catenin cellular localisation under identical conditions (**Figure 5.11**). In the absence of 2DG,  $\beta$ -catenin was distributed throughout the cytoplasm. Given the dotted staining pattern,  $\beta$ -catenin was likely localised to destruction complexes [328]. In the presence of 10mM 2DG, the  $\beta$ -catenin staining pattern was more diffuse throughout the cytoplasm, with an abrogation of a dotted staining pattern. Some  $\beta$ -catenin potentially localized to the plasma membrane; however, this could not be verified in the absence of a plasma membrane counterstain. Despite the result, it is not appropriate to draw any conclusions from this data. It is possible that this *in vitro* model of 2DG addition is not representative of what happens to tumour cells in the brain, when glycolysis is inhibited.

To further investigate the potential activity of  $\beta$ -catenin, cells were harvested for RNA isolation, at 0, 24 and 48h after addition of 2DG. *NKD1* and *AXIN2*, two genes that are exclusively expressed downstream of active Wnt signalling, were observed to be upregulated in dormant cells *in vivo*. Therefore, I performed qPCR analysis to determine their expression *in vitro* (**Table 5.4**). At the 24 and 48h time points, addition of 10mM 2DG did not alter expression of either gene. Therefore, one can conclude that aerobic glycolysis inhibition via the addition of 2DG does not influence the expression of *NKD1* and *AXIN2* in MDA-MB-231 cells, but only in an *in vitro* context. Further study is required to determine whether their expression is altered when cancer cell glycolysis is inhibited within the brain.



**Figure 5.11. Inhibition of aerobic glycolysis does not induce nuclear localisation of  $\beta$ -catenin *in vitro*.** MDA-MB-231 cells were cultured in the presence or absence (negative control) of 10mM 2DG for 48h, on plastic substrate *in vitro*. Immunofluorescence staining of beta-catenin was performed, with DAPI nuclear counterstain. Images are representative of 3 independent experiments. Scale bar = 100  $\mu$ m.

	<i>AXIN2</i>					
	0mM	10mM	0mM	10mM	0mM	10mM
	2DG (0h)	2DG (0h)	2DG (24h)	2DG (24h)	2DG (48h)	2DG (48h)
<b>Fold change (SE)</b>	0.99 (0.09)	1.19 (0.10)	1.28 (0.11)	1.13 (0.12)	2.24 (0.21)	2.23 (0.22)
<b>Significant difference?</b>	No		No		No	

**Table 5.4. Inhibition of aerobic glycolysis does not affect the expression of the Wnt obligate genes, *AXIN2* and *NKD1*, *in vitro*.** MDA-MB-231 cells were cultured in the presence or absence (negative control) of 10mM 2DG for 0, 24 and 48h. Cells were harvested at the indicated time points and were analysed by qPCR for the expression of *AXIN2* and *NKD1*. Expression of *NKD1* could not be calculated due to an absence of Ct value for all 2DG concentrations and time points. Shown is the expression of *AXIN2*, normalised to a housekeeping gene, *POLR2A*. N=3. Statistical analysis performed was a one-way ANOVA with multiple comparisons test. One representative experiment out of three is shown.

### 5.7 Proposed Schematic of Dormancy Induction Facilitated by a Downregulation of Aerobic Glycolysis

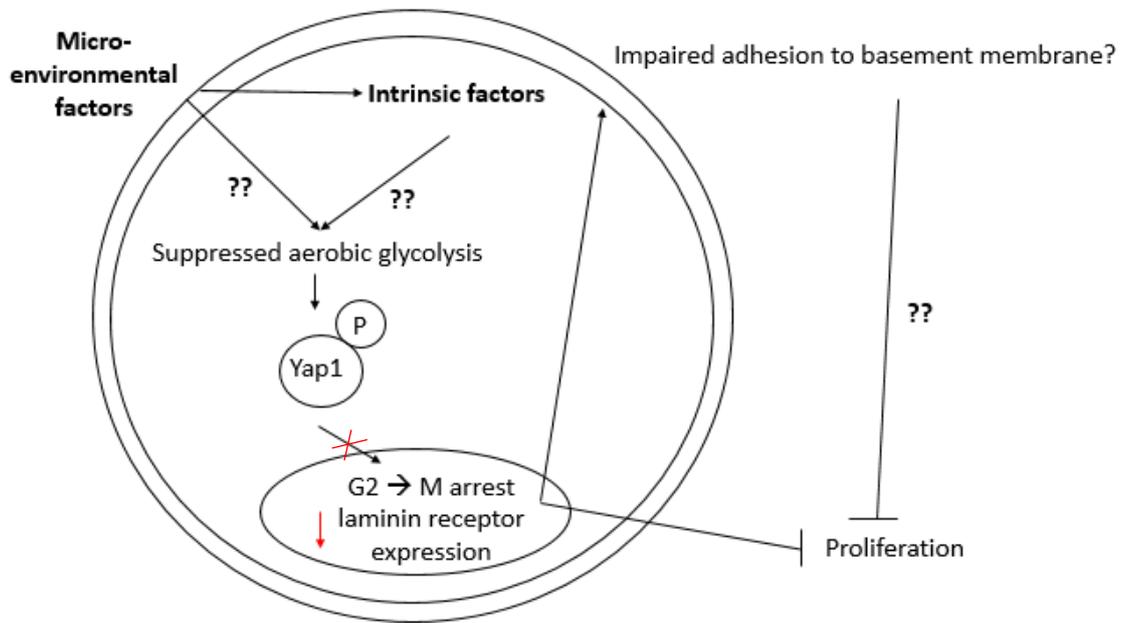
Based on a comparison of dormant cells *in vivo* and MDA-MB-231 cells *in vitro* following addition of 10mM 2DG (**Table 5.5**) I devised a schematic of the proposed pathway in which inhibited aerobic glycolysis induces growth arrest in MDA-MB-231 cells (**Figure 5.12**). Based on transcription factor analysis in Chapter 3, a downregulation of glycolysis genes may be modulated via suppression of HIF-1 and Myc, with an upregulation of p53. Suppression of aerobic glycolysis could then induce the phosphorylation and cytoplasmic retention of YAP, which blocks its ability to undergo nuclear shuttling and blocks proliferation-promoting transcription. One of the effects of aerobic glycolysis inhibition is a mid-cycle arrest in G2/M phase. Further work would be required to determine the involvement of YAP.

	<b>Dormant cells <i>In vivo</i></b>	<b>MDA-MB-231 cells <i>In vitro</i> (after addition of 10mM 2DG)</b>
<b>Glycolysis genes</b>	Significant downregulation of <i>ALDOA</i> , <i>GAPDH</i> , <i>PGK1</i> , <i>PGAM1</i> , <i>ENO1</i> and <i>LDHA</i> at the gene expression level	No change in gene expression to <i>PGK1</i> or <i>GAPDH</i>
<b>Biglycan gene expression</b>	Significantly upregulated at the gene expression level	No change in gene expression
<b>YAP status</b>	No evidence	Cytoplasmic localisation / phosphorylated
<b><i>L1CAM</i> expression</b>	Significantly upregulated at the gene expression level	No change in gene expression
<b>Wnt signalling pathway</b>	Enriched	Cannot determine $\beta$ -catenin cellular localisation No change to <i>AXIN2</i> expression <i>NKD2</i> not expressed in both dormant and proliferating cell populations
<b>Cell size and morphology</b>	Cannot determine	Normal size and morphology
<b>Extracellular pH</b>	Cannot determine	Extracellular pH does not reduce
<b>Cell cycle</b>	Predicted reduced G1 score, with increased G2/M and S phase scores	G2/M cell cycle arrest
<b>Integrin gene expression</b>	<i>ITGB4</i> significantly downregulation Decreased mean expression of <i>ITGA6</i> but not significant	Significant downregulation of <i>ITGB4</i> and <i>ITGA6</i>

**Table 5.5. A comparison of dormant cancer cells *in vivo* and MDA-MB-231 cells *in vitro* (after addition of 10mM 2DG)**

Furthermore, downregulation of *ITGB4* and *ITGA6* genes suggests that expression of laminin receptor may be reduced. A follow-up investigation should determine whether these observations are translatable *in vivo*, for example whether impairment of adhesion of cancer cells to the basement membrane, via YAP phosphorylation, can provide an additive effect in the inhibition of metastatic outgrowth. However, there

is currently little agreement between the two model systems, and my findings *in vitro* aren't immediately translatable *in vivo*. Going forward, my observations in chapter 5 will need to be replicated using suitable *in vivo* models, and it should be determined whether there may be a link to the overexpression of biglycan in dormant cells.



**Figure 5.12. Schematic of proposed dormancy mechanisms in breast cancer cells.** My *in vitro* studies identified that inhibition of aerobic glycolysis is associated with phosphorylation and cytoplasmic retention of YAP. In turn, this leads to G2/M phase arrest and reduced laminin receptor gene expression. Further study is required to determine whether this is comparable to dormant breast cancer cells *in vivo*, and whether reduced laminin receptor expression presents with any functional effects. In addition, it will be important to determine the mechanisms of aerobic glycolysis suppression, as this may reveal novel therapeutic targets.

## 5.8 Discussion

In this chapter, aerobic glycolysis was investigated as a potential driver of breast cancer cell dormancy. Inhibition of glycolysis using the glycolysis inhibitor, 2DG, induced growth arrest of MDA-MB-231 cells *in vitro*. Titration studies revealed that inhibition of aerobic glycolysis with 2DG induces growth arrest over the 9 days in which the drug was applied, whilst maintaining cellular viability. In addition, analysis of dormant (GFP+CV+) cell mRNAseq data, in Chapter 3, identified a reduction in G1 scores, with increases in S and G2/M phase scores. Analysis of 2DG-arrested MDA-MB-231 cells identified an arrest in the G2/M phase, corroborating the *in vivo* data. Pertinent to the dormancy phenomenon, growth arrest was reversible, restoring normal cancer cell proliferation upon withdrawal of 2DG. Interestingly, growth arrest was accompanied by increased phosphorylation and cytoplasmic retention of YAP, with reduced expression of genes controlling expression of laminin receptor subunits. As described in Chapter 3, YAP is the core transcriptional regulator of the HIPPO signalling pathway, which was predicted to be enriched in dormant cells *in vivo*. This suggests that aerobic glycolysis may be involved in dormancy induction and maintenance, either through direct growth inhibition, or its interaction with other intracellular pathways, such as the HIPPO pathway. However, the data presented in this study is not directly translatable to what is happening to dormant cells *in vivo* and should be viewed openly.

### 5.8.1 Relevance of 2DG in the Treatment of Triple Negative Breast Cancer

The use of 2DG as a chemotherapeutic agent has been explored in combination with docetaxel, in a phase I trial involving patients with solid tumours [329]. With the determined clinically tolerable dose of 2DG being 63 mg/kg, patients given 63 – 88 mg/kg all present with reversible hyperglycaemia, 6% with gastrointestinal bleeding and 22% with reversible grade 3 QTc prolongation [329]. Whilst these are tolerable symptoms, 66% of patients still present with progressive disease, 32% with stable disease and 3% with a partial response [329]. Whilst 10mM 2DG induced reversible growth arrest of triple negative MDA-MB-231 cells *in vitro* in my study, this may not be translatable in clinic if the intention was to induce a state of dormancy in established lesions. The average total blood volume for an adult male is 82.3 +/- 2.0 ml/kg [330]. Assuming the molecular weight of 2DG to be 164.16, and an average total blood volume

of 82.3 ml/kg, then reaching 10mM 2DG would require a dose of approximately 135 mg/kg, assuming 100% bioavailability. This is well above the threshold of tolerable side-effects. Therefore, it would be impractical to administer such doses without adverse side effects. It may be easier to sustain dormancy indefinitely at tolerable doses of 2DG, given that dormant cancer cells are already metabolically compromised. Triple negative and HER2+ breast cancers are particularly reliant on glycolysis, whilst ER+ breast cancers retain more functionality in oxidative phosphorylation [331]. As such, glycolysis inhibitors may display variable functionality between subtypes. A better approach may be to target downstream effectors of glycolysis.

### 5.8.2 Inhibited Glycolysis on Cell Survival, Growth Arrest and YAP

*In vitro*, I observed that MDA-MB-231 cells display a reduction in viability at 2DG concentrations higher than 10mM. However, at 10mM 2DG, G2/M cell cycle arrest was accompanied by sustained survival, for at least 9 days. Although unexplored in this study, further work should analyse the mRNAseq data presented in Chapter 3 for indicators of survival mechanisms, and whether those mechanisms are present upon addition of 2DG *in vitro*. For example, it has previously been shown that AMP-activated protein kinase (AMPK), an evolutionarily conserved energy sensor, is activated under glucose-limiting conditions which maintains cellular survival [332]. The addition of 2DG to MCF7 breast cancer cells induces short-term decreases in intracellular ATP, which has shown to be recovered by AMPK and abrogates 2DG-induced cytotoxicity [333]. Addition of AMPK antagonists in combination with CREB pathway inhibitors, depletes intracellular ATP and leads to enhanced cytotoxicity in cancer cells [332]. AMPK inhibitors do not induce the same cytotoxicity in non-cancerous MCF-10A cells [333], therefore if this, or a similar mechanism, exists in dormant MDA-MB-231 cells, then this may allow for targeted dormant cell cytotoxicity, given prior to relapse.

I also reported a significant downregulation of *ITGB4* in dormant cells, as identified by mRNAseq. Administration of 10mM 2DG to MDA-MB-231 cells *in vivo* induced a significant downregulation of both *ITGB4* and *ITGA6*. This was accompanied by phosphorylation and cytoplasmic retention of YAP, which is in line with a recent study which showed that YAP is involved in the activation of genes encoding integrins [416]. *ITGB4* and *ITGA6* heterodimerise to form the  $\alpha6\beta4$  integrin, functioning as a receptor

for laminins 332 and 511 [326]. Binding to laminin 332 and 511 allows cancer cells to adhere to the vascular basement membrane, a critical event in the establishment and outgrowth of brain metastases [334]. This may be important as I have observed a biglycan-expressing cancer cell adjacent to vascular endothelial cells in my *in vitro* model. Additionally,  $\alpha 6\beta 4$  has a wide range of downstream signalling targets. Downstream pathways include activation of the Ras-MEK-Erk and PI3K pathways, and Akt, Rac and mTOR targets [335]. This mechanism has been demonstrated in other models, for example a steroidal-based drug, curcubitacin B (CuB), has been shown to induce G2/M cell cycle arrest and inhibit the janus kinase (JAK)-signal transduction and activator of transcription (STAT) pathway [334]. Mechanistically, CuB induces downregulation of *ITGB4* and *ITGA6*, thereby suppressing integrin-mediated signalling [336]. Although speculation, there is enough evidence to potentially implicate YAP involvement in growth arrest, upon inhibition of aerobic glycolysis. It will be important for future work to determine whether this work translates to protein expression of these integrins, and whether there are any functional effects, involving YAP or otherwise pertaining to dormancy.

YAP has previously been associated with the development of brain metastases. In Chapter 4, I discussed my findings that biglycan-positive cancer cells localise along brain vascular endothelial cells. It has been shown in another study that a critical step for metastatic outgrowth is the spreading of cancer cells on resident vasculature, following which, L1CAM signals through YAP in the initial development of micrometastases [136]. Additionally, as a positive regulator of neovascularisation, VEGF-A has been implicated in tumour growth and metastasis, whose activity is positively regulated by YAP [337]. It is possible for solitary cells and micrometastases to persist without inducing neovascularisation. This involves the migration of cancer cells to pre-existing blood vessels of the host organ, in a process known as vessel co-option [338]. It is still unclear, however, as to whether dormancy is a phenomenon exclusive to solitary cells, or whether dormancy can occur in clusters of cells, such as micrometastases. One should take caution in the interpretation of these observations. Dormancy release and outgrowth of micrometastases, whether YAP is implicated or not, could be mediated by two distinct mechanisms.

As explored in chapter 4, upon biglycan over-expression *in vitro* at very high levels, I observed a downregulation of *GAPDH* in MDA-MB-231 cells. This preliminary data may point towards the origin of downregulated aerobic glycolysis. In p53-deficient cancer cells, two temporally distinct G2/M checkpoint networks exist: checkpoint kinase 1 (Chk1) mediates an early nuclear checkpoint and MAP kinase-activated protein kinase 2 (MK2) mediates a cytoplasmic checkpoint [339]. In response to DNA damage, where p53 is lost, as in breast cancer, the p38/MK2 complex undergoes nuclear to cytoplasmic translocation. This stabilises growth arrest and DNA damage inducible alpha (*Gadd45 $\alpha$* ) mRNA, blocking mitotic entry [339]. In dormant breast cancer cells *in vivo* in this study, downregulation of genes pertaining to aerobic glycolysis was associated with repression of the transcription factor MYC, and an indication of overexpression of *TP53* (significant by *p* value). This becomes relevant in the context of p38-mediated activation of TIAR, which is a potent inhibitor of MYC translation [340]. Furthermore, cells deficient in p38 activity also upregulate glycolysis, as p38 is known to inhibit the activity of HIF-1A transcription factor [341, 342]. HIF-1A is a critical transcription factor involved in the modulation of the Warburg effect, whose activity was predicted to be repressed in dormant breast cancer cells *in vivo* in this study. Therefore, future study should identify whether biglycan overexpression could induce G2/M arrest and the inhibition of aerobic glycolysis, of which p38 involvement should be explored.

There are also some limited studies documenting YAP in the context of G2/M-phase transition in cancer. YAP is phosphorylated at T119, S289 and S361 by CDK1, a process which is critical for neoplastic transformation [344]. In a recent study, a number of cancer cells were treated with a synthetic cyclizing-berberine (berberine of 1,13-cyclization), A35. This induced YAP phosphorylation (Ser127) and its cytoplasmic retention and was responsible for G2/M-phase arrest, independent of p53 [345]. It would therefore be useful to determine whether G2/M-phase arrest in my 2DG model is facilitated by YAP, given that I observed its phosphorylation and cytoplasmic retention. It would be important in this case to conduct protein expression and functional studies, for example to determine complexes involving YAP, potentially coupled with spatiotemporal localisation and expression of G2/M phase checkpoint proteins. However, it would be difficult at this stage to perform protein expression

studies without understanding if the mRNAseq data presented in Chapter 3 translated to protein expression. It may be worth the initial development of an *in vitro* co-culture model, for example using astrocytes, endothelial cells and MDA-MB-231 cells, in an attempt to replicate my *in vivo* data. This could be validated by gene expression. Subsequent proteomic analysis may then be possible if the yield of dormant cells is greater than what I achieved *in vivo*.

### 5.8.3 Decreased Acidity of the Extracellular Milieu and Growth Arrest

I observed that 2DG-treated, growth arrested MDA-MB-231 cell cultures don't present with a colour change of the growth medium over time, while medium of control cells changes towards yellow due to pH indicator phenol red contained in medium. Thus, reduced lactate and proton efflux, mediated by 2DG, correlates with a sustained alkaline microenvironment of growth arrested 231 cells. In a previous study, treatment of MCF7 breast cancer cells with 10mM 2DG *in vitro*, induced an increase in the pH of the growth medium as compared to untreated cells to 7 [331]. This correlated with an uptake of extracellular lactate by the cancer cells, actively sustaining a high pH [331]. As a pH gradient is crucial for metastatic outgrowth, it would be worth exploring whether an inability to generate an optimal pH gradient within the brain is supportive of a dormant phenotype. To support this theory, it has previously been shown that prior to G2/M entry, there is a transient intracellular increase in sodium-hydrogen exchanger 1 (NHE1/*SLC9A1*) activity and an increase in intracellular pH [346]. In breast cancer, intracellular pH is 7.1 in early S phase, 6.8 in G2, followed by a recovery in M phase, which is dependent on NHE1 [347]. If this can be established, this potentially opens up a novel area of therapeutics for breast cancer, ones which exploit the reduced metabolic state of dormant cells. Ultimately, it is still unknown as to whether 10mM 2DG, the optimum concentration to induce growth arrest in MDA-MB-231 cells, completely or partially inhibits aerobic glycolysis. A lack of acidification of the culture medium suggests that glycolysis may have been completely inhibited, therefore this requires downstream investigation. One such experiment may be to analyse the oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) of MDA-MB-231 cells by XF Extracellular Flux analyser [459], as they respond to various concentrations of 2DG. This approach has previously been used to determine the bioenergetics of two glioblastoma cell lines,

and whether they utilise aerobic glycolysis or oxidative phosphorylation in response to various substrates and metabolic perturbation agents [348]. This technique would allow quantification of the rate of glycolysis, and whether it is completely or partially inhibited upon addition of 10mM 2DG.

#### 5.8.4 Summary

The results of this chapter demonstrate that 2DG-mediated inhibition of aerobic glycolysis *in vitro* can induce and sustain reversible G2/M phase arrest in breast cancer cells. However, G2/M arrest would challenge the widely believed view that dormancy is characterised by chronic quiescence in G0 phase. Furthermore, YAP was phosphorylated downstream of inhibited glycolysis, resulting in its cytoplasmic retention. Downstream targets of YAP include genes which control expression of laminin receptor subunits, of which I observed *ITGB4* and *ITGA6* to be downregulated upon addition of 2DG *in vitro*. In Chapter 3, I determined that *ITGB4* was significantly downregulated in dormant cells *in vivo*. Further work is required to determine whether an inhibition of glycolysis in breast cancer cells can facilitate detachment from the ECM and induce growth arrest. Although these findings are important first steps, the work carried out in this chapter must be explored in greater depth using suitable *in vitro* and *in vivo* models. Additionally, it will be essential to determine whether the changes I observed at the gene expression level translate to changes in protein synthesis and functional effects. Finally, my work on biglycan in Chapter 4 can only be loosely tied to my work in this Chapter. To start to develop a possible dormancy mechanism, initial studies should focus on the overexpression of biglycan in MDA-MB-231 cells, and whether this results in the same observations made upon addition of 2DG, before commencing *in vivo* studies.

## Chapter 6

### Conclusions and Future Work

Of the approximately 2.1 million new breast cancer cases each year, the percentage of patients who develop brain metastases is 15-30% [1, 2, 20]. Surgery and current systemic therapies offer little to no benefit to survival rates post-diagnosis of brain metastases [30]. Median survival times which for patients with breast cancer brain metastases, from detection, ranges from 3.7 to 15 months, representing a survival range of 0.2 to 56 months [21]. Cancer dormancy is a phenomenon in which disseminated tumour cells persist in a clinically occult state of quiescence for extended periods of time [42, 43, 44]. Cancer cells may be able to escape from dormancy under more favourable conditions, leading to metastatic progression [42, 43, 44]. A deeper understanding of the mechanisms underpinning the induction, and maintenance of, dormancy may offer novel targets for the treatment of metastatic breast cancer.

#### 6.1 Chapter 3

##### 6.1.1 Conclusions

I have concluded from the data presented in Chapter 3 that the *in vivo* brain tumour xenograft model employed is viable in the isolation of dormant and proliferating breast adenocarcinoma cells. The advantage of this model, compared to others, is that I could isolate dormant and proliferating cells based upon the respective retention and loss of a vital dye *in vivo*. Given my observations that I observed a significantly far lower proportion of cancer cells positive for dye retention after 4 weeks *in vitro*, I could conclude that the cells identified as dormant *in vivo* were growth arrested due to being within the microenvironment of the brain. Dormant cells *in vivo* represent 0.5-2% of the total cancer cell population in the brain, based on my observations. Furthermore, I concluded that dormant and proliferating cancer cells display distinct transcriptomes *in vivo*, further demonstrating that dormancy in the model employed was a reversible phenomenon. Transcriptomic analysis identified a number of potentially implicated

genes and pathways, although few firm conclusions can be made based on this data alone. In dormant cells, the HIPPO, Wnt and TGF- $\beta$  signalling pathways were enriched. As the most significantly differentially expressed gene, biglycan was chosen as a target for downstream analysis. I could also conclude that dormant cells downregulate a number of genes encoding enzymes within aerobic glycolysis, in line with my prediction, based on a comparison to publicly available datasets, that the glycolysis-associated transcription factors MYC and HIF-1A, were repressed.

### 6.1.2 Critical Analysis

The *in vivo* model employed was limited by the low output of dormant cells for every mouse injected. Not only did this provide technical limitations with respect to yield, there was a higher risk of contamination from mouse cells, which I had to compensate for during mRNAseq data processing. The negative effects of this were not clear, despite attempts to mitigate contamination, however this could have obscured genes with small but significant effects, and downstream analyses could have still incorporated reads of mouse origin. Additionally, the low yield of dormant cells resulted in little material for cDNA synthesis, restricting the use of qPCR in validation studies. PCR pre-amplification of template DNA may overcome this problem, however increasing the number of cycles may also result in unequal amplification of DNA fragments, potentially distorting the data [349]. Furthermore, the low cell yield restricted the use of many molecular biology techniques, for example, traditional Western blotting. Dormant cells also cannot be propagated post-isolation, as they would no longer be dormant and therefore would lose the phenotype of interest. Finally, the low yield of dormant cells means that analysis of dormant cells by immunofluorescence within the microenvironment of the brain is difficult. With approximately 14 to 75 dormant cells per mouse brain, observing them *in vivo*, unaided by automated technology, is currently difficult.

### 6.1.3 Future Work

With respect to my gene expression data of dormant and proliferating cancer cells, future work should focus on a deeper exploration of differentially expressed genes, and

identified factors and pathways not explored in this study. For example, Sp1 transcription factor was predicted to be upregulated in dormant cells *in vivo*, which recognises a binding site on the biglycan promotor [350]. As yet to be explored, my list of differentially expressed genes could be used to predict microRNA (miRNA) involvement [351]. In glioblastoma, miRNAs are regulators of aerobic glycolysis, and worth investigating in the context of breast cancer dormancy in the brain [352]. Furthermore, much of the work in this study relies heavily on changes to gene expression. In recent years, the development of single-cell Western blotting may overcome antibody fidelity and sensitivity limitations, allowing for analysis of isolated dormant cells at the protein expression level [353]. This would allow insight into the phosphorylation status and relative protein expression of a number of proteins described as being implicated in the induction of dormancy in this study.

## **6.2 Chapter 4**

### **6.2.1 Conclusions**

No overarching conclusions can be made from the data presented in Chapter 4; however, some assumptions can be made based upon available data, which should inform future study. Based on publicly available datasets, I observed that high biglycan expression in primary breast cancers is correlated with increased potential of metastasis and relapse. Conversely, biglycan expression is reduced in all sites of metastases analysed. However, no conclusions can be made as to why these observations were made, nor potential mechanisms. I can conclude that overexpression of biglycan at very high levels in MDA-MB-231 cells *in vitro* can induce growth arrest. Preliminary data has shown a downregulation of *GAPDH* in response; however, no conclusions can be made regarding aerobic glycolysis as an implicated pathway in dormancy. Finally, given very few observable events of biglycan-expressing cancer cells in the brain, it would be premature to formulate any conclusions. However, preliminary data implicates vascular endothelial cells in dormancy induction and maintenance, given their localisation to the single dormant cell observed.

### 6.2.2 Critical Analysis

One of the main limitations in Chapter 4 was the scarcity of dormant cells in mouse brains. This made it exceptionally challenging to validate my mRNAseq data, evidenced by the observation of just 2 biglycan-expressing cancer cells, which may or may not be dormant. Interestingly, however, was the link between very high biglycan overexpression in MDA-MB-231 cells *in vitro*, with *GAPDH* downregulation and growth arrest, in my preliminary studies. Whilst this is yet to be explored further, this starts to suggest a potential link between SLRP/biglycan expression and glycolysis. To my knowledge, there is currently no literature which describes the effects of extracellular, or intracellular, biglycan or any proteoglycan on aerobic glycolysis in cancer, therefore this would be a novel biological finding. A significant problem I have encountered as a whole, however, is that *in vitro* validation experiments were not representative of dormant cells *in vivo*. There may be great deal of divergence of the effects of biglycan when overexpressed in cancer cells within the brain, therefore one must be careful with regards to drawing any major conclusions regarding dormancy from the work in this chapter.

### 6.2.3 Future Work

Future work should address the problem that biglycan expression positively correlated with a reduction in distant metastasis- and relapse-free survival in primary tumours, yet compared to primary tumours, biglycan expression was reduced at metastatic sites. It would be useful to develop *in vivo* biglycan overexpression models. One such approach could be the development of a spontaneous model of metastasis, for example injection of cancer cells into the mammary fat pad, then propagation of cancer cells within the fat pad, followed by removal of the primary tumour, then continuous monitoring for metastasis [354]. A complication is that spontaneous breast cancer metastasis models rarely develop brain metastases [355]. It may be appropriate to investigate more common sites of spontaneous metastasis. Spontaneous models of metastasis in the context of dormancy may involve the injection of MDA-MB-231 cells, and/or cell lines of

alternative subtypes, with constitutively active biglycan, or its inducible expression. A useful addition to this would be super-resolution microscopy (SRM) which bypasses diffraction limits and vastly improves optical resolution [356]. An example for its use would be to directly image the potential binding of biglycan to its predicted binding partners, or to observe whether its excretion from dormant cells induces cross-linking and/or alternative modification of the basement membrane.

## 6.3 Chapter 5

### 6.3.1 Conclusions

Based on my observations in Chapter 5, I have concluded that an inhibition of aerobic glycolysis *in vitro* is sufficient to induce reversible growth arrest of MDA-MB-231 cells, in line with an expected dormancy phenotype. Using the same model, I have concluded that growth arrest, mediated by 2DG, arrests breast cancer cells at the G2/M phase checkpoint. This is accompanied by phosphorylation and cytoplasmic retention of the core HIPPO transcriptional regulator, YAP. However, I cannot determine at this stage whether YAP is involved in G2/M arrest. Concurrent with my observations of YAP, inhibition of aerobic glycolysis *in vitro* downregulates laminin receptor subunits at the gene expression level, *ITGB4* and *ITGA6*, corroborating the gene expression data in dormant cells, as shown in Chapter 3. However, few conclusions can be made as to the effects of downregulated aerobic glycolysis in dormant cells *in vivo*, if any. Inhibition of glycolysis by 2DG *in vitro* does not modulate biglycan, glycolysis enzymes or *L1CAM* expression, and has no observable effects on Wnt signalling, according to  $\beta$ -catenin localisation.

### 6.3.2 Critical Analysis

One of the largest pitfalls to this approach is that *in vitro* studies of glycolysis inhibition is not representative of glycolysis enzyme downregulation in dormant cells *in vivo*. However, it is very promising that long-term survivable growth-arrest could be achieved, for as long as 2DG was administered. Following its removal, cells escaped growth arrest

and resumed proliferation, and mirrors the phenotype of the dormancy phenomenon I was expecting to observe. Recent studies are reporting that dormancy is characterised by G0/G1 phase arrest and a low metabolic state [162]. A low metabolic state does corroborate my data; however, it remains to be shown whether this is an effect or driver of the dormancy phenomenon. To my knowledge, there is currently no literature which functionally links low metabolic state to the induction of dormancy. The expectation of G0/G1 arrest remains a significant problem. In Chapter 3, my analyses directly conflict this, showing a reduction in G1 score in dormant cells *in vivo*, whilst I observed G2/M arrest *in vitro* upon addition of 2DG. This strong challenge against the wider literature certainly warrants further scrutiny.

### 6.3.3 Future Work

As stated, future work should focus on confirming whether what we have observed *in vitro* translates to dormant cells *in vivo*. One such approach may be the inducible knockdown of glycolysis-associated enzymes in my brain tumour xenograft model, avoiding potentially adverse systemic effects of administering 2DG. I also would not be able to reproducibly expose cancer cells in the brain to the same dose of 2DG. Furthermore, given the observations that aerobic glycolysis induces G2/M arrest, G2/M-phase specific markers, such as cyclin B1 and CDK1 [357], could be used in a number of techniques, for example flow cytometry, to pinpoint the precise cell cycle phase/checkpoint of isolated dormant and proliferating cells. Due to the low yield of dormant cells in my model, cell cycle FACS would not be appropriate with the current *in vivo* model, as this requires a minimum of 10,000 cells, for each gated phase [358]. However, this could and should be combined with a validated biglycan-over-expression model. Inhibition of aerobic glycolysis *in vitro* did not affect biglycan expression, however it remains to be observed whether biglycan overexpression can downregulate aerobic glycolysis, and what effects this has both *in vivo* and *in vitro*. Furthermore, determining the expression of specific checkpoint markers would allow for the identification of the exact point of arrest, offering an advantage over cell cycle FACS which broadly distinguishes cells according to major phase: G0, G1, S and G2/M [359]. I

would then be able to investigate the potential survival mechanisms and mechanisms associated with cell cycle arrest in single dormant breast cancer cells.

#### **6.4 Final Conclusion**

In conclusion, a downregulation of aerobic glycolysis at the gene expression level may be required to induce and sustain a state of growth-arrest in breast cancer cells, according to my *in vitro* data. Further study is warranted to explore whether downregulated aerobic glycolysis is translatable *in vivo*, as a driver dormancy, rather than an effect of the phenomenon. In conjunction, dormant breast cancer cells in the brain were associated with high levels of biglycan expression. Immunofluorescence validation of mRNAseq data using sections of brain tumour xenografts provided some data to suggest that solitary biglycan-expressing breast cancer cells localise with brain microvasculature. Finally, YAP, a major component of the HIPPO signalling pathway, was shown to be phosphorylated with cytoplasmic localisation in 2DG-mediated growth-arrested breast cancer cells *in vitro*. These events were associated with G2/M phase cell cycle arrest, in contrast to existing literature which states that dormancy is likely to be in G0 arrest. These results provide a new insight into potential mechanisms underpinning breast cancer dormancy in the brain. Future study should attempt to link biglycan overexpression to inhibited aerobic glycolysis, and whether this reflects the mechanisms of dormancy induction and maintenance *in vivo*.

## Appendix

### 7.1 List of Suppliers

Abcam	Discovery Drive, Cambridge Biomedical Campus, Cambridge, CB2 0AX
Acros Organics (Thermo Fisher)	168 Third Avenue, Waltham, MA, USA, 02451
Agilent	5301 Stevens Creek Blvd, Santa Clara, CA, 95051, USA
Alfa Aesar	Shore Road, Port of Heysham Industrial Park, Heysham, LA3 2XY, UK
American Type Culture Collection (ATCC)	LGC Standards, Queen's Road, Teddington, Middlesex, TW11 0LY, UK
Anaspec	34801 Campus Drive, Fremont, CA, 94555, USA
Applied Biosystems	168 Third Avenue, Waltham, MA, USA, 02451
BD Biosciences	Edmund Halley Road, Oxford Science Park, Oxford, OX4 4DQ, UK
BDH Laboratory Supplies	P.O Box 2716, Plot 7 Bombo Road, Kampala, Uganda
Beckman Coulter	Oakley Court, Kingsmead Business Park, London Road, High Wycombe, HP11 1JU, UK
Bio-Rad	Bio-Rad House, Maxted Road, Hemel Hempstead, Hertfordshire, HP2 7DX, UK
Cell Signalling Technology	Dellaertweg 9b, 2316 WZ Leiden, The Netherlands
Corning	Fogostraat 12, 1060 LJ Amsterdam, The Netherlands
eBioscience	3rd Floor, 1 Ashley Road, Altrincham, Cheshire, WA14 2DT, UK

Fisher Scientific (Thermo Fisher)	168 Third Avenue, Waltham, MA, USA, 02451
Geneflow	Elmhurst Business Park, Lichfield, WS13 8EX, UK
Genscript	860 Centennial Avenue, Piscataway, NJ, 08854, USA
Gibco (Thermo Fisher)	168 Third Avenue, Waltham, MA, USA, 02451
GraphPad	7825 Fay Avenue, Suite 230, La Jolla, CA, 92037, USA
Greiner Bio-One	Brunel Way, Stroudwater Business Park, Stonehouse, Gloucestershire, GL10 3SX, UK
HyClone	GE Healthcare Life Sciences, Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA, UK
Illumina	5200 Illumina Way, San Diego, CA, 92 122 USA
Invitrogen (Thermo Fisher)	168 Third Avenue, Waltham, MA, USA, 02451
Jackson Immunoresearch	872 W Baltimore Pike, West Grove, PA, 19390, USA
Kühner AG	Adolf Kühner AG, Dinkelbergstrasse 1, 4127, Birsfelden, Switzerland
Labtech	11 Browning Road, Heathfield, East Sussex, TN21 8DB
Leica	9 Police Street, Manchester, M2 7LQ, UK
Marvel (Premier Foods)	Premier House, Griffiths Way, St Albans, Hertfordshire, AL1 2RE, UK
Merck	2000 Galloping Hill Road, Kenilworth, NJ, 07033, USA
Merial (Boehringer Ingelheim)	Ellesfield Avenue, Bracknell, Berkshire, RG12 8YS, UK
Microsoft	1 Microsoft Way, Redmond, WA, 98052-6399, USA
Miltenyi Biotec	Almac House, Church Lane, Bisley, Surrey, GU24 9DR, UK

Nikon	Shinagawa Intercity Tower C, 2-15-3, Konan, Minato-ku, Tokyo, 108-6290, Japan
Novus Biologicals	10730 E Briarwood Avenue, Centennial, CO, 80112, USA
Proteintech	196 Deansgate, Manchester, M3 3WF
Qiagen	Skelton House, Lloyd Street North, Manchester, M15 6SH, UK
Roche	Konzern Hauptsitz, Grenzacherstrasse 124, CD- 4070, Basel, Switzerland
Santa Cruz Biotechnology	Bergheimer Street, 89-2, 69115, Heidelberg, Germany
Sigma Aldrich	The Old Brickyard, New Road, Gillingham, SP8 4XT, UK
Scientific Laboratory Supplies (SLS)	Nottingham South and Wilford Industrial Estate, 22- 23 Ruddington Lane, West Bridgford, Nottingham, NG11 7EP, UK
STAR LAB	5 Tanners Drive, Blakelands, Milton Keynes, MK14 5BU, UK
Takara	Nojihigashi 7-4-38, Kusatu, Shiga, 525-0058, Japan
Thermo Fisher Scientific	168 Third Avenue, Waltham, MA, USA, 02451
Tree Star Inc.	340 A St Suite 206, Ashland, OR, 97520, USA
Veolia	8th Floor, 210 Pentonville Road, London, N1 9JY, UK
VWR	One Radnor Corporate Center, Building One, Suite 200, 100 Matsonford Road, Radnor, PA, 19087, USA
Zeiss	ZEISS House, Building 1030, Cambourne Business Park, Cambourne, CB23 6DW, Cambridge, UK

## 7.2 Supplementary Figures

### 7.2.1 Custom R Scripts

**# This script loads the Human and Mouse reference genomes, then:  
## alters the chromosome names to be unique and saves a single output**

**# Load the R package Biostrings:**

```
library(Biostrings)
```

**# Set the run IO:**

```
io <- list(
  'input.dir' = file.path('.', 'input'),
  'output' = file.path('.', 'output', 'combined.fasta.gz')
)
```

**# Specify the genomes:**

```
species <- c('human'='H', 'mouse'='M')
```

**# Load the reference genomes:**

```
input.genomes <- sapply(names(species), function(s){
  input.fasta <- file.path(io$input.dir, sprintf('%s.fasta.gz', s))
  message(sprintf('loading %s genome from "%s"...', s, input.fasta))
  return(readDNASTringSet(input.fasta))
}, simplify=FALSE)
```

**# Simplify and disambiguate the chromosome names:**

```
for(s in names(species)){
  message(sprintf('updating %s chromosome names...', s))
  n <- gsub('([^\s]+) .*', '\\1', names(input.genomes[[s]]), perl=TRUE)
  names(input.genomes[[s]]) <- sprintf('%s_%s', species[s], n)
}
```

**# Output the two datasets combined:**

```
message(sprintf('saving combined genome to "%s"...', io$output))
writeXStringSet(append(input.genomes$human, input.genomes$mouse),
  filepath=io$output, compress=TRUE, format='fasta')
```

**Supplementary Figure S1. R script for the concatenation of human and mouse reference genomes prior to generation of human-mouse mRNAseq alignment index, using R package Biostrings.**

```
library(Rsubread)
```

```
#Path to alignment BAM files
```

```
bams <- list.files("/path/to/bam/files", pattern="bam$", full.names=TRUE)
```

```
#Path to GTF annotation files
```

```
gtf <- "/path/to/gtf/*.gtf"
```

```
## Counting reads overlapping genes
```

```
geneCounts <- featureCounts(bams, annot.ext= gtf, isGTFAnnotationFile = TRUE,
  useMetaFeatures=TRUE, allowMultiOverlap=TRUE,
  countMultiMappingReads=TRUE,
  fraction=TRUE, nthreads=4, ignoreDup=FALSE)
```

```
## Counting reads overlapping exons
```

```
exonCounts <- featureCounts(bams, annot.ext= gtf, isGTFAnnotationFile = TRUE,
  useMetaFeatures=FALSE, allowMultiOverlap=TRUE,
  countMultiMappingReads=TRUE,
  fraction=TRUE, nthreads=4, ignoreDup=FALSE)
```

```
dataGenes <- cbind(geneCounts$annotation, geneCounts$counts)
```

```
dataExons <- cbind(exonCounts$annotation, exonCounts$counts)
```

```
write.table(x=as.data.frame(dataGenes), quote=FALSE, sep="\t",
  file="gene_counts_dups.txt")
```

```
write.table(x=as.data.frame(dataExons), quote=FALSE, sep="\t",
  file="exon_counts_dups.txt")
```

**Supplementary Figure S2. R script for mRNAseq read count generation from BAM files using R package Rsubread.**

**#Import dataset .xlsx****##Requires normalised expression values and gene names as Ensembl IDs****###Statistical cutoff for differentially expressed genes is  $p \leq 0.05$** 

```
Simple_data_matrix_p005 <- read_excel("N:/path/to/file/file.xlsx")
view(Simple_data_matrix_p005)
```

**#Load R package Scrان**

```
library(Scrان)
```

**#Create a data matrix****##This uses the data.matrix tool to create a data matrix, and cbind, to combine the data together****###Note, GFP+ = proliferating, GFP+CV+ = dormant, S = sort**

```
mat <- data.matrix(cbind(Simple_data_matrix_p005$GFP+S1,
Simple_data_matrix_p005$GFP+S2, Simple_data_matrix_p005$GFP+S3,
Simple_data_matrix_p005$GFP+CV+S1, Simple_data_matrix_p005$GFP+CV+S2,
Simple_data_matrix_p005$GFP+CV+S3))
```

**#Create pairs of human genes for cell cycle phase reference using the Scrان package**

```
hm.pairs <- readRDS(system.file("exdata", "human_cycle_markers.rds", package =
"scrان"))
```

**#Perform the cyclone function**

```
cyclone(mat, hm.pairs, gene.names=Simple_data_matrix_p005$Ensembl, iter=1000,
min.iter=100, min.pairs=50, BPPARAM=SerialParam(), verbose=FALSE,
subset.row=NULL)
```

**#Creation of assignments for downstream R plots, where required**

```
assignments <- cyclone(mat, hm.pairs,
gene.names=Simple_data_matrix_p005$Ensembl, iter=1000, min.iter=100,
min.pairs=50, BPPARAM=SerialParam(), verbose=FALSE, subset.row=NULL)
```

**#Plotting cell cycle phase graphs in R, for example G1 against G2/M scores**

```
plot(assignments$score$G1, assignments$score$G2M, xlab = "G1 score", ylab = "G2/M
score", pch=16)
```

**#Can view specific cell cycle phase scores:**

```
view(assignments$score$G1)
```

**Supplementary Figure S3. R script for Cell cycle phase analysis using R package Scrان.**

**#This takes raw Affymetrix datasets (CEL file input) and normalises them**

```
setwd("/Users/Ashley/Documents/R_Studio/GSE2034")
celFiles <- list.celfiles()
affyRaw <- read.celfiles(celFiles)
library(pd.hugene.2.0.st)
eset <- rma(affyRaw)
write.exprs(eset, file="GSE2034.txt")
```

**Supplementary Figure S4. R script for the normalisation of microarray data (CEL files).**

7.2.2 Human Biglycan mRNA Sequence in Dormant MDA-MB-231 cells

```

Matches ( ): 1106
Mismatches ( # ): 1
Gaps ( ): 0
Unattempted ( . ): 0

1 atgtggcccttggtggcctcgtgtctctgtgcccctgagccaggccctgcccttgagcagagaggcttctgggaactcaccctggacgatgggccaat 100
1 ATGTGGCCCTGTGGCCCTCGTGTCTCTGTGGCCCTGAGCCAGGCCCTGCCCTTGAGCAGAGAGGCTTCTGGGACTTCACCTGGACGATGGGCCAI 100

101 tcgatgtaacgatgaggaagcttcgggcgctgacacctcggggctcctggaccggactctgtcacaccacctacagcgccatgtgtcctttoggctg 200
101 TCATGATGAACGATGAGGAAGCTTCGGGCGCTGACACCTCGGGCGTCTGGACCCGGACTCTGTACACCCACCTACAGCGCCATGTGTCTTCGGCTG 200

201 ccaactgccactcgggggtggtcagtgctccgacctgggtctgaagtctgtgcccagaagatctcccctgacaccacgctgtggacctgcagaacaac 300
201 CCACTGCCACTCGGGGTGTTTCAAGTCTCCGACCTGGGTCTGAAGTCTGTGCCAAGAGATCTCCCTGACACCACGCTGTGGACTGCAGAACCAAC 300

301 gacatctccgagctccgcaaggatgactcaagggtctccagcactctacgcccctcgtcctggtaacaacaagatctccaagatccatgagaaggcct 400
301 GACATCTCCGAGCTCCGCAAGGATGACTCAAGGTCTCCAGCACCCTACGCCCTCGTCTGGTGAACAACAAGATCTCCAAGATCCATGAGAAGGCCCT 400

401 tcagcccactgcggaagctgcagaagctctacatctccaagaaccctcgtggagatcccgcccacctaccagctccctggtagctccgcatcca 500
401 TCAGCCCCTGCGGAAGCTGCAGAAGCTCTACATCTCCAAGAACCCTCGTGGAGATCCCGCCCACTACCCAGCTCCCTGGTGGAGCTCCGCATCCA 500

501 cgacaaccgcatccgcaagggtcccgaagggtggtcaggggctccggaacatgaactgcacagatggggcggaaacccactggagaacagtggtctt 600
501 CGACAACCgCATCCGCAAGGTCGCCAAGGAGTGTTCAGTGGGCTCCGGAACATGAACCTGCATCGAGATGGCGGGAAACCCACTGGAGAACAGTGGCTTI 600

601 gaacctggagccttcgatggcctgaagctcaactacctcgcgcatctcagaggccaagctgactggcatcccacaagacctccctgagacctgaatgaac 700
601 GAACCTGGAGCCTTCGATGGCCTGAAGCTCAACTACCTCGGCATCTCAGAGGCCAAGCTGACTGGCATCCCCAAGACCTCCCTGAGACCTGAATGAAC 700

701 tccacctagaccacaacaaaatccaggccatogaactggaggacctgctctcactccaagctgtacaggctgggacctaggccacaaccagatcaggat 800
701 TCCACCTAGACCACAACAAAATCCAGGCCATCGAACTGGAGGACCTGCTCTCGCTACTCCAAGCTGTACAGGCTGGGCTAGGCCACAACCAGATCAGGAT 800

801 gatcgagaacggggagcctgagcttctcctcccaccctccgggagctccacttggacaacaacaagtggccagggtgcctcagggtccagacctcaag 900
801 GATCGAGAACGGGGAGCCTGAGCTTCTGCCACCCTCCGGGAGCTCCACTTGGACAACAACAAGTGGCCAGGGTGCCTCAGGGCTCCAGACCTCAAG 900

901 ctctccaggtggtctatctgcactccaacaacatcaccaaagtgggtgtaacgacttctgtcccattggggttcgggggtgaagcgggctactacaacg 1000
901 CTCTCCAGGTGGTCTAICTGCCTCCAACAACATCACAAAGTGGGTGTAACGACTTCTGTCCATGGGCTTCGGGGTGAAGCGGGCTACTACAACG 1000

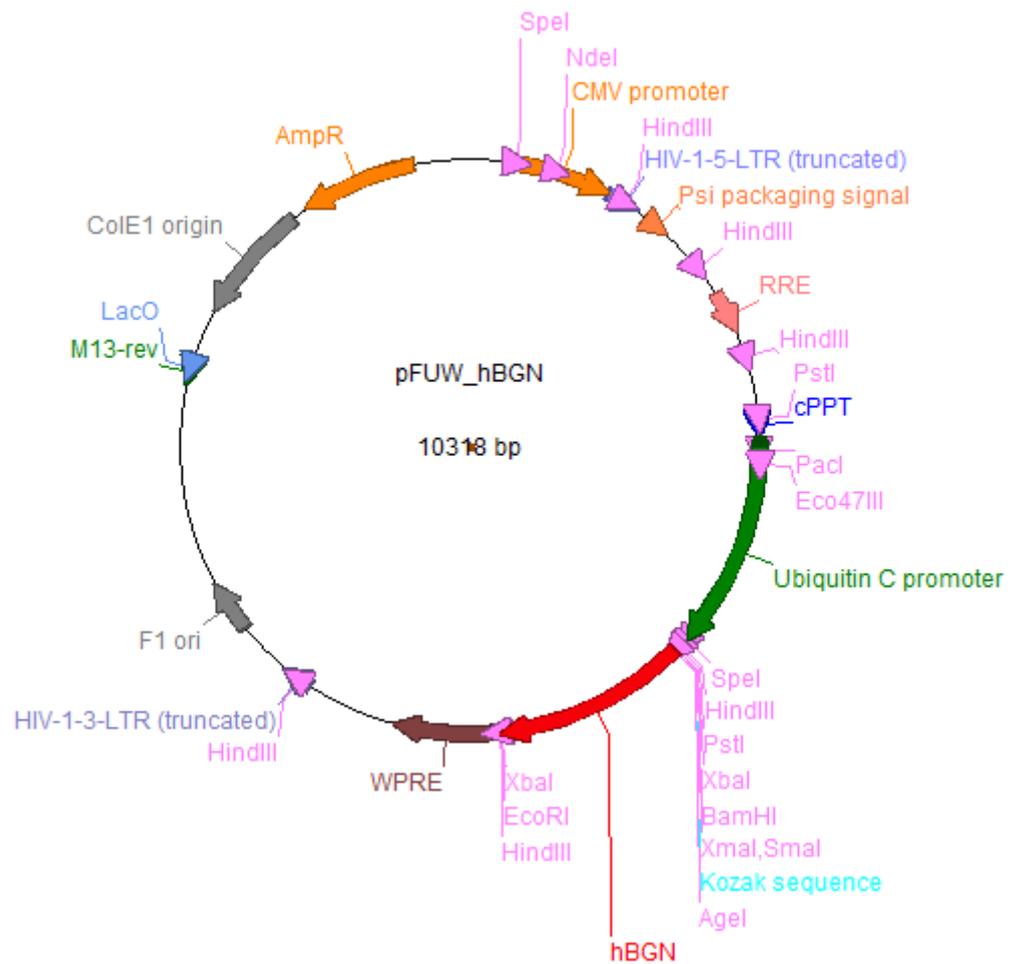
1001 gcatcagcctcttcaacaaccccgctccctactgggaggtgcagccggccacttccgctgctcactgaccgctggccatccagtttggcaactacaa 1100
1001 GCATCAGCCTCTTCAACAACCCCGTCCCTACTGGGAGGTGCAGCCGGCCACTTCCGCTGCTCAGTACCCTGGCCATCCAGTTTGGCAACTACAA 1100

1101 aaagtag 1107
1101 AAAGTAG 1107

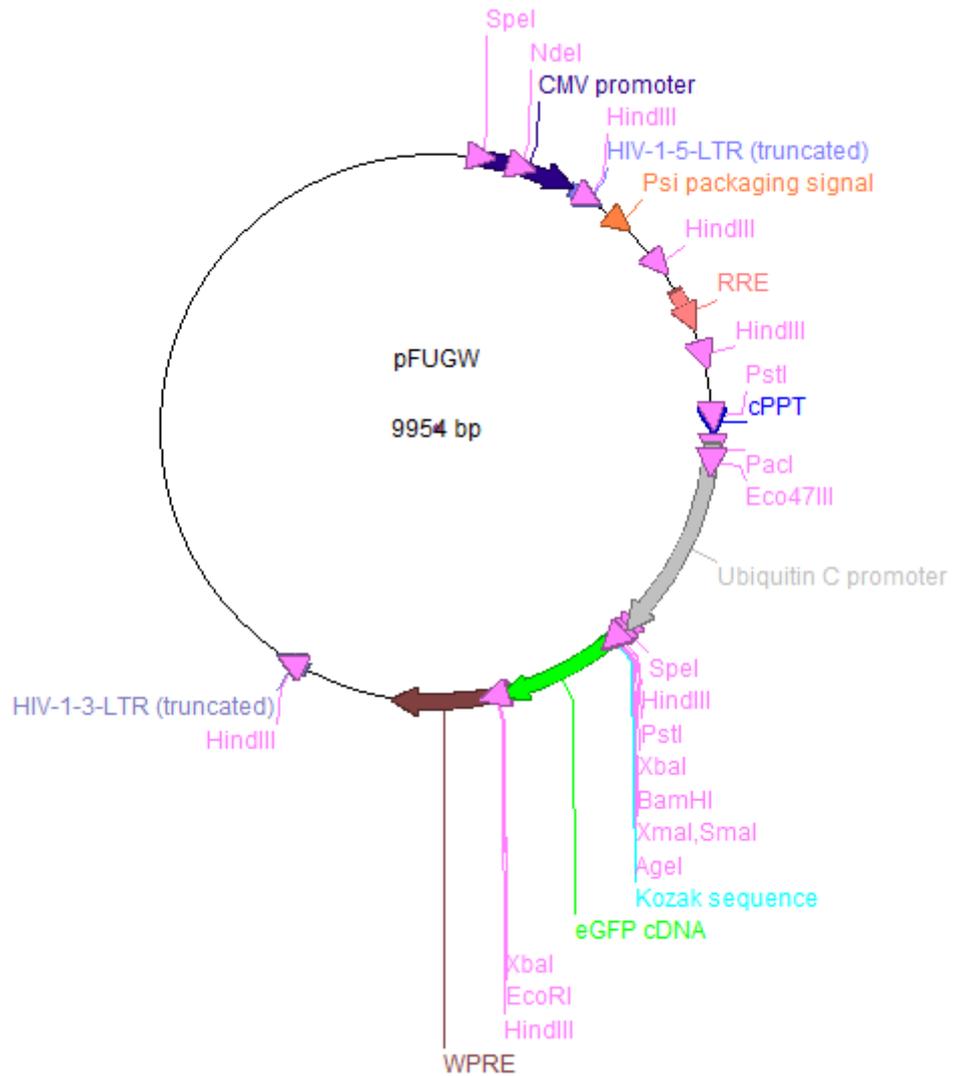
```

Supplementary Figure S5. Alignment of dormant cell human BGN mRNA sequence to known human BGN sequence. Lowercase = known consensus sequence for human biglycan mRNA (Ensembl ID = ENSG00000182492). Uppercase = mRNA sequence for dormant cell human biglycan, as determined by mRNAseq. | denotes alignment. # denotes mismatch.

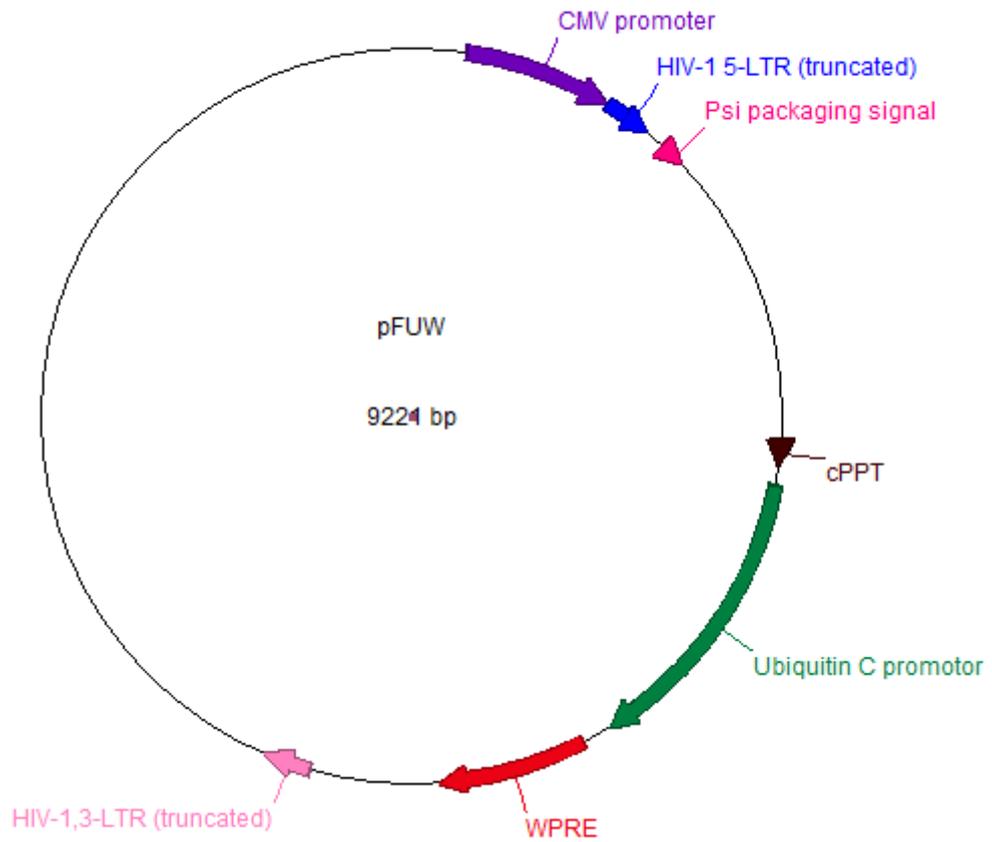
### 7.3.3 Plasmid Maps



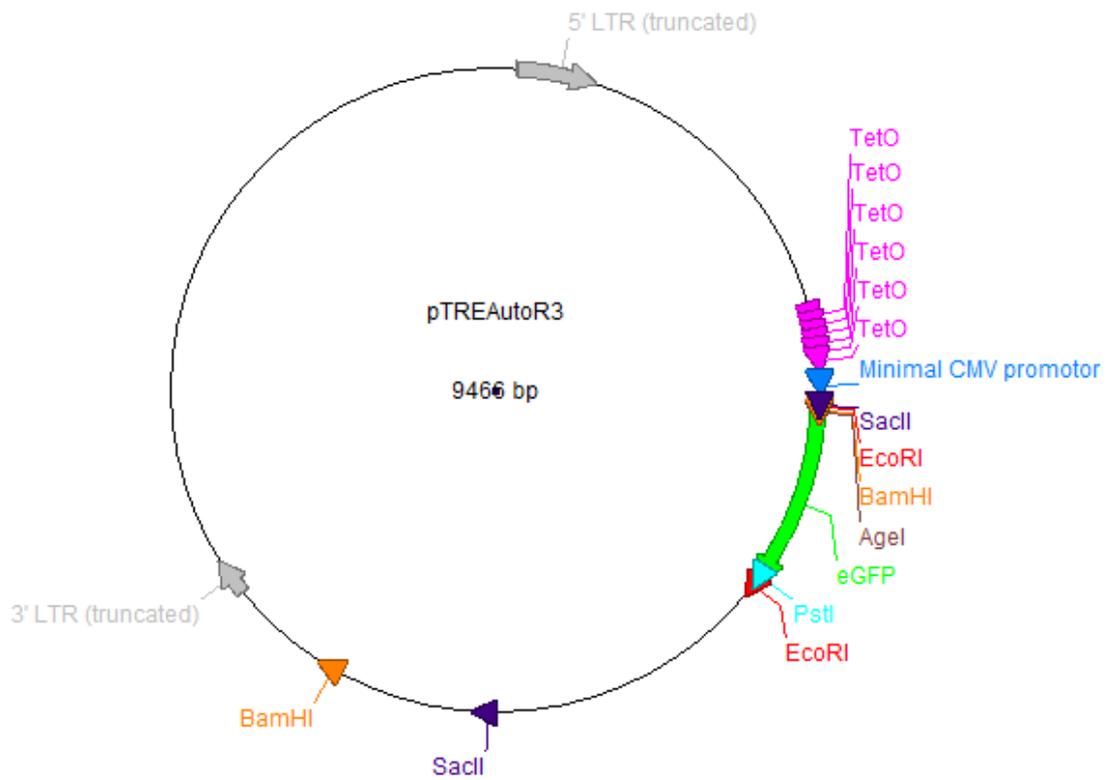
**Supplementary Figure S6. Plasmid map for the ubiquitin C promoter-controlled human biglycan lentiviral expression vector, pFUW\_hBGN.** Major components key: cytomegalovirus (CMV), Human Immunodeficiency Virus 1 (HIV-1), Rev response element (RRE), central polypurine tract (cPPT), human biglycan (hBGN) insert, Woodchuck Hepatitis Virus (WHP) posttranscriptional regulatory element (WPRE), long terminal repeat (LTR).



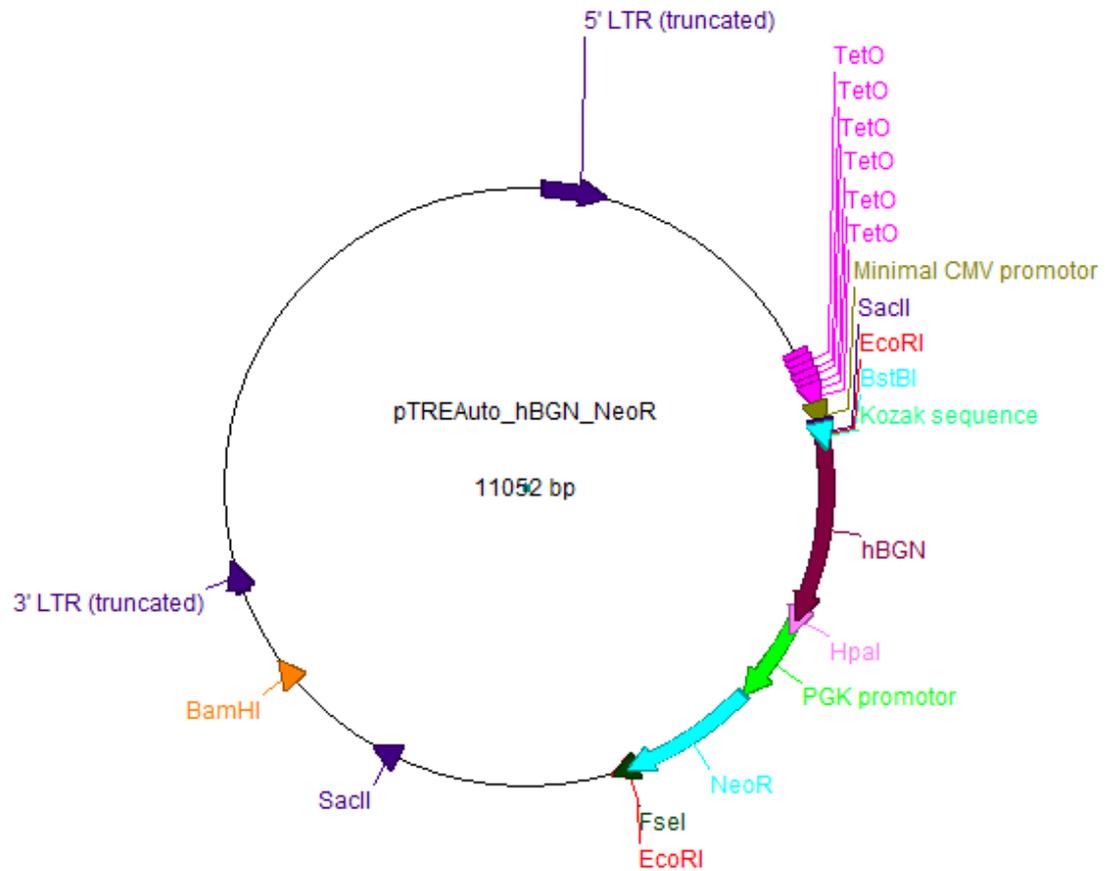
**Supplementary Figure S7. Plasmid map for the ubiquitin C promoter-controlled enhanced green fluorescent protein (eGFP) lentiviral expression vector, pFUGW.** Major components key: cytomegalovirus (CMV), Human Immunodeficiency Virus 1 (HIV-1), Rev response element (RRE), central polypurine tract (cPPT), Woodchuck Hepatitis Virus (WHP) posttranscriptional regulatory element (WPRE), long terminal repeat (LTR), enhanced GFP (eGFP).



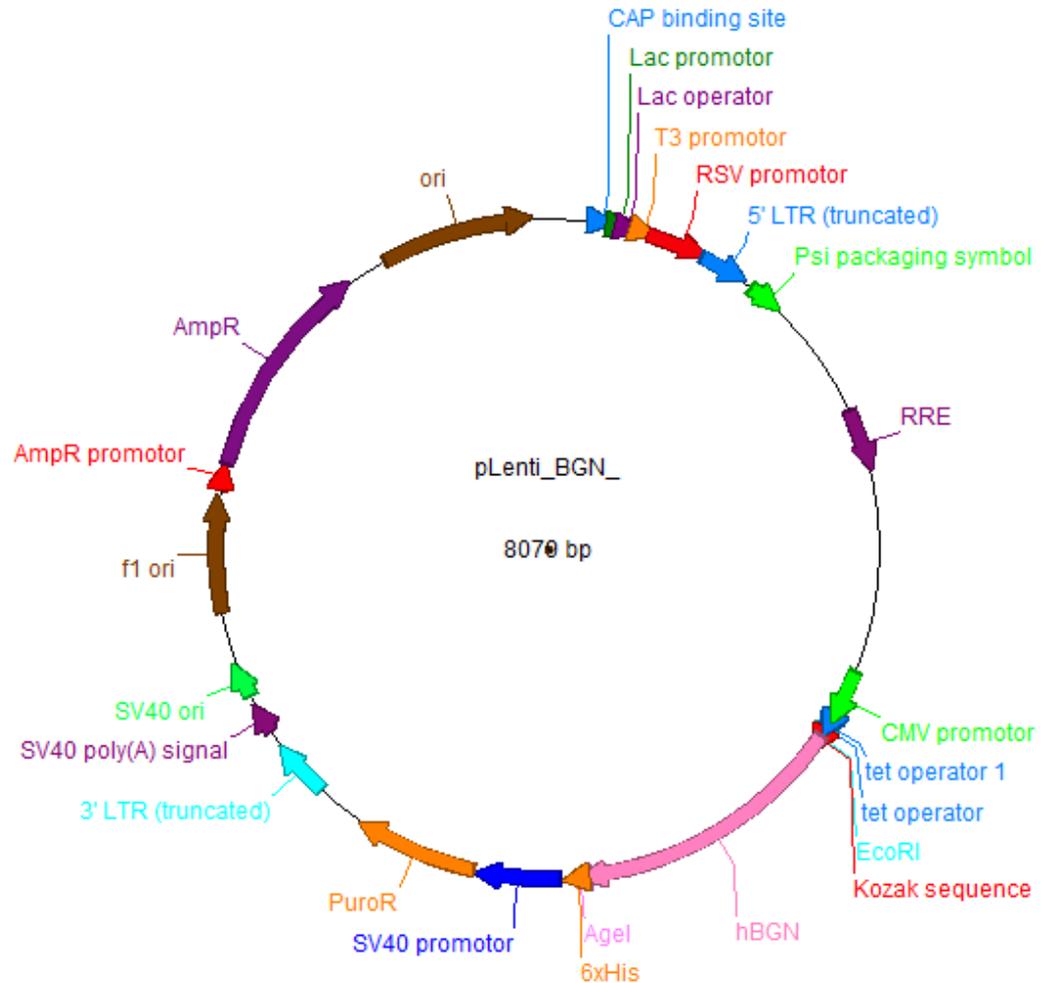
**Supplementary Figure S8. Plasmid map for the ubiquitin C promoter-controlled lentiviral expression vector, pFUW (empty vector).** Major components key: cytomegalovirus (CMV), Human Immunodeficiency Virus 1 (HIV-1), Rev response element (RRE), central polypurine tract (cPPT), long terminal repeat (LTR).



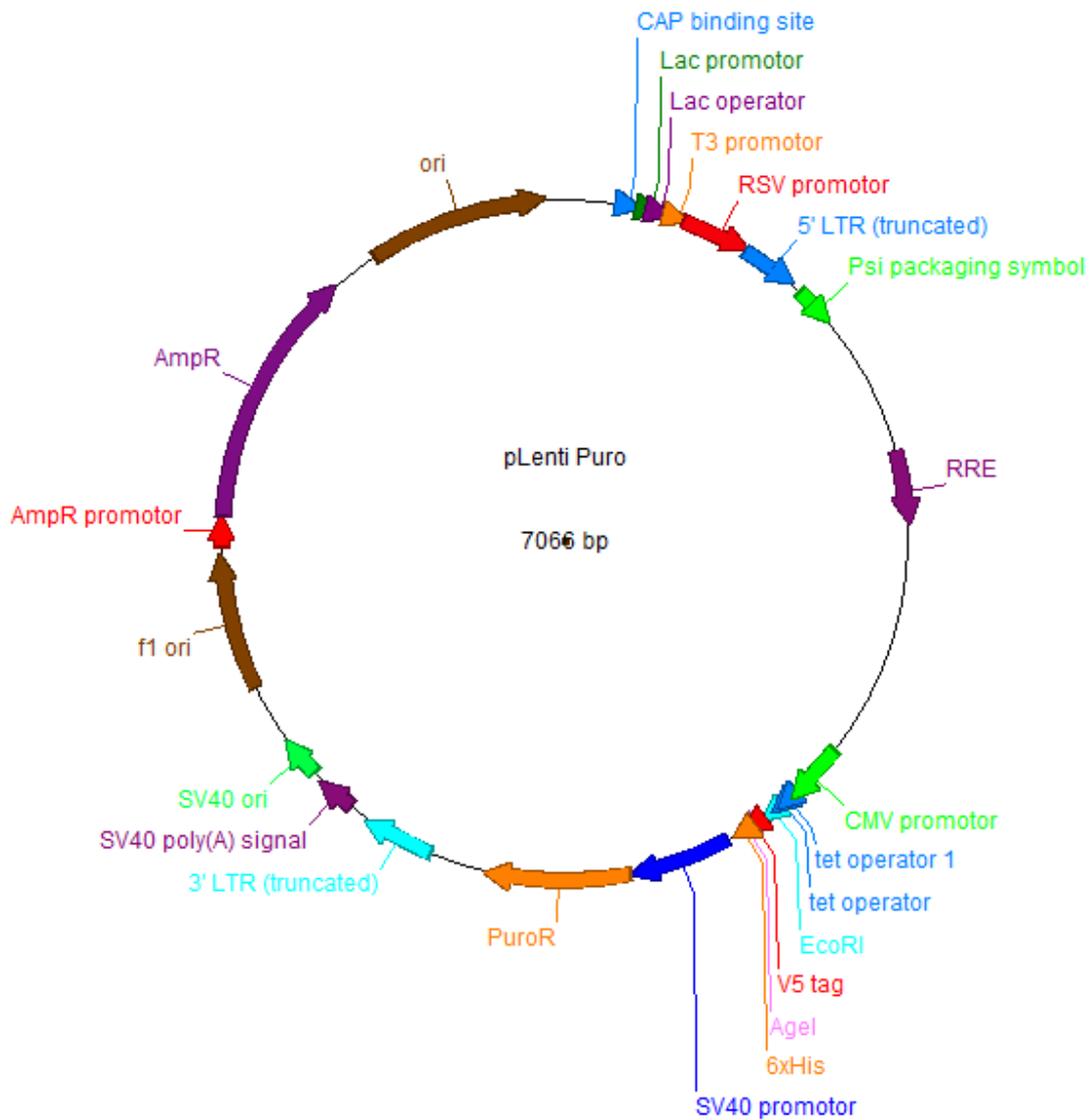
**Supplementary Figure S9. Plasmid map for the doxycycline-inducible (Tet-ON) enhanced green fluorescent protein (eGFP) lentiviral expression vector, pTREAutoR3.** Major components key: tetracycline-ON (TetO) long terminal repeat (LTR), enhanced GFP (eGFP).



**Supplementary Figure S10. Plasmid map for the doxycycline-inducible (Tet-ON) human biglycan lentiviral expression vector, pTREAuto\_hBGN\_NeoR (pTREA\_BGN).** Major components key: long terminal repeat (LTR), cytomegalovirus (CMV), human biglycan (hBGN), phosphoglycerate kinase (PGK), neomycin resistance (NeoR).



**Supplementary Figure S11. Plasmid map for the doxycycline-inducible (Tet-OFF) human biglycan lentiviral expression vector, pLenti\_BGN.** Major components key: Catabolite activator protein (CAP), lactose (Lac), Rous sarcoma virus (RSV), long terminal repeat (LTR), Rev response element (RRE), cytomegalovirus (CMV), tetracycline (tet), human biglycan (hBGN), puromycin resistance (PuroR), ampicillin resistance (AmpR).



**Supplementary Figure S12. Plasmid map for the doxycycline-inducible (Tet-OFF) lentiviral expression vector, pLenti Puro (empty vector).** Major components key: Catabolite activator protein (CAP), lactose (Lac), Rous sarcoma virus (RSV), long terminal repeat (LTR), Rev response element (RRE), cytomegalovirus (CMV), tetracycline (tet), puromycin resistance (PuroR), ampicillin resistance (AmpR).

## Bibliography

1. Muir, D., Kanthan, R., Kanthan, S. C., (2003). Male Versus Female Breast Cancers. *Arch Pathol Lab Med.* 27: 36-41
2. Global Cancer Observatory. (2019). Breast. Available: <https://gco.iarc.fr/today/data/factsheets/cancers/20-Breast-fact-sheet.pdf>. Last accessed 23 September 2019
3. Akram, M., Iqbal, M., Daniyal, M., Khan, A. U. (2017). Awareness and current knowledge of breast cancer. *Biological Research.* 50: 33
4. Apple, S. K., Bassett, L. W. (2018). 'Chapter 3 – Normal breast anatomy and histology' in Lee, C. I., Lehman, C. D., Bassett, L. W. eds. *Breast Imaging*. Oxford: Oxford University Press
5. Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., Ji, X., Liu, W., Huang, B., Luo, W., Liu, B., Lei, Y., Du, S., Vuppapapati, A., Luu, H. H., Haydon, R. C., He, Y., Ren, G. (2018). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes and Diseases.* 5: 77-106
6. Zucca-Matthes, G., Urban, C., Vallejo, A. (2016). Anatomy of the nipple and breast ducts. *Gland Surgery.* 5(1): 32-36
7. Dai, X., Ting, L., Bai, Z., Yang, Y., Liu, X., Zhan, J., Shi, B. (2015). Breast cancer intrinsic subtype classification, clinical use and future trends. *American Journal of Cancer Research.* 5(10): 2929-2943
8. Fentiman, I.S. (2016). Male breast cancer is not congruent with the female disease. *Critical Reviews in Oncology/Hematology.* 101: 119-124
9. Yalaza, M., Inan, A., Bozer, M. (2016). Male breast Cancer. *The Journal of Breast Health.* 12(1): 1-8
10. Kaminska, M., Ciszewski, T., Lopacka-Szatan, K., Miotla, P., Staroslawska, E. (2015). Breast cancer risk factors. *Przegląd Menopauzalny.* 14(3): 196 – 202
11. Chew, H. K. (2001). Adjuvant therapy for breast cancer: who should get what? *Western Journal of Medicine.* 174(4): 284-287
12. Office for National Statistics. (2016). *Cancer survival by stage at diagnosis for England (experimental statistics): Adults diagnosed 2012, 2013 and 2014 and followed*

up to 2015. Available at:

<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalbystageatdiagnosisforenglandexperimentalstatistics/adultsdiagnosed20122013and2014andfollowedupto2015#main-points>

[Accessed 11 September 2019]

13. Coleman, M.P. (2000). Trends in Breast Cancer Incidence, Survival, and Mortality. *Lancet*. 356(9229): 590-1
14. Zhang, X. H. F., Giuliano, M., Trivedi, M. V., Schliff, R., Osborne, C. K. (2014). Metastasis dormancy in estrogen receptor-positive breast cancer. *Clinical Cancer Research*. 19(23)
15. Gerber, B., Freund, M., Reimer, T. (2010). Recurrent breast cancer: Treatment strategies for maintaining and prolonging good quality of life. *Deutsches Arzteblatt International*. 107(6): 85 – 91
16. Kim, H., Choi, D. H., Park, W., Huh, S. J., Nam, S. J., Lee, J. E., Ahn, J. S., Im, Y. (2013). Prognostic factors for survivals from first relapse in breast cancer patients: Analysis of deceased patients. *Radiation Oncology Journal*. 31(4): 222-227
17. Carlson, P., Dasgupta, A., Grzelak, C. A., Kim, J., Barrett, A., Coleman, I. M., Shor, R. E., Goddard, E. T. Dai, J., Schweitzer, E. M., Lim, A. R., Crist, S. B., Cheresch, D. A., Nelson, P. S., Hansen, K. C., Ghajar, C. M. (2019). Targeting the perivascular niche sensitizes disseminated tumour cells to chemotherapy. *Nature Cell Biology*. 21: 238 – 250
18. Jin, X., Mu, P. (2015). Targeting Breast Cancer Metastasis. *Breast Cancer: Basic and Clinical Research*. 9(S1): 23-34
19. Leone, J. P., Leone, B. A. (2015). Breast cancer brain metastases: The last frontier. *Experimental Hematology and Oncology*. 4: 33
20. Witzel, I., Oliveira-Ferrer, L., Pantel, K., Muller, V., Wilkman, H. (2016). Breast cancer brain metastasis: Biology and new clinical perspectives. *Breast Cancer Research*. 18(8)
21. Niwinska, A., Murawska, M., Pogoda, K. (2010). Breast cancer brain metastases: Differences in survival depending on biological subtype, RPA, RTOG prognostic class and systemic treatment after whole-brain radiotherapy. *Annals of Oncology*. 21(5): 942-948

- 22.** Balkwill, F. R., Capasso, M., Hagemann, T. (2012). The tumour microenvironment at a glance. *Journal of Cell Science*. 125: 5591 – 5596
- 23.** Joyce, J. A., Pollard, J. W. (2009). Microenvironmental regulation of metastasis. *Nature Reviews Cancer*. 9: 239 – 252
- 24.** Polyak, K. (2011). Heterogeneity in breast cancer. *Journal of Clinical Investigation*. 121(10): 3786 – 3788
- 25.** Venur, V. A., Leone, J. P. (2016). Targeted therapies for brain metastases from breast cancer. *International Journal of Molecular Sciences*. 17(9): 1543
- 26.** Kabraji, S., Ni, J., Lin, N. U., Xie, S., Winer, E. P., Zhao, J. J. (2018). Drug resistance in HER2-positive breast cancer brain metastases: Blame the barrier or the brain? *Clinical Cancer Research*. 24(8): 1795-1804
- 27.** Takahashi, H., Isogawa, M. (2018). Management of breast cancer brain metastases. *Chinese Clinical Oncology*. 7(3)
- 28.** Engelhardt, B., Vajkoczy, P., Weller, R. O. (2017). The movers and shapers in immune privilege of the CNS. *Nature Immunology*. 18(2): 123 – 131
- 29.** Soto, M. S., Sibson, N. R. (2018). The multifarious role of microglia in brain metastasis. *Frontiers in Cellular Neuroscience*. 12: 414
- 30.** Ramakrishna, N., Temin, S., Chandarlapaty, S., Crews, J. R., Davidson, N. E., Esteva, F. J., Giordano, S. H., Gonzalez-Angulo, A. M., Kirshner, J. J., Krop, I., Levinson, J., Modi, S., Patt, D. A., Perez, E. A., Perimutter, J., Winer, E. P. Lin, N. U. (2014). Recommendations on disease management for patients with advanced human epidermal growth factor receptor 2-positive breast cancer and brain metastases: American Society of Clinical Oncology clinical practice guideline. *Journal of Clinical Oncology*. 32(9): 2100 – 2108
- 31.** Baumert, B. G., Rutten, I., Dehing-Oberjje, C., Twjinstra, A., Dirx, M. J., Debougnoux-Huppertz, R. M., Lambin, P., Kubat, B. (2006). A pathology-based substrate for target definition in radiosurgery of brain metastases. *International Journal of Radiation Oncology, Biology, Physics*. 66(1): 187 – 194
- 32.** Weil, R. J., Palmieri, D. C., Bronder, J. L., Stark, A. M., Steeg, P. S. (2005). Breast cancer metastasis to the central nervous system. *The American Journal of Pathology*. 167(4): 913 – 920

33. Nagao, E., Yoshiura, T., Obara, H. M., Yamashita, K., Kamano, H., Takayama, Y., Kobayashi, K., Honda, H. (2011). 3D turbo spin-echo sequence with motion-sensitized driven-equilibrium preparation for detection of brain metastases on 3T MR imaging. *American Journal of Neuroradiology*. 32(4): 664 – 670
34. Wise, J. (2014). Benefits of trastuzumab outweigh its harms, says Cochrane review. *British Medical Journal*. 348: g3835
35. Witzel, I., Oliveira-Ferrer, L., Pantel, K., Müller, V., Wikman, H. (2016). Breast Cancer Brain Metastases: Biology and New Clinical Perspectives. *Breast Cancer Research*. 18: 8
36. Bender, E. T., Tome, W. A. (2011). Distribution of brain metastases: Implications for non-uniform dose prescriptions. *The British Institute of Radiology*. 84(1003)
37. Kyeong, S., Cha, Y. J., Ahn, S. G., Suh, S. H., Son, E. J., Ahn, S. J. (2017). Subtypes of breast cancer show different spatial distributions of brain metastases. *PLoS One*. 12(11): e0188542
38. Anderson, A. (2013). Metastatic breast cancer to the brain: A clinical primer for translational investigation. In: Jandial, R. *Madame Curie Bioscience Database*. Austin (TX): Landes Bioscience
39. Maurer, C., Tulpin, L., Moreau, M., Dumitrescu, C., de Azambuja, E., Paesmans, M., Nogaret, J., Piccart, M. J., Awada, A. (2018). Risk factors for the development of brain metastases in patients with HER2-positive breast cancer. *ESMO Open*. 3: e000440
40. Paget, S. (1889). The distribution of secondary growth in cancer of the breast. *The Lancet*. 133(3421): 571 – 573
41. Weidle, U., Birzele, F., Kollmorgen, G., Ruger, R. (2016). Dissection of the process of brain metastasis reveals targets and mechanisms for molecular-based intervention. *Cancer Genomics and Proteomics*. 13(4): 245 – 258
42. Touil, Y., Segard, P., Ostyn, P., Begard, S., Aspod, C., El Machhour, R., Masselot, B., Vandomme, J., Flamenco, P., Idziorek, T., Flgeac, M., Formstecher, P., Quesnel, B., Polakowska, R. (2016). Melanoma dormancy in a mouse model is linked to GILZ/FOXO3A-dependent quiescence of disseminated stem-like cells. *Scientific Reports*. 6(30405)

43. Wang, S., Lin, S. (2013). Tumour dormancy: Potential therapeutic target in tumour recurrence and metastasis prevention. *Experimental Hematology and Oncology*. 2(29)
44. Yeh, A., Ramaswamy, S. (2015). Mechanisms of cancer cell dormancy – Another hallmark of cancer? *Cancer Research*. 75(23): 5014 – 5022
45. Willis, R. A. (1934). The spread of tumours in the human body. *British Journal of Surgery*. 22(85): 196
46. Galao, L., Criscitiello, C., Fumagalli, L., Locatelli, M., Manunta, A., Esposito, A., Minchella, I., Goldhirsch, A., Curigliano, G. (2013). Tumour dormancy and clinical implications in breast cancer. *ecancer*. 7(320)
47. Senft, D., Ronai, Z. (2015). Immunogenic, cellular and angiogenic drivers of tumour dormancy – A melanoma view. *Pigment Cell and Melanoma Research*. 29: 27 – 42
48. Fidler, I. (2002). The organ microenvironment and cancer metastasis. *Differentiation*. 70: 498 – 505
49. Folkman, J. (1986). How is Blood Vessel Growth Regulated in Normal and Neoplastic Tissue? G. H. A. Clowes Memorial Award Lecture. *Cancer Research*. 46: 467-473
50. Gimbrone, M.A. Cotran, R.S., Leapman, S.B., Folkman, J. (1974). Tumour Growth and Neovascularisation: an Experimental Model using Rabbit Cornea. *Journal of the National Cancer Institute*. 52: 413-427
51. Nuamov, G., Zurakowski, D., Kang, S.Y., Sampson, D., Flynn, E., Watnick, R.S., Straume, O., Akslen, L.A., Folkman, J., Almog, N. (2006). A Model of Human Tumour Dormancy: An Angiogenic Switch from the Nonangiogenic Phenotype. *Journal of the National Cancer Institute*. 98(5): 316-325
52. Natale, G., Bocci, G. (2018). Does metronomic chemotherapy induce tumour angiogenic dormancy? A review of available preclinical and clinical data. *Cancer Letters*. 432: 28 – 37
53. Hurst, R. E., Bastain, A., Bailey-Downs, L., Ihnat, M. A. (2016). Targeting dormant micrometastases: Rationale, evidence to date and clinical implications. *Therapeutic Advances in Medical Oncology*. 8(2): 126 – 137

- 54.** Sosa, M. S., Bragado, P., Aguirre-Ghiso, J. A. (2014). Mechanisms of disseminated cancer cell dormancy: An awakening field. *Nature Reviews Cancer*. 14: 611 – 622
- 55.** Hoffman, M., Pasch, S., Schamberger, T., Maneck, M., Mohlendick, B., Schumacher, S., Brockhoff, G., Knoefel, W. T., Izbicki, J., Polzer, B., Stoecklein, N. H., Klein, C. A. (2017). Diagnostic pathology of early systemic cancer: *ERBB2* gene amplification in single disseminated cancer cells determines patient survival in operable esophageal cancer. *International Journal of Cancer*. 42(4): 833 – 843
- 56.** Van't Veer, L. J., Dal, H., van de Vijver, M. J., He, Y. D., Hart, A. A. M., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Wilteveen, A. T., Schreiber, G. J., Kerbhoven, R. M., Roberts, C., Linsley, P. S., Berbaris, R. Friend, S. H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 415: 530 – 536
- 57.** Noltenius, C., Noltenius, H. (1985). Dormant tumour cells in liver and brain. An autopsy study on metastasizing tumours. *Pathology – Research and Practice*. 197(4-5): 504 – 11
- 58.** Udagawa, T. (2008). Tumour dormancy of primary and secondary cancers. *APMIS*. 116(7): 615 – 628
- 59.** Klein, C. A. (2009). Parallel progression of primary tumours and metastases. *Nature Reviews Cancer*. 9: 302 – 312
- 60.** Werner-Klein, M., Scheitler, S., Hoffman, M., Isabelle, H., Dietz, K., Lehnert, P., Naimer, V., Polzer, B., Treitschke, S., Werno, C., Markiewicz, A., Weidele, K., Czyz, Z., Hohenleutner, U., Hafner, C., Haferkamp, S., Berneburg, M., Rummele, P., Ulmer, A., Klein, C. A. (2018). Genertic alterations driving metastatic colony formation are acquired outside of the primary tumour in melanoma. *Nature Communications*. 9(595)
- 61.** Eyles, J., Puaux, A., Wang, X., Toh, B., Prakash, C., Hong, M., Tan, T. G., Zheng, L., Ong, L. C., Jin, Y., Kato, M., Prevost-Blondel, A., Chow, P., Yang, H., Abastado, J. (2010). Tumour cells disseminate early, but immunosurveillance limits metastatic outgrowth, in a mouse model of melanoma. *Journal of Clinical Investigation*. 120(6): 2030 – 2039
- 62.** Hüseman, Y., Geigl, J. B., Schubert, F., Musiani, P., Meyer, P., Meyer, M., Burghart, E., Forni, G., Elis, S., Fehm, T., Riethmuller, G., Klein, C. A. (2008). Systemic spread is an early step in breast cancer. *Cancer Cell*. 13: 58 – 68
- 63.** Harper, K. L., Sosa, M. S., Entenberg, D., Hosseini, H., Cheung, J., F., Nobre, R., Avivar-Valderas, A., Nagi, C., Girnius, N., Davis, R. J., Farias, E. F., Condeelis, J., Klein, C.

- A., Aguirre-Ghiso, J. A. (2016). Mechanism of early dissemination and metastasis in HER2<sup>+</sup> mammary cancer. *540*: 588 – 592
- 64.** Schardt, J. A., Meyer, M., Hartmann, C. H., Schubert, F., Schmidt-Kittler, O., Fuhrmann, C., Polzer, B., Petronio, M., Eils, R., Klein, C. A. (2005). Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer. *Cancer Cell*. *8*(3): 227 – 239
- 65.** Hosseini, H., Obradovic, M. M. S., Hoffman, M., Harper, K. L., Sosa, M. S., Werner-Klein, M., Nanduri, L. K., Werno, C., Ehrl, C., Maneck, M., Patwary, N., Haunschild, G., Guzvic, M., Reimelt, C., Grauvogl, M., Eichner, N., Weber, F., Hartkopf, A. D., Taran, F., Brucker, S. Y., Fehm, T., Rack, B., Buchholz, S., Spang, R., Meister, G. (2016). Early dissemination seeds metastasis in breast cancer. *Nature*. *540*: 552 – 558
- 66.** Sänger, N., Effenberger, K. E., Riethdorf, S., Van Haasteren, V., Gauwerky, J., Wiegratz, I., Strebhardt, K., Kaufmann, M., Pantel, K. (2011). Disseminated tumour cells in the bone marrow of patients with ductal carcinoma *in situ*. *International Journal of Cancer*. *129*(10)
- 67.** Cutuli, B., Cohen-Solal-Le-Nir, C., De Lafontan, B., Mignotte, H., Fichet, V., Fay, R., Servent, V., Giard, S., Charra-Brunaud, C., Auvray, H., Penault-Llorca, F., Charpentier, J. -C. (2001). Ductal carcinoma *in situ* of the breast results of conservative and radical treatments of 716 patients. *European Journal of Cancer*. *37*(18): 2365 – 2372
- 68.** Schmidt-Kittler, O., Ragg, T., Daskalakis, A., Granzow, M., Ahr, A., Blankenstein, T. J., Kaufmann, M., Diebold, J., Arnholdt, H., Muller, P., Bischoff, J. Harich, D., Schlimok, G., Elis, R., Klein, C. A. (2003). From latent disseminated cells to overt metastasis: Genetic analysis of systemic breast cancer progression. *PNAS*. *77*37 – 7742
- 69.** Heyn, C., Ronald, J.A., Ramadan, S.S., Snir, J.A., Barry, A.M., MacKenzie, L.T., Mikulis, D.J., Palmieri, D., Bronder, J.L., Steeg, P.S., Yoneda, T., MacDonald, I., Chambers, A.F., Rutt, B.K., Foster, P. (2006). In Vivo MRI of Cancer Cell Fate at the Single-Cell Level in a Mouse Model of Breast Cancer Metastasis to the Brain. *Magnetic Resonance in Medicine*. *56*: 1001-1010
- 70.** Luzzi, K. J., MacDonald, I. C., Schmidt, E. E., Kerkvliet, N., Morris, V. L., Chambers, A. F., Groom, A. C. (1998). Multistep nature of metastatic inefficiency: Dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *American Journal of Pathology*. *153*(3): 865 – 873

- 71.** Hunt, T., Naysmyth, K., Novák, B. (2011). Introduction: The cell cycle. *Philosophical Transactions: Biological Sciences*. 366(1584): 3494-3497
- 72.** Hustedt, N., Durocher, D. (2017). The control of DNA repair by the cell cycle. *Nature Cell Biology*. 19: 1-9
- 73.** Okhura, H. (2015). Meiosis: An overview of key differences from mitosis. 7(5): a015859
- 74.** Yaswen, P., MacKenzie, K. L., Keith, W. N., Hentosh, P., Rodier, F., Zhu, J., Firestone, G. L., Matheu, A., Carnero, A., Bilsland, A., Sundin, T., Honoki, K., Fujii, H., Georgakilas, A. G., Amedei, A., Amin, A., Helferich, B., Boosani, C. S., Guha, G., Ciriolo, M. S., Chen, S., Mohammed, S. I., Azmi, A. S., Bhakta, D. B., Halicka, D., Niccolia, E., Aguilano, K., Ashraf, S. S., Nowsheen, S., Yang, X. (2015). Therapeutic targeting of replicative immortality. *Seminars in Cancer Biology*. 35(Suppl): S104-S128
- 75.** Barnum, K. J., O'Connell, M. J. (2016). Cell cycle regulation by checkpoints. *Methods in Molecular Biology*. 1170: 29-40
- 76.** Wenzel, E. S., Singh, A. T. K. (2018). Cell-cycle checkpoints and aneuploidy on the path to cancer. *In Vivo*. 32(1): 1 – 5
- 77.** Harashima, H., Dissmeyer, N., Schnittger, A. (2013). Cell cycle control across the eukaryotic kingdom. *Trends in Cell Biology*. 23(7): 345 – 356
- 78.** Vermeulen, K., Van Bockstaele, D. R., Berneman, Z. N. (2003). The cell cycle: A review of regulation, deregulation and therapeutic targets in cancer. *Cell Proliferation*. 36(3)
- 79.** O'Grady, S., Lawless, M. W. (2015). 'Chapter 12 – Liver Cancer (Hepatocellular Carcinoma)' in Gray, S. G. ed. *Epigenetic Cancer Therapy*. Cambridge: Academic Press
- 80.** Visconti, R., Monica, R. D., Grieco, D. (2016). Cell cycle checkpoint in cancer: A therapeutically targetable double-edged sword. *Journal of Experimental and Clinical Cancer Research*. 35: 153
- 81.** Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K., Elledge, S. J. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*. 75(4): 805 – 816
- 82.** Dang, C. V. (2013). MYC on the path to cancer. *Cell*. 149(1): 22 – 35

- 83.** Tomura, M., Sakaue-Sawano, A., Mori, Y., Takase-Utsugi, M., Hata, A., Ohtawa, K., Kanagawa, O., Miyawaki, A. (2013). Contrasting quiescent G<sub>0</sub> phase with mitotic cell cycling in the mouse immune system. *PLoS One*. 8(9): e73801
- 84.** Gao, X., Zhang, M., Tang, Y., Liang, X. (2019). Cancer cell dormancy: mechanisms and implications of cancer recurrence and metastasis. *OncoTargets and Therapy*. 10: 5219 – 5228
- 85.** Zhang, J., Niu, C., Ye, L., Huang, H., He, X., Tong, W. G., Ross, J., Haug, J., Johnson, T., Feng, J. Q., Harris, S., Wiedemann, L. M., Mishina, Y., Li, L. (2003). Identification of the haematopoietic stem cell niche and control of the niche size. *Nature*. 425(6960): 836 – 841
- 86.** Oki, T., Nishimura, K., Kitaura, J., Togami, K., Maehara, A., Izawa, K., Sakaue-Sawano, A., Niida, A., Miyano, S., Aburatani, H., Kiyonari, H., Miyawaki, A., Kitamura, T. (2014). A novel cell-cycle-indicator, mVenus-p27K<sup>+</sup>, identifies quiescent cells and visualises G<sub>0</sub>-G<sub>1</sub> transition. *Scientific Reports*. 4: 4012
- 87.** Wheeler, S. E., Clark, A. M., Taylor, D. P., Young, C. L., Pillai, V. C., Stoiz, B., Venkataramanan, R., Lauffenburger, D., Griffith, L., Wells, A. (2014). Spontaneous dormancy of metastatic breast cancer cells in an all human liver microphysiologic system. *British Journal of Cancer*. 111: 2342-2350
- 88.** Kong, Y., Chen, J., Zhou, Z., Xia, H., Qiu, M., Chen, C. (2014). Cucurbitacin E induces cell cycle G<sub>2</sub>/M arrest and apoptosis in triple negative breast cancer. *PLoS ONE*. 9(7): e103760
- 89.** Zhao, T., Sun, Q., del Rincon, S. V., Lovato, A., Marques, M., Witcher, M. (2014). Gallotannin imposes S phase arrest in breast cancer cells and suppresses the growth of triple-negative tumours *in vivo*. *PLoS ONE*. 9(3): e92853
- 90.** Friedman, J. M., Fialkow, P. J. (1976). Cell marker studies of human tumorigenesis. *Transplantation Reviews*. 28: 17 – 33
- 91.** Balani, S., Nguyen, L. V., Eaves, C. J. (2017). Modeling the process of human tumorigenesis. *Nature Communications*. 8: 15244
- 92.** Nachman, M. W., Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*. 156(1): 297 – 304
- 93.** Hadrup, S., Donia, M., thor Straten, P. (2013). Effector CD4 and CD8 T cells and their role in the tumour microenvironment. *Cancer Microenvironment*. 6(2): 123 – 133

- 94.** Denton, A. E., Roberts, E. W., Fearon, D. T. (2018). Stromal cells in the tumour microenvironment. *Stromal Immunology*. 1060: 99 – 114
- 95.** Cazet, A. S., Hui, M. N., Elsworth, B. L., Wu, S. Z., Roden, D., Chan, C., Skhinas, J. N., Collot, R., Yang, J., Harvey, K., Johan, M. Z., Cooper, C., Nair, R., Herrmann, D., McFarland, A., Deng, N., Ruiz-Borrego, M., Rojo, F., Trigo, J. M., Bezares, S., Caballero, R., Lim, E., Timpson, P., O'Toole, S., Watkins, D. N., Cox, T. R., Samuel, M. S., Martin, M., Swarbrick, A. (2018). Targeting stromal remodelling and cancer stem cell plasticity overcomes chemoresistance in triple negative breast cancer. *Nature Communications*. 9: 2897
- 96.** Nallanthighal, S., Heiserman, J. P., Cheon, D. (2019). The role of the extracellular matrix in cancer stemness. *Frontiers in Cell and Developmental Biology*. 7:86
- 97.** Julich, D., Cobb, G., Melo, A. M., McMillen, P., Lawton, A. K., Mochrie, S. G. J., Rhoades, E., Holley, S. A. (2015). Cross-scale integrin regulation organises ECM and tissue topology. *Developmental Cell*. 34(1): 33 – 44
- 98.** Jang, M., Koh, I., Lee, J. E., Lim, J. Y., Cheong, J., Kim, P. (2018). Increased extracellular matrix density disrupts E-cadherin/ $\beta$ -catenin complex in gastric cancer cells. *Biomaterials Science*. 6: 2704 – 2713
- 99.** Yeh, Y., Ling, J., Chen, W., Lin, H., Tang, M. (2017). Mechanotransduction of matrix stiffness in regulation of focal adhesion size and number: Reciprocal regulation of caveolin-1 and  $\beta$ 1 integrin. *Scientific Reports*. 7: article 15008
- 100.** Kollmannsberger, P., Bidan, C. M., Dunlop, J. W. C., Fratzl, P., Vogel, V. (2018). Tensile forces drive a reversible fibroblast-to-myofibroblast transition during tissue growth in engineered clefts. *Science Advances*. 4: eaao4881
- 101.** Nam, J. M., Onodera, Y., Bissell, M. J., Park, C. C. (2010). Breast cancer cells in three-dimensional culture display an enhanced radioresponse after coordinate targeting of integrin  $\alpha$ 5 $\beta$ 1 and fibronectin. *Cancer Research*. 70(13): 5238 – 5248
- 102.** Quail, D. F., Joyce, J. A. (2018). The microenvironmental landscape of brain tumours. *Cancer Cell*. 31(3): 326 – 341
- 103.** Wang, L., Cossette, S. M., Rarick, K. R., Gershan, J., Dwinell, M. B., Harder, D. R., Ramchandran, R. (2013). Astrocytes directly influence tumour cell invasion and metastasis *in vivo*. *PLoS ONE*. 8(12): e80933

- 104.** Eble, J. A., Niland, S. (2019). The extracellular matrix in tumour progression and metastasis. *Clinical and Experimental Metastasis*. 36(3): 171 – 198
- 105.** Karamanos, N. K., Piperigkou, Z., Theocharis, A. D., Watanabe, H., Franchi, M., Baud, S., Brezillon, S., Gotte, M., Passi, A., Vigetti, D., Ricard-Blum, S., Sanderson, R. D., Neill, T., Iozzo, R. V. (2018). Proteoglycan and chemical diversity drives multifunctional cell regulation and therapeutics. *Chemical Reviews*. 118: 9152 – 9232
- 106.** Acerbi, I., Cassereau, L., Dean, I., Shi, Q., Au, A., Park, C., Chen, Y. Y., Liphardt, J., Hwang, E. S., Weaver, V. M. (2015). Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. *Integrative Biology*. 7(10): 1120 – 1134
- 107.** Drabsch, Y., ten Dijke, P. (2012). TGF- $\beta$  signalling and its role in cancer progression and metastasis. *Cancer and Metastasis Reviews*. 31(3-4): 553- 568
- 108.** Barnes, J. M., Przybyla, L., Weaver, V. M. (2017). Tissue mechanics regulate brain development, homeostasis and disease. *Journal of Cell Science*. 130: 71-82
- 109.** Discher, D. E., Janmey, P., Wang, Y. (2005). Tissue cells feel and respond to the stiffness of their substrate. *Science*. 310(5751): 1139-1143
- 110.** Sak, M. A., Littrup, P. J., Duric, N., Mullooly, M., Sherman, M. E., Gierach, G. L. (2015). Current and future methods for measuring breast density: A brief comparative review. *Breast Cancer Management*. 4(4): 209-221
- 111.** Conklin, M. W., Eickhoff, J. C., Riching, K. M., Pehlke, C. A., Eliceiri, K. W., Provenzano, P. P., Freidl, A., Keely, P. J. (2011). Aligned collagen is a prognostic signature for survival in human breast carcinoma. *American Journal of Pathology*. 178(3): 1221 – 1232
- 112.** Joyce, M. H., Lu, C., James, E. R., Hegab, R., Allen, S. C., Suggs, L. J., Brock, A. (2018). Phenotypic basis for matrix stiffness-dependent chemoresistance of breast cancer cells to doxorubicin. *Frontiers in Oncology*. 8: 337
- 113.** Schrader, J., Gordon-Walker, T. T., Aucott, R. L., Van Deemter, M., Quaas, S. W., Benten, D., Forbes, S. J., Wells, R. G., Iredale, J. P. (2011). Matrix stiffness modulates proliferation, chemotherapeutic response, and dormancy in hepatocellular carcinoma cells. *Hepatology*. 53(4): 1192 – 1205
- 114.** Chaudhuri, O., Koshy, S. T., Branco da Cunha, C., Shin, J., Verbeke, C. S., Allison, K. H., Mooney, J. (2014). Extracellular matrix stiffness and composition jointly regulate

the induction of malignant phenotypes in mammary epithelium. *Nature Materials*. 13: 970-978

**115.** Seong, J., Tajik, A., Guan, J., Humphries, M., Craig, S. E., Shekaran, A., Garcia, A., Lu, S., Lin, M. Z., Wang, N., Wang, Y. (2013). Distinct biophysical mechanisms of focal adhesion kinase mechanoactivation by different extracellular matrix proteins. *PNAS*. 110(48): 19372-19377

**116.** Wu, C. (2007). Focal adhesion: A focal point in current cell biology and molecular medicine. *Cell Adhesion and Migration*. 1(1): 13-18

**117.** Bae, Y. H., Mui, K. L., Hsu, B. Y., Liu, S., Cretu, A., Razinia, Z., Xu, T., Puré, E., Assoian, R. K. (2014). A FAK-cas-rac-lamellipodin signalling module transduces extracellular matrix stiffness into mechanical cell cycling. *Science Signaling*. 7(330): ra57

**118.** Liu, Y., Lv, J., Liang, X., Yin, X., Zhang, L., Chen, D., Jin, X., Fiskesund, R., Tang, K., Ma, J., Zhang, H., Dong, W., Mo, S., Zhang, T., Cheng, F., Zhou, Y., Xie, J., Wang, N., Huang, B. (2018). Fibrin stiffness mediates dormancy of tumour-repopulating cells via a cdc42-driven tet2 epigenetic program. *Cancer Research*. 78(14): 3926-3937

**119.** Ulrich, T. A., de Juan Pardo, E. M., Kumar. S. (2009). The mechanical rigidity of the extracellular matrix regulates the structure, motility, and proliferation of glioma cells. *Cancer Research*. 69(10): 4167-4174

**120.** Keeratichamoren, S., Lirdprapamonkol, K., Svasti, J. (2018). Mechanism of ECM-induced dormancy and chemoresistance in A549 lung carcinoma cells. *Oncology Reports*. 39(4): 1765 – 1774

**121.** Schaefer, L., Iozzo, R. V. (2008). Biological functions of the small leucine-rich proteoglycans: From genetics to signal transduction. *Journal of Biological Chemistry*. 283(31): 21305 – 21309

**122.** Costanza, B., Umelo, I. A., Bellier, J., Castronovo, V., Turtoi, A. Stromal modulators of TGF- $\beta$  in cancer. *Journal of Clinical Medicine*. 6(7)

**123.** Ren, D., Dai, Y., Yang, Q., Zhang, X., Guo, W., Ye, L., Huang, S., Chen, X., Lai, Y., Du, H., Lin, C., Peng, X., Song, L. (2019). Wnt5a induces and maintains prostate cancer cell dormancy in bone. *Journal of Experimental Medicine*. 216(2): 428 – 449

**124.** Liang, Y., Haring, M., Roughley, P. J., Margolis, R. K., Margolis, R. U. (1997). Glypican and biglycan in the nuclei of neurons and glioma cells: Presence of functional

nuclear localization signals and dynamic changes in glypical during the cell cycle. *Journal of Cell Biology*. 139(4): 851 – 864

**125.** Scott, P. G., Dodd, C. M., Bergmann, E. M., Sheehan, J. K., Bishop, P. N. (2006). Crystal structure of the biglycan dimer and evidence that dimerization is essential for folding and stability of class I small leucine-rich proteoglycans. *Journal of Biological Chemistry*. 281: 13324 – 13332

**126.** Nastase, M., Young, M. F., Schaefer, L. (2012). Biglycan: A multivalent proteoglycan providing structure and signals. *Journal of Histochemistry and Cytochemistry*. 60(12): 963 – 975

**127.** Kim, J., Lee, S. K., Shin, J., Jeoun, U., Jang, Y. J., Park, H. S., Kim, J., Gong, G., Lee, T., Hong, J., Lee, Y. J., Heo, Y. (2016). Enhanced biglycan gene expression in the adipose tissues of obese women and its association with obesity-related genes and metabolic parameters. *Scientific Reports*. 6: 30609

**128.** Hsieh, L. T., Nastase, M., Zeng-Brouwers, J., Iozzo, R., Schaefer, L. (2014). Soluble biglycan as a biomarker of inflammatory renal diseases. *International Journal of Biochemistry and Cellular Biology*. 0: 223 – 235

**129.** Hen, O., Barkan, D. (2019). Dormant disseminated tumour cells and cancer stem/progenitor-like cells: Similarities and opportunities. *Seminars in Cancer Biology*. [In Press]

**130.** Chambers, A. F., Groom, A. C., MacDonald, I. C. (2002). Dissemination and growth of cancer cells in metastatic sites. *Nature Reviews Cancer*. 2: 563-572

**131.** Prager, B. C., Xie, Q., Bao, S., Rich, J. N. (2019). Cancer stem cells: The architects of the tumour ecosystem. *Cell Stem Cell*. 24: 41-53

**132.** Seano, G. (2018). Targeting the perivascular niche in brain tumours. *Current Opinion in Oncology*. 30(1): 54-60

**133.** Ghajar, C. M., Peinado, H., Mori, H., Matei, I. R., Evason, K. J., Brazier, H., Almeida, D., Koller, A., Hajjar, K. A., Stainier, D. Y. R., Chen, E. I., Lyden, D., Bissell, M. J. (2013). The perivascular niche regulates breast tumour dormancy. *Nature Cell Biology*. 15(7): 807-817

**134.** Loriger, M., Felding-Habermann, B. (2010). Capturing changes in the brain microenvironment during the initial steps of breast cancer brain metastasis. *American Journal of Pathology*. 176(6): 2958 – 2971

- 135.** Yoneda, T., Williams, P. J, Hiraga, T., Niewolna, M., Nishimura, R. (2009). A bone-seeking clone exhibits different biological properties from the MDA-MB-231 parents human breast cancer cells and a brain-seeking clone in vivo and in vitro. *Journal of Bone and Mineral Research*. 16(8): 1486 – 1495
- 136.** Er, E. E., Valiente, M., Ganesh, K., Zou, Y., Agrawal, S., Hu, J., Griscom, B., Rosenblum, M., Boire, A., Brogi, E., Giancotti, F. G., Schnachner, M., Malladi, S., Massague, J. (2018). Pericyte-like spreading by disseminated cancer cells activates YAP and MRTF for metastatic colonization. *Nature Cell Biology*. 20(8): 966 – 978
- 137.** Baluk, P., Morikawa, S., Haskell, A., Mancuso, M., McDonald, D. M. (2003). Abnormalities of basement membrane on blood vessels and endothelial sprouts in tumours. *American Journal of Pathology*. 163(5): 1801 – 1815
- 138.** Pogany, G., Timar, F., Olah, J., Harisi, R., Polony, G., Paku, S., Bocsi, J., Jeney, A., Laurie, G. W. (2001). Role of the basement membrane in tumour cell dormancy and cytotoxic resistance. *Oncology*. 60(3): 274 – 281
- 139.** Yurchenco, P. D. (2011). Basement membranes: Cell scaffoldings and signalling platforms. *Cold Spring Harbor Perspectives in Biology*. 3(2): a004911
- 140.** Fujita, M., Khazenzon, N. M., Bose, S., Kiyotoshi, K., Sasaki, T., Carter, W. G., Ljubimov, A. V., Black, K. L. Ljubimova, J. Y. (2005). Overexpression of  $\beta$ 1-chain-containing laminins in capillary basement membranes of human breast cancer and its metastases. *Breast Cancer Research*. 7(4): R411 – R 421
- 141.** DiGiacomo, V., Meruelo, D. (2017). Looking into laminin receptor: critical discussion regarding the non-integrin 37/67-kDa laminin receptor/RPSA protein. *Biological Reviews of the Cambridge Philosophical Society*. 91(2): 288 – 310
- 142.** Domogatskata, A., Rodin, S., Tryggvason, K. (2012). Functional diversity of laminins. *Annual Review of Cell and Development Biology*. 28: 523 – 553
- 143.** Givant-Horwitz, V., Davidson, B., Reich, R. (2005). Laminin-induced signalling in tumor cells. *Cancer Letters*. 223(1): 1 – 10
- 144.** Schwartz, M. A., Assoian, R. K. (2001). Integrins and cell proliferation. *Journal of Cell Science*. 114: 2553 – 2560
- 145.** Ljubimova, J. Y., Fugita, M., Khaenzon, N. M., Das, A., Pikul, B. B., Newman, D., Seikiguchi, K., Sorokin, L. M., Sasaki, T., Black, M. D. (2004). Association between laminin-8 and glial tumour grade, recurrence, and patient survival. *Cancer*. 101(3): 604 – 612

- 146.** Ljubimova, J. Y., Lakhter, A. J., Loksh, A., Yong, W. H., Riedinger, M. S., Miner, J. H., Sorokin, L. M., Ljubimov, A. V., Black, K. L. (2001). Overexpression of  $\alpha 4$  chain-containing laminins in human glial tumours identified by gene microarray analysis. *Cancer Research*. 61: 5601 – 5610
- 147.** Albregues, J., Shields, M. A., Ng, D., Park, C. G., Ambrico, A., Poindexter, M. E., Upadhyay, P., Uyeminami, D. L., Pommier, A., Kuttner, V., Bruzas, E., Maiorino, L., Bautista, C., Carmona, E. M., Gimotty, P. A., Fearon, D. T., Chang, K., Lyons, S. K., Pinkerton, K. E., Trotman, L. C., Goldberg, M. S., Yeh, J. T., Egeblad, M. (2018). Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. *Science*. 361(6409): eaao4227
- 148.** Grimbert, P., Bouguermouh, S., Baba, N., Nakajima, T., Allakhverdi, Z., Braun, D., Saito, H., Rubio, M., Delespesse, G., Sarfati, M. (2006). Thrombospondin/CD47 interaction: A pathway to generate regulatory T cells from Human CD4+CD25- T cells in response to inflammation. *The Journal of Immunology*. 177(6): 3534 – 3541
- 149.** Ghajar, C. M., Peinado, H., Mori, H., Matei, I. R., Evason, K. J., Brazier, H., Almeida, D., Koller, A., Hajjar, K. A., Stainier, D. Y. R., Chen, E. I., Lyden, D., Bissell, M. J. (2014). The perivascular niche regulates breast tumour dormancy. *Nature Cell Biology*. 15(7): 807 – 817
- 150.** Valiente, M., Obenaus, A. C., Jin, X., Chen, Q., Zhang, X. H. -F., Lee, D. J., Chaft, J. E., Kris, M. G., Huse, J. T., Brogi, E., Massagué, J. (2014). Serpins promote cancer cell survival and vascular co-option in brain metastasis. *Cell*. 156: 1002-1016
- 151.** Heit, C., Jackson, B. C., McAndrews, M., Wright, M. W., Thompson, D. C., Silverman, G. A., Nebert, D. W., Vasiliou, V. (2013). Update of the human and mouse *SERPIN* gene superfamily. *Human Genomics*. 7:22
- 152.** Pantel, K., Schlimok, G., Braun, S., Kutter, D., Lindemann, F., Schaller, G., Funke, I., Izbicki, J. R., Riethmuller, G. (1993). Differential expression of proliferation-associated molecules in individual micrometastatic carcinoma cells. *Journal of the National Cancer Institute*. 85(17): 1419 – 1424
- 153.** Hambardzumyan, D., Becher, O. J., Rosenblum, M. K., Pandolfi, P. P., Manova-Todorova, K., Holland, E. C. (2008). PI3K pathway regulates survival of cancer stem cells residing in the perivascular niche following radiation in medulloblastoma in vivo. *Genes and Development*. 22: 436 – 448

- 154.** Eales, K. L., Hollinshead, K. E. R., Tennant, D. A. (2016). Hypoxia and metabolic adaptation of cancer cells. *Oncogenesis*. 5: e190
- 155.** Endo, H., Inoue, M. (2019). Dormancy in cancer. *Cancer Science*. 110(2): 474 – 480
- 156.** Pouyssegur, J., Frédéric, D., Mazure, N. M. (2006). Hypoxia signalling in cancer approaches to enforce tumour regression. *Nature*. 441: 437 – 443
- 157.** Wang, G. L., Semenza, G. L. (1995). Purification and characterization of hypoxia-inducible factor 1. *Journal of Biological Chemistry*. 270(3): 1230 – 1237
- 158.** Butturini, E., Carcereri de Prati, A., Boriero, D., Mariotto, S. (2019). Tumor dormancy and interplay with hypoxic tumour microenvironment. *International Journal of Medical Sciences*. 20(17): 4305
- 159.** Semenza, G. L. (2003). Targeting HIF-1 for cancer therapy. *Nature Reviews Cancer*. 3: 721 – 732
- 160.** Bachelder, R. E., Crago, A., Chung, J., Wendt, M. A., Shaw, L. M., Robinson, G., Mercurio, A. M. (2001). Vascular endothelial growth factor is an autocrine survival factor for neuropilin-expressing breast adenocarcinoma cells. *Cancer Research*. 61 5736 – 5740
- 161.** Chen, X., Liu, J., He, B., Li, Y., Liu, S., Wu, B., Wang, S., Zhang, S., Xu, X., Wang, J. (2015). Vascular endothelial growth factor (VEGF) regulation by hypoxia inducible factor-1 alpha (HIF1A) starts and peaks during endometrial breakdown, not repair, in a mouse menstrual-like model. *Human Reproduction*. 30(9): 2160 – 2170
- 162.** Carcereri de Prati, A., Butturini, E., Rigo, A., Oppici, E., Rossin, M., Boriero, D., Mariotto, S. (2017). Metastatic breast cancer cells enter into dormant state and express cancer stem cells phenotype under chronic hypoxia. *Journal of Cellular Biochemistry*. 118(10)
- 163.** Fluegen, G., Avivar-Valderas, A., Wang, Y., Padgen, M. R., Williams, J. K., Nobre, A. R., Calvo, V., Cheung, J. F., Bravo-Coredero, J. J., Entenberg, D., Castracane, J., Verkhusha, V., Keely, P. J., Condeelis, J., Aguirre-Ghiso, J. A. (2017). Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic environments. *Nature Cell Biology*. 19(2): 120 – 132
- 164.** Gilkes, D. M., Semenza, G. L., Wirtz, D. (2014). Hypoxia and the extracellular matrix: drivers of tumour metastasis. *Nature Reviews Cancer*. 14(6): 430 – 439

- 165.** Zhan, T., Rindtorff, N., Boutros, M. (2017). Wnt signalling in cancer. *Oncogene*. 36: 1461 – 1473
- 166.** Koval, A., Kataneav, V. L. (2018). Dramatic disbalancing of the Wnt pathway in breast cancers. *Scientific Reports*. 8: 7329
- 167.** Katoh, M. (2017). Canonical and non-canonical WNT signalling in cancer stem cells and their niches: Cellular heterogeneity, omics reprogramming, targeted therapy and tumour plasticity. *International Journal of Oncology*. 51(5): 1357 – 1369
- 168.** Zhou, Y., Su, Y., Zhu, H., Wang, X., Li, X., Dai, C., Xu, C., Zheng, T., Mao, C., Chen, D. (2019). Interleukin-23 receptor signalling mediates cancer dormancy and radioresistance in human esophageal squamous carcinoma cells via the Wnt/Notch pathway. *Journal of Molecular Medicine*. 97(2): 177 – 188
- 169.** Holdman, X. B., Welte, T., Rajapakshe, K., Pond, A., Coarfa, C., Mo., Q., Huang, S., Hilsenbeck, S. G., Edwards, D. P., Zhang, X., Rosen, J. M. (2015). Upregulation of EGFR signalling is correlated with tumour stroma remodelling and tumour recurrence in FGFR1-driven breast cancer. *Breast Cancer Research*. 17(141)
- 170.** Guagnano, V., Furet, P., Spanka, C., Bordas, V., Le Douget, M., Stamm, C., Brueggen, J., Jensen, M. R., Schnell, C., Schmid, H., Wartmann, H., Berghausen, J., DruECKes, P., Zimmerlin, A., Bussiere, D., Murray, J., Graus Porta, D. (2011). Discovery of 3-(2,6-dichloro-3,5-deimethoxy-phenyl)-1{6-[4-(4-ethyl-piperazin-1-yl)-phenylamino]-pyrimidin-4-yl}-1-methyl-urea (NVP-BGJ398), a potent and selective inhibitor of the fibroblast growth factor receptor family of receptor tyrosine kinase. *Journal of Medicinal Chemistry*. 54(20): 7066 – 7083
- 171.** Mentink, R. A., Rella, L., Radaszkiewicz, T. W., Gybel, T., Betist, M. C., Bryja, V. (2018). The planar cell polarity protein VANG-1/Vangl negatively regulates Wnt/  $\beta$ -catenin signalling through a Dvl dependent mechanism. *PLoS Genetics*. 14(12): e1007840
- 172.** Ren, D., Dai, Y., Yang, Q., Zhang, X., Guo, W., Ye, L., Huang, S., Chen, X., Lai, Y., Du, H., Lin, C., Peng, X., Song, L. (2019). Wnt5a induces and maintains prostate cancer cells dormancy in bone. *Journal of Experimental Medicine*. 216(2): 428 – 449
- 173.** Malladi, S., Macalinao, D. G., Jin, X., He, L., Basnet, H., Zou, Y., de Stanchina, E., Massague, J. (2016). Metastatic latency and immune evasion through autocrine inhibition of WNT. *Cell*. 165(1): 45 – 60

- 174.** Thorpe, L. M., Yuzugullu, H., Zhao, J. J. (2015). PI3K in cancer: Divergent roles of isoforms, modes of activation and therapeutic targeting. *Nature Reviews Cancer*. 15(1): 7 – 24
- 175.** Baselga, J. (2011). Targeting the phosphoinositide-3 (PI3) kinase pathway in breast cancer. *The Oncologist*. 16(1): 12 – 19
- 176.** Martini, M., Chiara De Santis, M., Braccini, L., Gulluni, F., Hirsch, E. (2014). PI3K/AKT Signaling Pathway and Cancer: an Updated Review. *Annals of Medicine*. 46: 372-383
- 177.** Yu, J. S., Cui, W. (2016). Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. *Development*. 143(17): 3050 – 3060
- 178.** Yang, L., He, C., Chen, X., Su, L., Liu, B., Zhang, H. (2016). Aurora kinase A revives dormant laryngeal squamous cell carcinoma cells via FAK/PI3K/Akt pathway activation. *Oncotarget*. 7(30): 48346 – 48359
- 179.** Humtsoe, J. O., Kramer, R. H. (2010). Differential epidermal growth factor receptor signaling regulates anchorage-independent growth by modulation of the PI3K/AKT pathway. *Oncogene*. 29: 1214 – 1226
- 180.** Vaziri-Gohar, A., Zheng, Y., Houston, K. D. (2017). IGF-1 receptor modulates FoxO1-mediated tamoxifen response in breast cancer cells. *Molecular Cancer Research*. 15(4): 489 497
- 181.** Zheng, W., Quirion, R. (2006). Insulin-like growth factor-1 (IGF-1) induces the activation/phosphorylation of Akt kinase and cAMP response element-binding protein (CREB) by activating different signalling pathways in PC12 cells. *BMC Neuroscience*. 7: 51
- 182.** Christopoulos, P. F., Msaoul, P., Koutsilieris, M. (2015). The role of the insulin-like growth factor-1 system in breast cancer. *Molecular Cancer*. 14: 43
- 183.** Jo, H., Jia, Y., Subramanian, K. K., Hattori, H. Luo, H. R. (2008). Cancer cell-derived clusterin modulates the phosphatidylinositol 3'-kinase-Akt pathway through attenuation of insulin-like growth factor 1 during serum deprivation. *Molecular and Cellular Biology*. 28(13): 4285 – 4299
- 184.** Shi, Y., Massague, J. (2003). Mechanisms of TGF- $\beta$  signalling from cell membrane to the nucleus. *Cell*. 113: 685 – 700

- 185.** Padua, D., Zhang, X. H., Wang, Q., Nadal, C., Gerald, W. L., Gomis, R. R., Massague, J. (2008). TGFbeta primes breast tumours for lung metastases seeding through angiopoietin-like 4. *Cell*. 133(1): 66 – 77
- 186.** Gong, X., Hou, Z., Endsley, M. P., Gronseth, E. I., Rarick, K. R., Jorns, J. M., Yang, Q., Du, Z., Yan, K., Bordas, M. L., Gershan, J., Deepak, P., Geethadevi, A., Chaluvally-Raghavan, P., Fan, Y., Harder, D. R., Ramchandran, R., Wang, L. (2019). Interaction of tumor cells and astrocytes promotes breast cancer brain metastases through TGF- $\beta$ 2/ANGPTL4 axes. *npj Precision Oncology*. 3: 24
- 187.** Datto, M. B., Li, Y., Panus, J. F., Howe, D. J., Xiong, Y., Wang, X. (1995). Transforming growth factor  $\beta$  induces the cyclin-dependent kinase inhibitor p21 through a p53-independent mechanism. *Biochemistry*. 92: 5545 – 5549
- 188.** Polyak, K., Kato, J. Y., Solomon, M. J., Sherr, C. J., Massague, J., Roberts, J. M., Koff, A. (1994). p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. *Genes and Development*. 8: 9 – 22
- 189.** Hannon, G. J., Beach, D. (1994). P15<sup>INK4B</sup> is a potential effector of TGF- $\beta$ -induced cell cycle arrest. *Nature*. 371: 257 – 261
- 190.** Kobayashi, A., Okuda, H, Xing, F., Pandey, P. R., Watable, M., Hirota, S., Pai, S. K., Liu, W., Fukuda, K., Chambers, C., Wilber, A., Watabe, K. (2011). Bone morphogenic protein 7 in dormancy and metastasis of prostate cancer stem-like cells in bone. *Journal of Experimental Medicine*. 208(13): 2641
- 191.** Bragado, P., Estrada, Y., Parikh, F., Krause, S., Capobianco, C., Farina, H. G., Schewe, D. M., Aguirre-Ghiso, J. A. (2013). TGF- $\beta$ 2 dictates disseminated tumour cell fate in target organs through TGF- $\beta$ -RIII and p38  $\alpha/\beta$  signalling. *Nature Cell Biology*. 15: 1351 – 1361
- 192.** Aguirre-Ghiso, J. A., Estrada, Y., Liu, D., Ossowski, L. (2003). ERK(MAPK) activity as a determinant of tumour growth and dormancy; regulation by p38(SAPK). *Cancer Research*. 63(7): 1684 – 1695
- 193.** Chen, S. X., Zhao, F., Huang, X. J. (2018). MAPK signalling pathway and erectile dysfunction. *National Journal of Andrology*. 24(5): 442 – 446
- 194.** Zhang, W., Liu, H.T. (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research*. 12: 9 – 18

- 195.** Chatterjee, M., Van, G. K. L. (2011). Dissecting the cellular mechanism of induction and sustenance of dormancy by Rho GTPases in breast cancer. *Cancer Research*. 71(24)
- 196.** Adam, A. P., George, A., Schewe, D., Bragado, P., Iglesias, B. V., Ranganathan, A. C., Kourtidis, A., Conklin, D. S., Aguirre-Ghiso, J. A. (2009). Computational identification of a p38SAPK-regulated transcription factor network for tumour cell quiescence. *Cancer Research*. 69(14): 5664 – 5672
- 197.** Marshall, J. A., Collins, J. W., Nakayama, J., Horak, C. E., Liewehr, D. J., Steinberg, S. M., Albaugh, M., Vidal-Vanaclocha, F., Palmieri, D., Barbier, M., Murone, M., Steeg, P. S. (2012). Effect of inhibition of the lysophosphatidic acid receptor 1 on metastasis and metastatic dormancy in breast cancer. *Journal of the National Cancer Institute*. 104(17): 1306 – 1319
- 198.** Morrison, D. K. (2012). MAP kinase pathways. *Cold Spring Harbor Perspectives in Biology*. 4: a011254
- 199.** Schewe, D. M., Aguirre-Ghiso, J. A. (2008). ATF6 $\alpha$ -Rheb-mTOR signalling promotes survival of dormant tumour cells *in vivo*. *PNAS*. 105(30): 10519 – 10524
- 200.** Aguirre-Ghiso, J. A., Kovalski, K., Ossowski, L. (1999). Tumor dormancy induced by downregulation of urokinase receptor in human carcinoma involves integrin and MAPK signalling. *Journal of Cell Biology*. 147(1): 89 – 104
- 201.** Harvey, K. F., Zhang, X., Thomas, D. M. (2013). The Hippo pathway and human cancer. *Nature Reviews Cancer*. 13: 246 – 257
- 202.** Yu, F., Zhao, B., Guan, K. (2015). Hippo pathway in organ size control, tissue homeostasis, and cancer. *Cell*. 163(4): 811 – 828
- 203.** Touil, Y., Igoudjil, W., Corvaisier, M., Dessen, A. F., Vandomme, J., Monte, D., Stechly, L., Skrypek, N., Langlois, C., Grard, G., Millet, G., Leteurtre, E., Dumont, P., Truant, S., Pruvot, F. R., Hebbar, M., Fan, F., Ellis, L. M., Formstecher, P., Van Seuning, I., Gespach, C., Polakowska, R., Huet, G. (2014). Colon cancer cells escape 5FU chemotherapy-induced cell death by entering stemness and quiescence associated with the c-Yes/YAP axis. *Clinical Cancer Research*. 20(4): 837 – 846
- 204.** Guo, L., Chen, Y., Luo, J., Zheng, J., Shao, G. (2019). YAP1 overexpression is associated with poor prognosis of breast cancer patients and induces breast cancer cell growth by inhibiting PTEN. *FEBS Open Bio*. 9(3): 437 – 445

- 205.** Cao, J., Zhao, X., Wang, D., Chen, K., Sheng, X., Li, W., Li, M., Liu, W., He, J. (2014). YAP is overexpressed in clear renal cell carcinoma and its knockdown reduces cell proliferation and induces cell cycle arrest and apoptosis. *Oncology Reports*. 32: 1594 – 1600
- 206.** Zhang, X., Li, Y., Ma, Y., Yang, L., Wang, T., Meng, T., Zong, Z., Sun, X., Hua, X., Li, H. (2018). Yes-associated protein (YAP) binds to HIF-1 $\alpha$  and sustains HIF-1 $\alpha$  protein stability to promote hepatocellular carcinoma cell glycolysis under hypoxic stress. *Journal of Experimental and Clinical Cancer Research*. 37(1): 216
- 207.** Qin, Z., Xia, W., Fisher, G. J., Voorhees, J. J., Quan, T. (2018). YAP/TAZ regulates TGF- $\beta$ /Smad3 signaling by induction of Smad7 via AP-1 in human skin dermal fibroblasts. *Cell Communication and Signalling*. 16(1): 18
- 208.** Zhao, Y., Montminy, T., Azad, T., Lightbody, E., Hao, Y., SenGupta, S., Asselin, E., Nicol, C., Yang, X. (2018). PI3K positively regulates YAP and TAZ in mammary tumorigenesis through multiple signalling pathways. *Molecular Cancer Research*. 16(6): 1046 – 1058
- 209.** Elbediwy, A., Vincent-Mistiaen, Z. I., Spencer-Dane, B., Stone, R. K., Boeing, S., Wculek, S. K., Cordero, J., Tan, E. H., Ridgway, R., Brunton, V. G., Sahai, E., Gerhardt, H., Behrens, A., Malanchi, I., Sansom, O. J., Thompson. (2016). Integrin signalling regulates YAP and TAZ to control skin homeostasis. *Development*. 143(10): 1674 – 1687
- 210.** Nardone, G., Oliver-De la Cruz, J., Vrbsky, J., Martini, C., Pribyl, J., Skladal, P., Pesl, M., Caluori, G., Pagliari, S., Martino, F., Macexkova, Z., Hajduch, M., Sanz-Garcia, A., Pugno, N. M., Stokin, G. B., Forte, G. (2017). YAP regulates cell mechanics by controlling focal adhesion assembly. *Nature Communications*. 8: 15321
- 211.** Juan, W. C., Hong, W. (2016). Targeting the Hippo signalling pathway for tissue regeneration and cancer therapy. *Genes*. 7(9)
- 212.** Fletcher, G. C., Elbediwy, A., Khanal, I., Ribeiro, P. S., Tapon, N., Thompson, B. J. (2015). The spectrin cytoskeleton regulates the Hippo signalling pathway. *EMBO*. 34(7): 940 – 954
- 213.** Li, X., Gu, J. Zhou, Q. (2015). Review of aerobic glycolysis and its key enzymes – new targets for lung cancer therapy. *Thoracic Cancer*. 6(1): 17 – 24
- 214.** Wilson, D. F. (2017). Oxidative phosphorylation: Regulation and role in cellular and tissue metabolism. *Journal of Physiology*. 595(23): 7023 – 7038

- 215.** Ciminera, A. K., Jandial, R., Termini, J. (2017). Metabolic advantages and vulnerabilities in brain metastases. *Clinical and Experimental Metastasis*. 34(6 – 7): 401 – 410
- 216.** Endo, H., Okuyama, H., Ohue, M., Inoue, M. (2014). Dormancy of cancer cells with suppression of AKT activity contributes to survival in chronic hypoxia. *PLoS One*. 9(6): e98858
- 217.** Swietach, P., Vaughan-Jones, R. D., Harris, A. L., Hulikova, A. (2014). The chemistry, physiology and pathology of pH in cancer. *Philosophical Transactions B*. 369(1638): 20130099
- 218.** Sharma, M., Astekar, M., Soi, S., Manjunatha, B. S., Shetty, D. C., Radhakrishnan, R. (2015). pH gradient reversal: An emerging hallmark of cancers. *Recent Patents on Anti-Cancer Drug Discovery*. 10(3): 244 – 258
- 219.** HiberCell Inc. (2019). *Our Science*. Available at: <https://www.hibercell.com/our-science/> [Accessed 11 September 2019]
- 220.** Touny, L. H., Vieira, A., Mendoza, A., Khanna, C., Hoenerhoff, M. J., Green, J. E. (2013). Combined SFK/MEK inhibition prevents metastatic outgrowth of dormant tumour cells. *The Journal of Clinical Investigation*. 124(1): 156 – 168
- 221.** Brastianos, P. K., Carter, S. L., Santagata, S., Cahill, D. P., Taylor-Weiner, A., Jones, R. T., Van Allen, E. M., Lawrence, M. S., Horowitz, P. M., Cibulskis, K., Ligon, K. L., Taberner, J., Seoane, J., Martinez-Saez, E., Curry, W. T., Dunn, I. F., Paek, S. H., Park, S. H., McKenna, A., Chevalier, A., Rosenberg, M., Barker, F. G., Gill, C. M., Van Hummelen, P., Thorner, A. R., Johnson, B. E., Hoang, M. P., Choueiri, T. K., Signoretti, S., Sougnez, C., Rabin, M. S., Lin, N. U., Winer, E. P., Stemmer-Rachamimov, A., Meyerson, M., Garraway, L., Gabriel, S., Lander, E. S., Beroukhi, R., Batchelor, T. T., Baselga, J., Louis, D. N., Getz, G., Hahn, W. C. (2015). Genomic characterization of brain metastases reveals branched evolution and potential therapeutic targets. *Cancer Discovery*. 5(11): 1164 – 1177
- 222.** National Institute of Health: U.S. National Library of Medicine. (2019). *ClinicalTrials.gov*. Available at: <https://clinicaltrials.gov/ct2/search/advanced> [Accessed 12 December 2019]
- 223.** Workman, P., Aboagye, E. O., Balkwill, F., Balmain, A., Bruder, G., Chaplin, D. J., Double, J. A., Everitt, J., Farningham, D. A. H., Glennie, M. J., Kelland, L. R., Robinson, V., Stratford, I. J., Tozer, G. M., Watson, S., Wedge, S. R., Eccles, S. A., An ad hoc committee

of the National Cancer Institute. (2010). Guidelines for the welfare and use of animals in cancer research. *British Journal of Cancer*. 102: 1555 – 1577

**224.** Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S., Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols*. 9: 171 – 181

**225.** Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. [Last accessed 02 September 2019]

**226.** Bolger, A.M., Lohse, M., Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*. 30(15): 2114-2120

**227.** Martin, M. (2011). Cutadapt removes adaptor sequences from high-throughput sequencing reads. *EMBnet.journal*. 17(1): 10-12

**228.** Karolchik, D., Hinrichs, A. S., Furey, T. S., Roskin, K. M., Sugnet, C. W., Haussler, D., Kent, W. J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Research*. 1(32): D493-496

**229.** Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Carbonell Sala, S., Chrast, J., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I.T., García Girón, C., Gonzalez, J.M., Grego, T., Hardy, M., Hourlier, T., Hunt, T., Izuogu, O.G., Lagarde, J., Martin, F.J., Martínez, L., Mohanan, S., Muir, P., Navarro, F.C.P., Parker, A., Pei, B., Pozo, F., Ruffier, M., Schmitt, B.M., Stapleton, E., Suner, M.M., Sycheva, I., Uszczynska-Ratajczak, B., Xu, J., Yates, A., Zerbino, D., Zhang, Y., Aken, B., Choudhary, J.S., Gerstein, M., Guigó, R., Hubbard, T.J.P., Kellis, M., Paten, B., Reymond, A., Tress, M.L., Flicek, P. (2018). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research*. 47(D1): D766-D773

**230.** Pages, H., Aboyoun, P., Gentleman, R., DebRoy, S. (2019). *Biostrings: Efficient manipulation of biological strings*. R package version 2.54.0

**231.** Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 29(1): 15-21

- 232.** Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*. 25(16): 2078-2079
- 233.** Okonechnikov, K., Conesa, A., Garcia-Alcalde, F. (2012). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*. 32(2): 292-294
- 234.** Wysoker A., Tibbetts K., Fennell T. (2013). *Picard Tools Version 1.90*. Available at: <http://picard.sourceforge.net> [Last accessed 17 December 2019]
- 235.** Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., Mesirov, J. P. (2011). Integrative Genomics Viewer. *Nature Biotechnology*. 29: 24-26
- 236.** Trapnell, C., Williams, B., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, J., Salzberg, S., Wold, B., Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*. 28(5): 511-515
- 237.** Liao, Y., Smyth, G. K., Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*. 47(8): e47
- 238.** Love, M. I., Huber, W., Anders, S. (2014). Moderated estimation of fold change and dispersion of RNAseq data with DESeq2. *Genome Biology*. 15(550)
- 239.** Metsalu, T., Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principle Component Analysis and heatmap. *Nucleic Acids Research*. W1: W566-570
- 240.** Saurin, A. (2019). *Genomics Biotoools*. [online] Available at: <https://www.biotoools.fr/> [Last accessed 01 August 2019]
- 241.** Tate, J. G., Bamford, S., Jubb, H. C, Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S. C., Kok, C. Y., Noble, K., Ponting, L., Ramshaw, C. C., Rye, C. E., Speedy, H. E., Stefancsik, R., Thompson, S. L., Wang, S., Ward, S., Campbell, P. J., Forbes, S. A. (2018). COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research*. 47(D1): D941-D947

- 242.** Yu, G., Wang, L. G., Han, Y., He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*. 16(5): 284 – 287
- 243.** Yu, G., He, Q. (2016). ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualisation. *Molecular BioSystems*. 12(12): 477 – 479
- 244.** Essaghir, A., Toffalini, F., Knoop, L., Kallin, A., van Helden, J., Demoulin, J. B. (2010). Transcription factor regulation can be accurately predicted from the presence of target gene signatures in microarray data. *Nucleic Acids Research*. 38(11): e120
- 245.** Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., von Mering, C. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*. 47: D607 – D613
- 246.** Lun, A. T. L., McCarthy, D. J., Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research*. 5: 2122
- 247.** Varešlija, D., Priedigkeit, N., Fagan, A., Purcell, S., Cosgrove, N., O'Halloran, P. J., Cocchiglia, S., Hartmaier, R., Castro, C. A., Zhu, L., Tseng, G. C., Lucas, P. C., Puhalla, S. L., Brufsky, A. M., Hamilton, R. L., Mathew, A., Leone, J. P., Basudan, A., Hudson, L., Dwyer, R., Das, S., O'Connor, D. P., Buckley, P. G., Farrell, M., Hill, A. D. K., Oesterreich, S., Lee, A. V., Young, L. S. (2019). Transcriptome characterization of matched primary breast and brain metastatic tumors to detect novel actionable targets. *Journal of the National Cancer Institute*. 111(4): 388 – 398
- 248.** Gautier, L., Cope, L., Bolstad, B. M., Irizarry, R. A. (2004). affy – analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 20(3): 307 – 315
- 249.** Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M., Look, M. P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M. E., Yu, J., Jatkoa, T., Berms, E. M., Atkins, D., Foekens, J. A. (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet*. 365(9460): 671 – 679
- 250.** Minn, A. J., Gupta, G. P., Padua, D., Bos, P., Nguyen, D. X., Nuyten, D., Kreike, B., Zhang, Y., Wang, Y., Ishwaran, H., Foekens, J. A., van de Vijver, M., Massague, J. (2007).

Lung metastasis genes couple breast tumour size and metastatic spread. *PNAS*. 104(16): 6740 – 6745

**251.** Bos, P. D., Zhang, X. H., Nadal, C., Shu, W., Gomis, R. R., Nguyen, D. X., Minn, A. J., van de Vijver, M. J., Gerald, W. L., Foekens, J. A., Massague, J. (2009). Genes that mediate breast cancer metastasis to the brain. *Nature*. 459(7249): 1005 – 1009

**252.** Zhang, X. H., Wang, Q., Gerald, W., Hudis, C. A., Norton, L., Smid, M., Foekens, J. A., Massague, J. (2009). Latent bone metastasis in breast cancer tied to Src-dependent survival signals. *Cancer Cell*. 16(1): 67 – 78

**253.** Xu, J., Acharya, S., Sahin, O., Zhang, Q., Saito, Y., Yao, J., Wang, H., Li, P., Zhang, L., Lowery, F. J., Kuo, W. L., Xiao, Y., Ensor, J., Sahin, A. A., Zhang, X. H., Hung, M. C., Zhang, J. D., Yu, D. (2015). 14-3-3 $\zeta$  turns TGF- $\beta$ 's function from tumor suppressor to metastasis promoter in breast cancer by contextual changes of Smad partners from p53 to Gli2. *Cancer Cell*. 27(2): 177 – 192

**254.** McMullin, R. P., Wittner, B. S., Yang, C., Denton-Schneider, B. R., Hicks, D., Singavarapu, R., Moulis, S., Lee, J., Akbari, M. R., Narod, S. A., Aldape, K. D., Steeg, P. S., Ramaswamy, S., Sgroi, D. C. (2014). A BRCA1 deficient-like signature is enriched in breast cancer brain metastases and predicts DNA damage-induced poly (ADP-ribose) polymerase inhibitor sensitivity. *Breast Cancer Research*. 16(2): R25

**255.** Nagy, A., Lanczky, A., Meynhart, O., Gyorffy, B. (2018). Validation of miRNA as a prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific Reports*. 8: 9227

**256.** Schindelin, J., Arganda-Carreras, I., Frise, E. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 9(7): 676-682

**257.** Lois, C., Hong, E. J., Pease, S., Brown, E. J., Baltimore, D. (2002). Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science*. 295(5556): 868 – 872

**258.** Markusic, D., Oude-Elferink, R., Das, A., T., Berkhout, B., Seppen, J. (2005). Comparison of single regulated lentiviral vectors with rtTA expression driven by an autoregulatory loop or a constitutive promotor. *Nucleic Acids Research*. 33(6): e63

**259.** Guan, B., Wang, T. L., Shih, M. (2011). ARID1A, a factor that promotes a SWI/SNF-mediated chromatin remodeling, is a tumour suppressor in gynecologic cancers. *Cancer Research*. 71(21): 6718 - 6727

- 260.** Li, M., Husic, N., Lin, Y., Snider, B. J. (2012). Production of lentiviral vectors for transducing cells from the central nervous system. *Journal of Visualized Experiments*. 63: e4031
- 261.** Bantly, A. D., Gray, B. D., Breslin, E., Weinstein, E. G., Muirhead, K. A., Ohlsson-Wilhelm, B. M., Moore, J. S. (2007). CellVue® Claret, a new far-red dye, facilitates polychromatic assessment of immune cell proliferation. *Immunological Investigations*. 36(5-6): 581 – 605
- 262.** Ching, T., Huang, S., Garmire, L. X. (2014). Power analysis and sample size estimation for RNA-Seq differential expression. *RNA*. 20(11): 1684 – 1696
- 263.** Liu, Y., Zhou, J., White, K. P. (2014). RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics*. 30(3): 301 – 304
- 264.** Pasaniuc, B., Zaitlen, N., Halperin, E. (2011). Accurate estimation of expression levels of homologous genes in RNA-seq experiments. *Journal of Computational Biology*. 18(3)
- 265.** Singhal, S.K., Usmani, N., Michiels, S., Metzger-Filho, O., Saini, K.S., Kovalchuk, O., Parliament, M. (2015). Towards understanding the breast cancer epigenome: a comparison of genome-wide DNA methylation and gene expression data. *Oncotarget*. 7(3): 3002-3017
- 266.** Kanehisa, M., Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. 28: 27-30
- 267.** Simillion, C., Liechti, R., Lischer, H. E. L., Ioannidis, V., Bruggmann, R. (2017). Avoiding the pitfalls of gene set enrichment analysis with SetRank. *BMC Bioinformatics*. 18: 151
- 268.** Li, L., Liang, Y., Kang, L., Liu, Y., Gao, S., Chen, S., Li, Y., You, W., Dong, Q., Hong, T., Yan, Z., Jin, S., Wang, T., Zhao, W., Mai, H., Huang, J., Han, X., Ji, Q., Song, Q., Yang, C., Zhao, S., Xu, X. Ye, Q. (2018). Transcriptional regulation of the Warburg effect in cancer by SIX1. *Cancer Cell*. 33: 368 – 385
- 269.** Tario Jr, J. D., Conway, A. N., Muirhead, K. A., Wallace, P. K. (2018). Monitoring cell proliferation by dye dilution: considerations for probe selection. *Methods in Molecular Biology*. 1678: 249 – 299

- 270.** Miller, I., Min, M., Yang, C., Tian, C., Gookin, S., Carter, D., Spencer, S. L. (2018). Ki67 is a graded rather than a binary marker of proliferation versus quiescence. *Cell Reports*. 24(5): 1105 – 1112
- 271.** Jurikova, M., Danihel, L., Polak, S., Varga, I. (2016). Ki67, PCNA, and MCM proteins: Markers of proliferation in the diagnosis of breast cancer. *Acta Histochemica*. 118(5): 544 – 552
- 272.** Vieth, B., Parekh, S., Ziegenhain, C., Enard, W., Hellmann, I. (2019). A systematic evaluation of single cell RNA-seq analysis pipelines. *Nature Communications*. 10(4667)
- 273.** Bradford, J. R., Farren, M., Powell, S. J., Runswick, S., Weston, S. L., Brown, H., Delpuech, O., Wappett, M., Smith, N. R., Carr, T. H., Dry, J., Gibson, N. J., Barry, S. T. (2013). RNA-Seq differentiates tumour and host mRNA expression changes induced by treatment of human tumour xenografts with the VEGFR tyrosine kinase inhibitor cediranib. *PLoS One*. 8(6): e66003
- 274.** Barkan, D., Green, J. E., Chambers, A. F. (2010). Extracellular matrix: A gatekeeper in the transition from dormancy to metastatic growth. *European Journal of Cancer*. 46(7): 1181 – 1188
- 275.** Almog, N., Ma, L., Raychowdhury, R., Schwager, C., Erber, R., Short, S., Hlatky, L., Vajkoczy, P., Huber, P. E., Folkman, J., Abdollahi, A. (2009). Transcriptional switch of dormant tumors to fast-growing angiogenic phenotype. *Cancer Research*. 69(3): 836 – 844
- 276.** Kim, R. S., Avivar-Valderas, A., Estrada, Y., Bragado, P., Sosa, M. S., Aguirre-Ghiso, J. A., Segall, J. E. (2012). Dormancy signatures and metastasis in estrogen receptor positive and negative breast cancer. *PLoS One*. 7(4): e35569
- 277.** Pearson, H.B., McGlinn, E., Phesse, T.J., Schluter, H., Srikumar, A., Godde, N.J., Woelwer, C.B., Ryan, A., Phillips, W.A., Matthias, E., Kaur, P., Humbert, P. (2015). The polarity protein Scrib mediates epidermal development and exerts a tumor suppressive function during skin carcinogenesis. *Molecular Cancer*. 14(169)
- 278.** Johnson, R., Halder, G. (2014). The two faces of Hippo: targeting the Hippo pathway for regenerative medicine and cancer treatment. *Nat Rev Drug Discov*. 13(1): 63-79
- 279.** Guo, L., Teng, L. (2015). YAP/TAZ for cancer therapy: opportunities and challenges (review). *Int J Oncol*. 46(4): 1444-52

- 280.** Oku, Y., Nishiya, N., Shito, T., Yamamoto, R., Yamamoto, Y., Oyama, C., Uehara, Y. (2015). Small molecules inhibiting the nuclear localization of YAP/TAZ for chemotherapeutics and chemosensitizers against breast cancers. *FEBS Open Bio.* 5: 542-549
- 281.** Frost, K., Seir, K., Lackner, A., Grusch, M., Grasi-Kraupp, B., Schulte-Hermann, R., Rodgarkia-Dara, C. (2011). Inhibin/activin expression in human and rodent liver: subunits  $\alpha$  and  $\beta$ B as new players in human hepatocellular carcinoma? *British Journal of Cancer.* 104: 1303-1312
- 282.** Liu, Q.Y., Niranjana, B., Gomes, P., Gomm, J.J., Davies, D., Coombes, R.C., Buluwela, L. (1996). Inhibitory effects of activin on the growth and morphogenesis of primary and transformed mammary epithelial cells. *Cancer Res.* 56(5): 1155-1163
- 283.** Shimonaka, M., Inoye, S., Shimasaki, S., Ling, N. (1991). Follistatin binds to both activin and inhibin through the common beta-subunit. *Endocrinology.* 128(6): 3313-3315
- 284.** McMahon, J.A., Takada, S., Zimmerman, L.B., Fan, C., Harland, R.M., McMahon, A.P. (1998). Noggin-mediated antagonism of BMP signalling is required for growth and patterning of the neural tube and somite. *Genes & Development.* 12: 1438-1452
- 285.** Ling, M.T., Wang, X., Zhang, X., Wong, Y.C. (2006). The multiple roles of Id-1 in cancer progression. *Differentiation.* 74(9-10): 481-487
- 286.** Weber, C.K., Sommer, G., Michl, P., Fensterer, H., Weimer, M., Gansauge, F., Leder, G., Adler, G., Gress, T.M. (2001). Biglycan is Overexpressed in Pancreatic Cancer and Induces G1-Arrest in Pancreatic Cancer Cell lines. *Gastroenterology.* 121: 657-667
- 287.** Bischof, A.G., Yuksel, D., Mammoto, T., Mammoto, A., Krause, S., Ingber, D.E. (2013). Breast cancer normalization induced by embryonic mesenchyme is mediated by extracellular matrix biglycan. *Integr Biol.* 5(8): 1045-56
- 288.** Kim, M., Jho, E. (2014). Cross-talk between Wnt/ $\beta$ -catenin and Hippo signalling pathways: a brief review. *BMB Reports.* 47(10): 540-545
- 289.** Kida, J., Hata, K., Nakamura, E., Yagi, H., Takahata, Y., Murakami, T., Maeda, Y., Nishimura, R. (2018). Interaction of LEF1 with TAZ is necessary for the osteoblastogenic activity of Wnt3a. *Scientific Reports.* 8: 10375

- 290.** Jho, E., Zhang, T., Domon, C., Joo, C. (2002). Wnt/ $\beta$ -Catenin/Tcf Signaling Induces the Transcription of Axin2, a Negative Regulator of the Signaling Pathway. *Molecular and Cellular Biology*. 22(4): 1172-1183
- 291.** Angonin, D., Van Raay, T.J. (2013). Nkd1 Functions as a Passive Antagonist of Wnt Signalling. *PLoS ONE*. 8(8)
- 292.** Li, N., Lu, N., Xie, C. (2019). The Hippo and Wnt signalling pathways: crosstalk during neoplastic progression in gastrointestinal tissue. *FEBS*. 286(19)
- 293.** Enzo, E., Santinon, G., Pocaterra, A., Aragona, M., Bresolin, S., Forcato, M., Grifoni, D., Pession, A., Zanconato, F., Guzzo, G., Bicciato, S., Dupont, S. (2015). Aerobic glycolysis tunes YAP/TAZ transcriptional activity. *EMBO*. 1349 – 1370
- 294.** Niedworok, C., Rock, K., Kretschmer, I., Freudenberger, T., Nagy, N., Szarvas, T., vom Dorp, F., Reis, H., Rubben, H., Fischer, J. W. (2013). Inhibitory role of the small leucine-rich proteoglycan biglycan in bladder cancer. *PLoS One*. 8(11): e80084
- 295.** Fisher, L. W., Termine, J. D., Young, M. F. (1989). Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. *The Journal of Biological Chemistry*. 264(8): 4571 – 4576
- 296.** Markusic, D., Oude-Elferink, R., Das, A., Berkhout, B., Seppen, J. (2005). Comparison of single regulated lentiviral vectors with rtTA expression driven by an autoregulatory loop or a constitutive promoter. *Nucleic Acids Research*. 33(6): e63
- 297.** Maishi, N., Ohba, Y., Akiyama, K., Ohga, N., Hamada, J., Nagao-Kitamoto, H., Alam, M. T., Yamamoto, K., Kawamoto, T., Inoue, N., Taketomi, A., Shindoh, M., Hida, Y., Hida, K. (2016). Tumour endothelial cells in high metastatic tumours promote metastasis via epigenetic dysregulation of biglycan. *Scientific Reports*. 28039
- 298.** Heergaard, A.M., Xie, Z., Young, M.F., Nielsen, K.L. (2004). Transforming growth factor beta stimulation of biglycan gene expression is potentially mediated by sp1 binding factors. *J Cell Biochem*. 93(3): 463-75
- 299.** Ungeforen, H., Groth, S., Ruhnke, M., Kalthoff, H., Fandrich, F. (2005). Transforming growth factor- $\beta$  (TGF- $\beta$ ) type I receptor/ALK5-dependent activation of the GADD45 $\beta$  gene mediates the induction of biglycan expression by TGF- $\beta$ . *Journal of Biological Chemistry*. 280: 2644 – 2652

- 300.** Schmitz, B., Salomon, A., Rotrige, A., Ritter, M., Ringelstein, B., Fishcer, J. W., Paul, M., Brand, E., Brand, S. (2013). Interindividual transcriptional regulation of the human biglycan gene involves three common molecular haplotypes. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 33: 871 – 880
- 301.** Wasyluk, B., Hagman, J., Guierrez-Hartmann, A. (1998). Ets transcription factors: Nuclear effectors of the Ras-MAPK-kinase signalling pathway. *Trends in Biochemical Sciences*. 23(6): 213 – 216
- 302.** Hildebrand, A., Romaris, M., Rasmussen, L. M., Heinegard, D., Twardzik, D. R., Border, W. A., Ruoslahti, E. (1994). Interaction of the small interstitial proteoglycans biglycan, decorin and fibromodulin with transforming growth factor beta. *Biochemical Journal*. 302(Pt 2): 527 – 534
- 303.** Lewandowski, S. L., Janardhan, H. P., Trivedi, C. M. (2015). Histone deacetylase 3 coordinates deacetylase-independent epigenetic silencing of transforming growth factor- $\beta$ 1 (TGF- $\beta$ 1) to orchestrate second heart field development. *Journal of Biological Chemistry*. 290(45): 27067 – 27089
- 304.** Chavez, K. J., Garimella, S. V., Lipkowitz, S. (2012). Triple negative breast cancer cell lines: One tool in the search for better treatment of triple negative breast cancer. *Breast Disease*. 31(1-2): 35 – 48
- 305.** Mostofa, A. G. M., Punganuru, S. R., Madala, H. R., Al-Obaide, M., Srivenugopal, K. S. (2017). The process and regulatory components of inflammation in brain oncogenesis. *Biomolecules*. 7(2): 34
- 306.** Claassen, G. F., Hann, S. R. (2000). A role for transcriptional repression of p21CIP1 by c-Myc in overcoming transforming growth factor beta-induced cell-cycle arrest. *PNAS*. 97(17): 9498 – 9503
- 307.** Jongwattapanisan, P., Terajima, M., Miguez, P. A., Querido, W., Nagaoka, H., Sumida, N., Gurysh, E. G., Ainslie, K. M., Pleshko, N., Perera, L., Yamauchi, M. (2018). Identification of the effector domain of biglycan that facilitates BMP-2 osteogenic function. *Scientific Reports*. 7022
- 308.** Chen, A., Wang, D., Liu, X., He, S., Yu, Z., Wang, J. (2012). Inhibitory effect of BMP-2 on the proliferation of breast cancer cells. *Molecular Medicine Reports*. 6(3): 1791 – 2997

- 309.** Schaefer, L., Babelova, A., Kiss, E., Hausser, H. J., Baliova, M., Krzyzankova, M., Marsche, G., Young, M. F., Mihalik, D., Gotte, M., Malle, E., Schaefer, R. M., Grone, H. J. (2005). The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages. *Journal of Clinical Investigation*. 115(8): 2223 – 2233
- 310.** Yang, H., Zhou, H., Feng, P., Zhou, X., Wen, H., Xie, X., Shen, H., Zhu, X. (2010). Reduced expression of Toll-like receptor 4 inhibits human breast cancer cells proliferation and inflammatory cytokines secretion. *Journal of Experimental and Clinical Cancer Research*. 29: 92
- 311.** El-Shennawy, L., Dubrovskiy, O., Kastrati, I., Danes, J. M., Zhang, Y., Whiteley, H. E., Creighton, C. J., Frasar, J. (2017). Coactivation of estrogen receptor and IKK $\beta$  induces a dormant metastatic phenotype in ER-positive breast cancer. *Tumour Biology and Immunology*. 78(4)
- 312.** Prunier, C., Baker, D., ten Dijke, P., Ritsma, L. (2019). TGF- $\beta$  Family Signaling Pathways in Cellular Dormancy. *Trends in Cancer*. 5(1): 66 – 78
- 313.** Epstein, T., Gatenby, R. A., Brown, J. S. (2017). The Warburg effect as an adaptation of cancer cells to rapid fluctuations in energy demand. *PLoS ONE*. 12(9): e0185085
- 314.** Gruetter, R., Novotny, E. J., Boulware, S. D., Rothman, D. L., Mason, G. F., Shulman, G. J., Shulman, R. G., Tamborlane, W. V. (1992). Direct measurement of brain glucose concentrations in humans by  $^{13}\text{C}$  NMR spectroscopy. *PNAS*. 89: 1109 – 1112
- 315.** Li, X., Gu, J., Zhou, Q. (2014). Review of aerobic glycolysis and its key enzymes – new targets for lung cancer therapy. *Thoracic Cancer*. 6(1)
- 316.** Zhang, D., Li, J., Wang, F., Hu, J., Wang, S., Sun, Y. (2014). 2-deoxy-D-glucose targeting of glucose metabolism in cancer cells as a potential therapy. *Cancer Letters*. 355(2): 176 – 183
- 317.** Zhang, X. D., Deslandes, E., Villedieu, M., Poulain, L., Duval, M., Gauduchon, P., Schwartz, L., Icard, P. (2006). Effect of 2-deoxy-D-glucose on various malignant cell lines in vitro. *Anticancer Research*. 26(5A): 3561 – 3566
- 318.** Kato, Y., Ozawa, S., Miyamoto, C., Maehata, Y., Suzuki, A., Maeda, T., Baba, Y. (2013). Acidic extracellular microenvironment and cancer. *Cancer Cell International*. 13: 89

- 319.** Goh, W., Sleptsova-Freidrich, I., Petrovic, N. (2014). Use of proton pump inhibitors as adjunct treatment for triple-negative breast cancers. An introductory study. *Journal of Pharmacy and Pharmaceutical Sciences*. 17(3): 439 – 446
- 320.** Lee, Y., Kim, N. H., Cho, E. S., Yang, J. H., Cha, Y. H., Kang, H. E., Yun, J. S., Cho, S. B., Lee, S., Paclikova, P. Radaszkiewicz, T. W., Bryja, V., Kang, C. G., Yuk, Y. S., Cha, S. Y., Kim, S., Kim, H. S., Yook, J. I. (2018). Dishevelled has a YAP nuclear export function in a tumour suppressor context-dependent manner. *Nature Communications*. 9: 2301
- 321.** Khanal, P., Yeung, B., Zhao, Y., Yang, X. (2019). Identification of prolyl isomerase Pin1 as a novel positive regulator of YAP/TAZ in breast cancer cells. *Scientific Reports*. 9: 6394
- 322.** Lin, K. C., Park, H. W., Guan, K. (2017). Regulation of the HIPPO pathway transcription factor TEAD. *Trends in Biochemical Sciences*. 42(11): 862 – 872
- 323.** Qin, Z., Xia, W., Fisher, G. J., Voorhees., Quan, T. (2018). YAP/TAZ regulates TGF- $\beta$ /Smad3 signaling by induction of Smad7 via AP-1 in human skin dermal fibroblasts. *Cell Communication and Singaling*. 16: 18
- 324.** Wang, J., Sinnott-Smith, J., Stevens, J. V., Young, S. H., Rozengurt, E. (2016). Biphasic regulation of Yes-associated protein (YAP) cellular localization, phosphorylation, and activity by G protein-coupled receptor agonists in intestinal epithelial cells: A novel role for protein kinase D (PKD). *Journal of Biological Chemistry*. 291: 17988 – 18005
- 325.** Nardone, G., Oliver-De La Cruz, J., Vrbsky, J., Martini, C., Pribyl, J., Skladal, P., Pesl, M., Caluori, G., Pagliari, S., Martino, F., Maceckova, Z., Hajduch, M., Sanz-Garcia, A., Pugno, N. M., Stokin, G. B., Forte, G. (2017). YAP regulates cell mechanics by controlling focal adhesion assembly. *Nature Communications*. 8(15321)
- 326.** Barczyk, M., Carracedo, S., Gullberg, D. (2010). Integrins. *Cell and Tissue Research*. 339(1): 269 – 280
- 327.** Maltseva, D. V., Rodin, S. A. (2018). Laminins in metastatic cancer. *Molekuliarnaia Biologiya*. 52(3): 411 – 434
- 328.** Wu, H., Lu, X., Wang, J., Yang, T., Li, X., He, X., Li, Y., Ye, W., Wu, Y., Gan, W., Guo, P., Li, J. (2019). TRAF6 inhibits colorectal cancer and metastasis through regulating selective autophagic CTNNB1/ $\beta$ -catenin degradation and is targeted for GSK3 $\beta$ /GSK3 $\beta$ -mediated phosphorylation and degradation. *Autophagy*. 15(9): 1506 – 1522

- 329.** Reaz, L. E., Papadopoulos, K., Ricard, A. D., Chiorean, E. G., DiPaola, R. S., Stein, M. N., Rocha Lima, C. M., Schlesselman, J. J., Tolba, K., Langmuir, V. K., Kroll, S., Jung, D. T., Kurtoglu, M., Rosenblatt, J., Lampidis, T. J. (2012). A phase I dose-escalation trial of 2-deoxy-D-glucose alone or combined with docetaxel in patients with advanced solid tumours. *Cancer Chemotherapy and Pharmacology*. 71(2): 523 – 530
- 330.** Yiengst, M. J., Shock, N. W. (1962). Blood and plasma volume in adult males. *Journal of Applied Physiology*. 17(2): 195 – 198
- 331.** Lucantoni, F., Dussmann, H., Prehn, J. H. M. (2018). Metabolic targeting of breast cancer cells with the 2-deoxy-D-glucose and the mitochondrial bioenergetics inhibitor MDIV1-1. *Frontiers in Cell and Developmental Biology*. 6(113)
- 332.** Chaube, B., Malvi, P., Singh, S. V., Mohammad, N., Viollet, B., Bhat, M. K. (2015). AMPK maintains energy homeostasis and survival in cancer cells via regulating p38/PGC-1 $\alpha$ -mediated mitochondrial biogenesis. *Cell Death Discovery*. 1: 15063
- 333.** Wu, Y., Sarkissyan, M., Mcghee, E., Lee, S., Vadgama, J. V. (2015). Combined inhibition of glycolysis and AMPK induces synergistic breast cancer cell killing. *Breast Cancer Research and Treatment*. 151(3): 529 – 539
- 334.** Carbonell, W. S., Ansorge, O., Sibson, N., Muschel, R. (2009). The vascular basement membrane as “soil” in brain metastasis. *PLoS ONE*. 4(6): e5857
- 335.** Beaulieu, J. (2010). Integrin  $\alpha 6\beta 4$  in colorectal cancer. *World Journal of Gastrointestinal Pathophysiology*. 1(1): 3 – 11
- 336.** Gupta, P., Srivastava, S. K. (2013). Cucurbitacin B suppresses breast cancer growth by inhibiting TGF $\beta$ -mediated integrin-HER2 signaling. *Cancer Research*. 73(8)
- 337.** Di Benedetto, M., Toullec, A., Buteau-Lozano, H., Abdelkarim, M., Vacher, S., Velasco, G., Christofari, M., Pocard, M., Bieche, I., Perrot-Applanat, M. (2015). MDA-MB-231 breast cancer cells overexpressing single VEGF isoforms display distinct colonization characteristics. *British Journal of Cancer*. 113: 773 – 785
- 338.** Donnem, T., Hu, J., Ferguson, M., Adighibe, O., Snell, C., Harris, A. L., Gatter, K. C., Pezzella, F. (2013). Vessel co-option in primary human tumours and metastases: An obstacle to effective anti-angiogenic treatment? *Cancer Medicine*. 2(4): 427 – 436
- 339.** Reinhardt, H. C., Hasskamp, P., Schmedding, I., Weaver, D., Carr, S. A., Yaffe, M. B. (2010). DNA damage activates a spatially distinct late cytoplasmic cell-cycle

checkpoint network controlled by MK2-mediated RNA stabilization. *Molecular Cell*. 40(1): 34 – 49

**340.** Liao, B., Hu, Y., Brewer, G. (2007). Competitive binding of AUF1 and TIAR to MYC mRNA controls its translation. *Nature Structural and Molecular Biology*. 14(6): 511 – 518

**341.** Simoes-Sousa, S., Littler, S., Thomson, S. L., Minshall, P., Whalley, H., Bakker, B., Belkot, K., Moralli, D., Bronder, D., Tighe, A., Spierings, D. C. J., Bah, N., Graham, J., Nelson, L., Green, C. M., Foijer, F., Townsend, P. A., Taylor, S. S. (2018). The p38 $\alpha$  stress kinase suppresses aneuploidy tolerance by inhibiting Hif-1 $\alpha$ . *Cell Reports*. 25(3): 749 – 760

**342.** Huang, D., Li, X., Sun, L., Huang, P., Ying, H., Wu, J., Song, H. (2016). Regulation of Hippo signalling by p38 signalling. *Journal of Molecular Cell Biology*. 8(4): 328 – 337

**343.** Lin, K. C., Moroishi, T., Meng, Z., Jeong, H. S., Plouffe, S. W., Sekido, Y., Han, J., Park, H. W., Guan, K. L. (2017). Regulation of the Hippo pathway transcription activator TEAD by p38 MAPK-induced cytoplasmic translocation. *Nature Cell Biology*. 19(8): 996 – 1002

**344.** Yang, S., Zhang, L., Liu, M., Chong, R., Ding, S., Chen, Y., Dong, J. (2013). CDK1 phosphorylation of YAP promotes mitotic defects and cell motility and is essential for neoplastic transformation. *Molecular and Cellular Pathology*. 73(22): 6722 – 6733

**345.** Zhao, W., Liu, H., Wang, J., Wang, M., Shao, R. (2018). Cyclizing-berberine A35 induces G2/M arrest and apoptosis by activating YAP phosphorylation (Ser127). *Journal of Experimental and Clinical Cancer Research*. 37: 98

**346.** Putney, L. K., Barber, D. L., (2003). Na-H exchange-dependent increase in intracellular pH times G2/M entry and transition. *Journal of Biological Chemistry*. 278(45): 44645 – 44649

**347.** Flinck, M., Kramer, S. H., Schnipper, J., Anderson, A. P., Pedersen, S. F. (2018). The acid-base transport proteins NHE1 and NBCn1 regulate cell cycle progression in human breast cancer cells. *Cell Cycle*. 17(9): 1056 – 1067

**348.** Pike Winer, L. S., Wu, M. (2014). Rapid analysis of glycolytic and oxidative substrate flux of cancer cells in a microplate. *PLoS One*. 9(10): e109916

**349.** Keschull, J. M., Zador, A. M. (2015). Sources of PCR-induced distortions in high-throughput sequencing data sets. *Nucleic Acids Research*. 43(21): e143

- 350.** Peng, Y., Croce, C. M. (2016). The role of microRNAs in human cancer. *Signal Transduction and Targeted Therapy*. 15004
- 351.** Paneru, B. D., Al-Tobasei, R. Kenney, B., Leeds, T. D., Salem, M. (2017). RNA-seq reveals microRNA expression signature and genetic polymorphism associated with growth and muscle quality traits in rainbow trout. *Scientific Reports*. 7: 9078
- 352.** Alfaridus, H., McIntyre, A., Smith, S. (2017). MicroRNA regulation of glycolytic metabolism in glioblastoma. *BioMed Research International*. 2017: 9157370
- 353.** Hughes, A. J., Spelke, D. P., Xu, Z., Kang, C., Schaffer, D. V., Herr, A. E. (2014). Single-cell Western blotting. *Nature Methods*. 11: 749 – 755
- 354.** Paschall, A. V., Liu, K. (2016). An orthotopic mouse model of spontaneous breast cancer metastasis. *Journal of Visualised Experiments*. 114: 354040
- 355.** Bailey-Downs, L. C., Thorpe, J. E., Disch, B. C., Bastian, A., Hauser, P. J., Farasyn, T., Berry, W. L., Hurst, R. E., Ihnat, M. A. (2014). Development and characterization of a preclinical model of breast cancer lung micrometastatic to macrometastatic progression. *PLoS One*. 9(5): e98624
- 356.** Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C. Sauer, M., Biehlmaier, O., Drummen, G. P. C. (2019). Super-resolution microscopy demystified. *Nature Cell Biology*. 21: 72 – 84
- 357.** Shangguan, W. J., Li, H., Zhang, Y. H. (2014). Induction of G2/M phase cell cycle arrest and apoptosis by ginsenoside Rf in human osteosarcoma MG-63 cells through the mitochondrial pathway. *Oncology Reports*. 31(1): 305 – 313
- 358.** Kumar, A., Sasmal, D., Sharma, N. (2014). Deltamethrin induced an apoptogenic signalling pathway in murine thymocytes: exploring the molecular mechanism. *Journal of Applied Toxicology*. 34(12): 1303 – 1310
- 359.** Pozarowski, P., Darzynkiewicz, Z. (2004). Analysis of cell cycle by flow cytometry. *Methods in Molecular Biology*. 281: 301 – 311