

# **A proteolytic landscape: understanding the role of myoepithelial cells in early-stage breast cancer progression**

**Shayin Vivien Gibson**

Submitted in partial fulfilment of the requirements of the Degree of  
Doctor of Philosophy

April 2023

Queen Mary University of London  
Barts Cancer Institute  
Centre for Tumour Biology  
John Vane Science Centre  
Charterhouse Square  
London  
EC1M 6BQ

## Statement of originality

I, Shayin Vivien Gibson, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature:

Shayin Vivien Gibson

Date: 6<sup>th</sup> July 2023

## Acknowledgements

Firstly, I would like to thank both Richard and Ed for their supervision throughout my PhD. You have been the most amazing mentors and I cannot believe how lucky I am to have had both of you guide me over the last few years. Your invaluable advice, support and encouragement means the world.

Secondly, thank you to the Grose Lab group members, both old and new, for making this the most enjoyable group to work in. To Demi, Elena, Nick and Reza – it's been an amazing journey and I am so glad we got to experience it together. To our incredible postdocs Lucia, Abi and Kubra, thank you for your words of encouragement and guidance. To Cerian and Qiaoying, thank you for your support over the last part of my PhD and for keeping my spirits up during thesis writing! You are all amazing friends, and I will forever cherish the memories we have made. I would also like to say an additional thank you Mike, Louise, John, Barrie and Angus for their scientific support and advice.

To my family, thank you for being my biggest supporters and for always being proud of me. And finally, to Matt, you have been my rock – thank you for all your love and support.

## Publications

- Gibson SV, Madzharova E, Tan AC, Allen MD, Keller UAD, Jones JL, Carter EP, Grose RP. ADAMTS3 restricts cancer invasion in models of early breast cancer progression through enhanced fibronectin degradation. *Matrix Biol.* 2023 (In press)
- Gibson SV, Tomas Bort E, Rodríguez-Fernández L, Allen MD, Gomm JJ, Goulding I, auf dem Keller U, Agnoletto A, Brisken C, Peck B, Cameron AJ, Marshall JF, Jones JL, Carter EP, Grose RP. TGFβ-mediated MMP13 secretion drives myoepithelial cell dependent breast cancer progression. *NPJ Breast Cancer.* 2023 Mar 2;9(1):9.
- Gibson SV, Roozitalab RM, Allen MD, Jones JL, Carter EP, Grose RP. Everybody needs good neighbours: the progressive DCIS microenvironment. *Trends Cancer.* 2023 Apr;9(4):326-338.
- Carter EP, Roozitalab R, Gibson SV, Grose RP. Tumour microenvironment 3D-modelling: simplicity to complexity and back again. *Trends Cancer.* 2021 Nov;7(11):1033-1046.

## Invited oral/poster presentations

- GRC Mammary Gland Conference (Poster), Italy, 2022
- City of London Tumour Microenvironment Cancer Models (Talk), Barts Cancer Institute, 2022
- CRUK City of London PhD Symposium (Talk), 2021/2022
- Crick Cancer Conference (Poster), The Francis Crick Institute, 2021
- Goodbye Flat Biology (Poster), London, 2021
- NC3Rs Conference (Poster), London, 2021
- Organ-on-a-Chip (Talk), Queen Mary University London, 2020



## Awards

- 1<sup>st</sup> Place, PhD Day Best Poster, Barts Cancer Institute, London, 2022
- 1<sup>st</sup> Place, William Harvey Day Flash Talk, William Harvey Research Institute, London 2021
- 1<sup>st</sup> Place, PhD Public Research Engagement Talk, Barts Cancer Institute, London, 2019

## Collaborations

- S. Gibson and R. Roozitalab, Queen Mary University London, designed Figure 1.4, which was created using Illustrator software by R. Roozitalab.
- The 1089 myoepithelial cell line was provided by Dr. M. Allen, Queen Mary University London.
- Primary cells were provided by Dr. J. Gomm and I. Goulding, Queen Mary University London, Breast Cancer Now Cell Bank.
- Dr. E. Carter, Queen Mary University London, constructed the ITGB6 pINDUCER plasmid and generated Imaris reconstructions for Figures 3.2 and 4.17. For Figures 3.4 and 3.10, Dr. E. Carter performed two out of four biological repeats, while S. Gibson performed the remaining replicates and quantification analyses across all replicates.
- Prof. J. Marshall, Queen Mary University London, provided scientific advice on ITGB6-related investigations.
- The SMAD-SBE luciferase reporter construct used in Figure 4.8 was provided by Dr. A. Cameron, Queen Mary University London.
- Dr. B. Peck, Queen Mary University London, provided the SGCBP30 inhibitor used in Figures 4.13 and 4.14 and provided input on EP300-related investigations.
- RNA sequencing and initial data processing was performed by Dr. C. Mein, Barts Genome Centre, Queen Mary University London.

- Differential expression analysis, using normalised counts from RNAseq, was performed by E. Bort, Queen Mary University London.
- Clinical breast tissue sections displayed in Figures 4.19 and 4.20 were provided by Prof. L. Jones, Queen Mary University London, Breast Cancer Now Tissue Bank.
- Integrin  $\beta 6$  immunohistochemistry in Figures 4.19 and 4.20 were performed by N. Rahman, Barts Cancer Institute Pathology Facility, Queen Mary University London.
- Unstained tissue sections from the intraductal mouse model were provided by Dr. A. Agnoletto and Prof. C. Briskin, École Polytechnique fédérale de Lausanne, for RNAscope analysis in Figure 4.18.
- A. Tan, Queen Mary University London, supported initial ADAMTS3 studies, performing two out of three biological replicates in Figure 5.3.
- Degradomic workflows were performed in collaboration with Dr. E. Madzharova and Prof. U. auf dem Keller, Technical University of Denmark. Dr. E. Madzharova performed the TAILS workflow while Prof. U. auf dem Keller performed the data processing.

## Funding

This work was supported by Cancer Research UK (C10847/A27781), Breast Cancer Now (2017NovPR988), Barts Charity (MRC0173), Rosetrees Trust (M501-F1), CRUK Microscopy Core Service Grant (C16420/A18066) and by a CRUK UK Centre Grant to Barts Cancer Institute (C355/A25137). Prof. U. auf dem Keller is supported by a Novo Nordisk Foundation Young Investigator Award (NNF16OC0020670).

## Abstract

Ductal carcinoma *in situ* (DCIS) is a non-obligate precursor of invasive breast cancer. Despite evidence suggesting that up to half of DCIS cases would remain stable and non-threatening, virtually all women with DCIS are treated, highlighting overtreatment as a pressing issue in DCIS management.

To understand the role of the tumour-associated myoepithelium, we present a 3D *in vitro* model that incorporates both luminal and myoepithelial cells in physiomimetic conditions. We demonstrate that integrin  $\beta 6$ -expressing DCIS-associated myoepithelial cells promote striking myoepithelial-led invasion of luminal cells via upregulation of the collagenase MMP13, through a non-canonical TGF $\beta$ /EP300 pathway. *In vivo*, MMP13 expression is associated with stromal invasion in a murine model of DCIS progression and is elevated in myoepithelial cells of clinical high-grade DCIS cases.

In addition to the upregulation of tumour-promoting MMP13, we also demonstrate the repressive function of myoepithelial ADAMTS3, with increased invasion observed upon ADAMTS3 siRNA-mediated loss. Degradomic analysis, using a terminal amine isotopic labelling of substrates approach, combined with functional assays, implicate fibronectin as a direct ADAMTS3 substrate. We further show that loss of ADAMTS3 enhances fibronectin levels in the microenvironment and promotes invasion through canonical integrin  $\alpha 5 \beta 1$  activation. Our data identify a key role for myoepithelial-derived MMP13 and ADAMTS3 loss in facilitating DCIS progression, pointing the way towards robust markers for risk stratification in DCIS patients.

## Table of contents

Statement of originality .....	2
<b>Abstract.....</b>	<b>7</b>
List of tables.....	15
List of figures.....	16
List of appendix figures .....	19
List of abbreviations.....	20
<b>Chapter 1. Introduction .....</b>	<b>30</b>
1.1. Breast Cancer .....	31
1.1.1. Epidemiology and risk factors.....	31
1.1.2. Cellular architecture of the breast.....	32
1.1.3. Breast cancer development and progression.....	34
1.1.4. Classification of breast cancer .....	36
1.2. DCIS.....	39
1.2.1. Screening and current treatment guidelines.....	39
1.2.2. Overdiagnosis and overtreatment.....	41
1.2.3. Tools for predicting risk of recurrence .....	42
1.2.4. Active surveillance trials for DCIS .....	44
1.2.5. Comparative genomics of DCIS and IDC .....	46
1.3. The tumour microenvironment in DCIS progression.....	47
1.3.1. Overview .....	47

1.3.2. Myoepithelial cells .....	47
1.3.3. Adipocytes .....	48
1.3.4. Immune cells .....	50
1.3.5. Vasculature .....	51
1.3.6. Fibroblasts .....	52
1.3.7. Matrix .....	53
1.3.8. Potential tumour-extrinsic markers of DCIS progression .....	55
1.4. Metzincins .....	60
1.4.1. Overview .....	60
1.4.2. Classification and structure of metzincins .....	61
1.4.3. Regulation and activation of metzincins .....	65
1.4.4. Role of metzincins in cancer .....	66
1.5. Integrins .....	70
1.5.1. Overview .....	70
1.5.2. Integrin signalling in cancer .....	72
1.5.3. Integrin $\alpha\beta 6$ .....	74
1.6. TGF $\beta$ signalling .....	75
1.6.1. TGF $\beta$ signalling overview .....	75
1.6.2. Relevance of TGF $\beta$ signalling in cancer .....	78
1.7. Fibronectin .....	80
1.7.1. Overview and structure .....	80

1.1.2. Function of fibronectin and relevance in cancer .....	82
1.8. Hypothesis and project aims .....	83
<b>Chapter 2. Methods .....</b>	<b>85</b>
2.1. Cell culture .....	86
2.1.1. Cell culture conditions .....	86
2.1.2. 3D cell culture models .....	89
2.1.3. Transient cell transfections .....	96
2.1.4. Lentiviral production and cell transduction .....	99
2.1.5. Fibronectin cleavage validation .....	100
2.2. Molecular cloning .....	102
2.2.1. HER2 and ITGB6 expression plasmids.....	102
2.2.2. MMP13 expression plasmid .....	102
2.2.3. ADAMTS3 expression plasmids .....	108
2.2.4. Bacterial transformation and plasmid extraction .....	108
2.3. Chromatin immunoprecipitation (ChIP) .....	110
2.3.1. Chromatin isolation .....	110
2.3.2. Immunoprecipitation.....	111
2.3.3. ChIP-qPCR .....	111
2.4. Transcriptomic analysis of cells .....	113
2.4.1. PCR.....	113
2.4.2. RNAseq.....	116

2.5. Proteomic analysis of cells.....	118
2.5.1. Immunofluorescence.....	118
2.5.2. Western blotting.....	121
2.5.3. Degradomic workflow .....	124
2.6. Clinical breast tissue .....	129
2.6.1. Ethical approval and cohort data .....	129
2.6.2. RNAscope.....	131
2.6.3. Immunohistochemistry.....	132
2.7. Animal experiments.....	135
2.7.1. Animals.....	135
2.7.2. Intraductal xenograft .....	135
2.8. Statistical analysis .....	136
<b>Chapter 3. Integrin <math>\beta</math>6 expression drives myoepithelial-led invasion.....</b>	<b>137</b>
3.1. Introduction .....	138
3.2. Generation and characterisation of integrin $\beta$ 6-inducible DCIS models.....	140
3.2.1. Generation of integrin $\beta$ 6-inducible myoepithelial cells.....	140
3.2.2. Integrin $\beta$ 6 expression drives invasion in primary ductal model of DCIS .....	142
3.2.3. Myoepithelial integrin $\beta$ 6 expression drives invasion in cell line sphere model .....	144
3.2.4. Myoepithelial integrin $\beta$ 6 expression drives invasion in primary sphere model .....	146
3.2.5. HB2 invasion is led by myoepithelial cells .....	148
3.3. Integrin $\beta$ 6 expression drives invasion through a protease-dependent mechanism.....	151

3.3.1. Invading spheroids remodel their basement membrane upon integrin $\beta 6$ expression .....	151
3.3.2. Integrin $\beta 6$ -dependent invasion requires metzincin activity .....	153
3.3.3. Integrin $\beta 6$ alters metzincin expression .....	157
3.4. Discussion.....	159
<b>Chapter 4. MMP13 is required for integrin <math>\beta 6</math>-dependent myoepithelial-led invasion .....</b>	<b>161</b>
4.1. Introduction .....	162
4.2. Integrin $\beta 6$ -driven invasion is MMP13 dependent.....	163
4.2.1. Knock-down of myoepithelial MMP13 reduces invasion .....	163
4.2.2. Treatment with MMP13i reduces $\beta 6$ -driven invasion.....	165
4.2.3. Integrin $\beta 6$ -independent overexpression of MMP13 drives invasion.....	169
4.3. Integrin $\beta 6$ -driven MMP13 upregulation is TGF $\beta$ -dependent.....	172
4.3.1. Integrin $\beta 6$ expression drives canonical TGF $\beta$ signalling .....	172
4.3.2. Inhibition of TGF $\beta$ signalling blocks integrin $\beta 6$ -driven MMP13 upregulation .....	175
4.4. TGF $\beta$ -dependent expression of MMP13 requires activation of EP300.....	178
4.4.1. Identification of integrin $\beta 6$ /TGF $\beta$ -dependent transcriptional regulators.....	178
4.4.2. Validation of EP300 as a driver of MMP13 expression .....	181
4.4.3. Knock-down of EP300 reduces integrin $\beta 6$ -driven invasion.....	184
4.4.4. EP300 expression is associated with poor prognosis in breast cancer .....	184
4.4.5. Integrin $\beta 6$ -driven invasion is dependent on EP300/MMP13 in a duct model.....	187
4.5. MMP13 expression <i>in vivo</i> and in clinical cases .....	189
4.5.1. Stromal Mmp13 expression is associated with invasion <i>in vivo</i> .....	189



4.5.2. Myoepithelial cells express MMP13 in high-grade DCIS .....	191
4.5.3. MMP13 is expressed in IDC cases.....	193
4.6. Discussion.....	195
<b>Chapter 5. Loss of ADAMTS3 in DCIS-myoepithelial cells contributes to invasive progression through increased fibronectin/integrin <math>\alpha 5\beta 1</math> signalling .....</b>	<b>199</b>
5.1. Introduction .....	200
5.2. Myoepithelial ADAMTS3 loss drives invasion.....	201
5.2.1. ADAMTS3 expression is lost in DCIS and IDC.....	201
5.2.2. Integrin $\beta 6$ -induction drives ADAMTS3 loss in myoepithelial cells .....	203
5.2.3. ADAMTS3 loss drives myoepithelial-led invasion.....	205
5.2.4. Myoepithelial ADAMTS3 expression inhibits invasion .....	208
5.3. Degradomic analysis reveals ADAMTS3 substrate candidates.....	210
5.3.1. Analysis of ADAMTS3 degradome .....	210
5.3.2. ADAMTS3 substrates are associated with matrix organisation and migration.....	212
5.4. Proteolytic activity of ADAMTS3 is required for cleavage of fibronectin.....	214
5.4.1. ADAMTS3 mediates fibronectin cleavage .....	214
5.4.2. ADAMTS3 expression regulates levels of fibronectin in conditioned media .....	216
5.5. Fibronectin-dependent RGD integrin signalling drives invasion upon ADAMTS3 loss.....	218
5.5.1. Increased fibronectin is localised to protrusions of invading spheroids .....	218
5.5.2. Knockdown of myoepithelial fibronectin blocks invasion upon ADAMTS3 loss .....	218
5.5.3. ADAMTS3 loss drives Rho GTPase signalling .....	222
5.5.4. Treatment with RGD peptide blocks invasion upon ADAMTS3 loss .....	222

5.5.5. Invading spheroids display increased integrin $\alpha 5 \beta 1$ and FAK activation .....	225
5.5.6. Knockdown of myoepithelial-integrin $\alpha 5$ blocks invasion upon ADAMTS3 loss .....	225
5.6. Discussion.....	228
<b>Chapter 6. Final Discussion.....</b>	<b>231</b>
References.....	240
Appendix.....	256

## List of tables

Table 1.1 Intrinsic classification of breast cancer subtypes .....	38
Table 1.2 Summary of active surveillance trials .....	45
Table 1.3 Microenvironmental markers based on recurrence data .....	58
Table 1.4 Microenvironmental markers based on comparing disease states.....	59
Table 2.1 Collagen-I gel recipe for primary duct model .....	90
Table 2.2 Corresponding methylcellulose and cell suspension volumes .....	92
Table 2.3 Collagen-I gel recipe for HB2/1089 cell line model .....	92
Table 2.4 Collagen-I gel recipe for HB2/1° Myo model .....	94
Table 2.5 SiRNA targets and catalogue identifiers .....	97
Table 2.6 Cell culture treatment conditions.....	101
Table 2.7 MMP13 gateway cloning primers .....	105
Table 2.8 Genomic chIP qPCR primers .....	112
Table 2.9 List of qPCR primers .....	115
Table 2.10 Antibodies for immunofluorescence .....	120
Table 2.11 Antibody for western blotting .....	123
Table 2.12 Age at surgery and diagnosis of patients.....	130
Table 2.13 DCIS and IDC grade .....	130
Table 2.14 Antibodies for immunohistochemistry .....	134
Table 6.1 Myoepithelial biomarker panel for validation.....	238

## List of figures

Figure 1.1 Anatomy of the human breast. ....	33
Figure 1.2 Classical model of breast cancer progression. ....	35
Figure 1.3 Breast carcinoma in situ trends over time. ....	40
Figure 1.4 Role of the DCIS microenvironment in facilitating DCIS progression. ....	57
Figure 1.5 Classification and structure of MMPs. ....	63
Figure 1.6 Classification and structure of ADAM and ADAMTS proteases. ....	64
Figure 1.7 Overview of integrin signalling. ....	71
Figure 1.8 Overview of TGF $\beta$ activation and signalling. ....	77
Figure 1.9 Structure of fibronectin and overview of fibronectin fibrillogenesis. ....	81
Figure 2.1 Schematic summarising gateway cloning strategy. ....	103
Figure 2.2 Gateway cloning and construction of pLV CMV MMP13 plasmid. ....	107
Figure 2.3 Schematic summarising proteomic workflow. ....	128
Figure 3.1 Generation of integrin $\beta$ 6-inducible cells. ....	141
Figure 3.2 Integrin $\beta$ 6 expression drives invasion in primary ductal model of DCIS. ....	143
Figure 3.3 Myoepithelial integrin $\beta$ 6 expression drives invasion in HB2/1089 <sup>i<math>\beta</math>6</sup> model. ....	145
Figure 3.4 Myoepithelial integrin $\beta$ 6 expression drives invasion in HB2/Myo <sup>i<math>\beta</math>6</sup> model. ....	147
Figure 3.5 HB2 invasion is led by myoepithelial cells in HB2/1089 <sup>i<math>\beta</math>6</sup> model. ....	149
Figure 3.6 HB2 invasion is led by myoepithelial cells in HB2/Myo <sup>i<math>\beta</math>6</sup> model. ....	150
Figure 3.7 BM remodelling is associated with invasion in the HB2/1089 <sup>i<math>\beta</math>6</sup> model. ....	152
Figure 3.8 Integrin $\beta$ 6-driven invasion is protease-dependent in the HB2/1089 <sup>i<math>\beta</math>6</sup> model. ....	154
Figure 3.9 Protease activity is required for collagen-I cleavage in HB2/1089 <sup>i<math>\beta</math>6</sup> model. ....	155
Figure 3.10 Integrin $\beta$ 6-driven invasion is protease-dependent in the HB2/Myo <sup>i<math>\beta</math>6</sup> model. ....	156
Figure 3.11 Overexpression of integrin $\beta$ 6 in myoepithelial cells alters metzincin expression. ....	158

Figure 4.1 Integrin $\beta 6$ -driven invasion is MMP13 dependent in the HB2/1089 <sup>i<math>\beta 6</math></sup> model. ....	164
Figure 4.2 Integrin $\beta 6$ -driven invasion is dependent on MMP13 activity in the HB2/1089 <sup>i<math>\beta 6</math></sup> model. ....	166
Figure 4.3 Integrin $\beta 6$ -driven invasion is dependent on MMP13 activity in the HB2/Myo <sup>i<math>\beta 6</math></sup> model. ....	167
Figure 4.4 Invasion requires MMP13-dependent collagen-I cleavage in HB2/1089 <sup>i<math>\beta 6</math></sup> model. ....	168
Figure 4.5 Validation of pLV CMV MMP13 HA-Tag construct. ....	170
Figure 4.6 MMP13 drives invasion independently of integrin $\beta 6$ in the HB2/1089 model.....	171
Figure 4.7 TGF $\beta$ signalling is enriched following myoepithelial integrin $\beta 6$ induction. ....	173
Figure 4.8 Integrin $\beta 6$ induction drives canonical TGF $\beta$ signalling. ....	174
Figure 4.9 Integrin $\beta 6$ -driven upregulation of MMP13 is dependent on TGF $\beta$ signalling.....	176
Figure 4.10 $\beta 6$ -driven invasion is dependent on TGF $\beta$ signalling in the HB2/1089 <sup>i<math>\beta 6</math></sup> model. ....	177
Figure 4.11 Activation Z-scores of integrin $\beta 6$ /TGF $\beta$ -dependent transcriptional regulators.....	179
Figure 4.12 SiRNA screen reveals EP300 as a potential transcriptional regulator of MMP13. ....	180
Figure 4.13 MMP13 upregulation is dependent on expression and activity of EP300. ....	182
Figure 4.14 EP300 regulates MMP13 expression through H3K27ac enrichment. ....	183
Figure 4.15 EP300 expression is required for $\beta 6$ -driven invasion. ....	185
Figure 4.16 EP300 expression is associated with poor prognosis in breast cancer. ....	186
Figure 4.17 Integrin $\beta 6$ -driven invasion is EP300/MMP13-dependent in ductal model.....	188
Figure 4.18 Stromal Mmp13 is associated with invasion in intraductal mouse model. ....	190
Figure 4.19 MMP13 is localised to integrin $\beta 6$ -expressing myoepithelial cells in DCIS. ....	192
Figure 4.20 Expression of MMP13 is increased in IDC. ....	194
Figure 4.21 Proposed mechanism of the $\beta 6$ /MMP13 axis in myoepithelial-led invasion.....	198
 Figure 5.1 ADAMTS3 loss is observed in DCIS and IDC. ....	 202
Figure 5.2 Integrin $\beta 6$ expression drives ADAMTS3 loss. ....	204
Figure 5.3 Loss of myoepithelial ADAMTS3 drives invasion in HB2/1089 model. ....	206
Figure 5.4 Loss of myoepithelial ADAMTS3 drives invasion in HB2/Myo model. ....	207

Figure 5.5 Myoepithelial expression of ADAMTS3 inhibits integrin $\beta$ 6-induced invasion .....	209
Figure 5.6 Summary of detected neo-N-term peptides. ....	211
Figure 5.7 Pathway enrichment analysis of candidate ADAMTS3 substrates. ....	213
Figure 5.8 Proteolytic activity of ADAMTS3 is required for cleavage of fibronectin.....	215
Figure 5.9 ADAMTS3 regulates levels of fibronectin in culture medium. ....	217
Figure 5.10 Fibronectin is localised to invading protrusions. ....	219
Figure 5.11 Loss of myoepithelial fibronectin reduces invasion in HB2/1089 model. ....	220
Figure 5.12 Fibronectin gels drive invasion in HB2/1089 model.....	221
Figure 5.13 Analysis of enriched pathways following ADAMTS3 knockdown.....	223
Figure 5.14 Inhibition of RGD integrins blocks invasion upon ADAMTS3 loss. ....	224
Figure 5.15 Enhanced integrin $\alpha$ 5 activation is observed following ADAMTS3 knockdown.....	226
Figure 5.16 Integrin $\alpha$ 5 is required for invasion upon ADAMTS3 loss.....	227
Figure 5.17 Proposed tumour-suppressive mechanism of myoepithelial ADAMTS3. ....	230
 Figure 6.1 Proposed tumour-promoting mechanism of myoepithelial cells in DCIS progression. ....	 235

## List of appendix figures

Appendix 1.1 Matrix expression changes following integrin $\beta 6$ induction in 1089 <sup>i<math>\beta</math>6</sup> cells.....	256
Appendix 1.2 HB2 monoculture spheroids.....	257
Appendix 2.1 Validation of MMP13 siRNA knockdown in 1089 <sup>i<math>\beta</math>6</sup> cells. ....	258
Appendix 2.2 MMP13 expression following SMAD4 siRNA knockdown in 1089 <sup>i<math>\beta</math>6</sup> cells.....	259
Appendix 2.3 Validation of EP300 siRNA knockdown in 1089 <sup>i<math>\beta</math>6</sup> cells. ....	260
Appendix 3.1 ADAMTS3 expression in cell types of the human breast. ....	261
Appendix 3. 2 Validation of ADAMTS3 siRNA knockdown in 1089 cells. ....	262
Appendix 3.3 Validation of ADAMTS3 expression constructs in 1089 cells. ....	263
Appendix 3.4 Validation of fibronectin siRNA knockdown in 1089 cells.....	264
Appendix 3.5 Integrin expression in 1089 and primary myoepithelial cells. ....	265
Appendix 3.6 Validation of integrin $\alpha 5$ siRNA knockdown in 1089 cells.....	266

## List of abbreviations

ACTB	Actin beta
ADAM	A disintegrin and metalloproteinase
ADAMTS	A disintegrin and metalloproteinase with thrombospondin motifs
AKT	AKT Serine/threonine kinase
AURKA	Aurora kinase A
BIRC5	Baculoviral IAP repeat containing 5
BM	Basement membrane
Bp	Base pairs
BSA	Bovine serum albumin
CA	Cysteine array
CAA	Cancer-associated adipocyte
CAA	Chloroacetamide
CAF	Cancer-associated fibroblast
CCL	Chemokine C-C motif
CCNB1	Cyclin B1
CDH3	P-cadherin
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CK	Cytokeratin



CMV	Cytomegalovirus
CO <sub>2</sub>	Carbon dioxide
COX	Prostaglandin endoperoxide synthase
CRUK	Cancer research UK
CTCF	Connective tissue growth factor
CXCL	C-X-C motif chemokine ligand
Cys	Cysteine-rich domain
Cyto	Cytoplasmic tail
DAB	3,3'-diaminobenzidine
DAPI	4,6-diamidino-2-phenylindole
DCIS	Ductal carcinoma <i>in situ</i>
Dis	Disintegrin-like domain
DMEM	Dulbecco's modified eagle medium
DMEM/F12	Dulbecco's modified eagle medium/Nutrient mixture F-12
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPX	Distyrene-tricresyl
DTT	Dithiothreitol
dNTP	Deoxynucleoside triphosphate
ECM	Extracellular matrix

EDA/B	Extr domain A/B
EDTA	Ethylendiaminetetraacetic acid
EGF	Epidermal growth factor
EMT	Epithelial mesenchymal transition
EP300	E1A binding protein p300
EpCAM	Epithelial cell adhesion molecule
ER	Oestrogen receptor
EtOH	Ethanol
FACS	Fluorescence-activated cell sorting
FAK	Focal adhesion kinase
FBS	Foetal bovine serum
FDR	False discovery rate
FGF	Fibroblast growth factor
FKBP12	FKBP Prolyl Isomerase 1A
FLT-4	Fms related receptor tyrosine kinase 4
FN1	Fibronectin
FOX	Forkhead box transcription factor
FSP1	Fibroblast-specific protein 1
Fu	Furin
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase

GFP	Green fluorescent protein
GPI	Glycosylphosphatidylinositol
GSEA	Gene set enrichment analysis
GSTM1	Glutathione s-transferase mu 1
GuHCl	Guanidine hydrochloride
h	Hour(s)
H2B	Histone 2B
HB2	Luminal cell line
HCl	Hydrochloric acid
HEK293T	Human embryonic kidney cell line
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2	Human epidermal growth factor receptor 2
HGF	Hepatocyte growth factor
HLADR	Human leukocyte antigen D-related receptor
HMS-1-6	Myoepithelial tumour cell line series
HPG-ALD	Hyperbranched aldehyde-derivatized polyglycerol polymer
HRP	Horse radish peroxidase
HSC70	Heat shock protein family A
hTERT	Human telomerase
HuMEC	Human mammary epithelial cell

IDC	Invasive ductal carcinoma
Ig-like	Immunoglobulin-like
IL	Interleukin
IP	Immunoprecipitation
IRES	Internal ribosome entry site
ITGB6	Integrin subunit beta 6
JNK	Mitogen-activated protein kinase 8
Ki67	Marker of proliferation ki67
LAP	Latency associated peptide
LB	Luria Bertani
LCIS	Lobular carcinoma <i>in situ</i>
LLC	Large latent complex
LTBP	Latent TGF $\beta$ binding protein
MAPK	Mitogen-activated protein kinase 1
MCF10DCIS	Mammary epithelial MCF10A progression series cell line
MET	MET proto-oncogene receptor tyrosine kinase
min	Minute(s)
MIND	Mouse intraductal model
MMG	Mammography
MMP	Matrix metalloproteinase

MRI	Magnetic resonance imaging
mRNA	Messenger RNA
MT-MMP	Membrane type matrix metalloproteinase
MYBL2	MYB proto-oncogene like 2
Myo	Primary myoepithelial cell
MYO10	Myosin 10
NaBH <sub>3</sub> CN	Sodium cyanoborohydride
NaOH	Sodium hydroxide
NES	Normalised enrichment score
NF-κB	Nuclear factor κ B
NH <sub>4</sub> HCO <sub>3</sub>	Ammonium bicarbonate
NHS	National Health Service
NIH	National Institute of Health
NSG	NOD scid gamma mouse
NTC	Non targeting control
P15	Cyclin dependent kinase inhibitor 2B
P16	Cyclin dependent kinase inhibitor 2A
P21	Cyclin dependent kinase inhibitor 1A
P38	Mitogen-activated protein kinase 14
PBS	Phosphate buffered saline

PBST	PBS-Tween
PDL-1	Programmed cell death ligand
PI3KCA	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PLAC	Protease and lacunin
PMSF	Phenylmethanesulfonylfluoride
PR	Progesterone receptor
Pro	Zymogenic pro-peptide
qPCR	Quantitative polymerase chain reaction
RAHBT	Resource Archival Human Breast Tissue
RAS	RAS Proto-oncogene
REC	Research ethics committee
RFP	Red fluorescent protein
RGD	Arg-Gly-Asp
RIN	RNA integrity number
RNAseq	RNA sequencing
RT	Room temperature
RTK	Receptor tyrosine kinase
rtTA	Reverse tetracycline transactivator
SBE	Smad binding element
SDS	Sodium dodecyl sulphate

SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEMA7A	Semaphorin 7A
SH	Thiol
SiRNA	Small interfering RNA
SLC	Small latent complex
SNAI1	Snail family transcriptional repressor 1
SOX	SRY-box transcription factor
SP	Signal peptide
Spa	Spacer region
SRC	SRC Proto-oncogene
STAT3	Signal transducer and activator of transcription 3
STRING	Search tool for the retrieval of interacting genes/proteins
SV40 TAG	Simian virus 40 large-tumour antigen
TAE	Tris acetate-EDTA
TAILS	Terminal amine isotopic labelling of substrates
TBS	Tris buffered saline
TBST	TBS-Tween
TCEP	Tris (2-carboxyethyl) phosphine
TDLU	Terminal ductal lobular units
TE	Tris-EDTA

TGF $\beta$ (R)	Transforming growth factor $\beta$ (receptor)
TIL	Tumour infiltrating lymphocyte
TIMP	Tissue inhibitors of metalloproteinases
TM	Transmembrane
TMT	Tandem mass tag
TNBC	Triple negative breast cancer
TNF $\alpha$	Tumour necrosis factor $\alpha$
TP53	Tumour protein 53
TRE	Tetracycline response element
TS	Thrombospondin repeat
TSP-1	Thrombospondin 1
UK	United Kingdom
US	Ultrasound
V	Variable region
VEGF(R)	Vascular endothelial growth factor (receptor)
VIM1	Vimentin
Vn	Vitronectin-like insert
Wnt-1	Wnt family member 1
ZEB1/2	Zinc finger E-box binding homeobox 1/2
Zn	Zinc binding site



$\alpha$ SMA	$\alpha$ smooth muscle actin
$\kappa$	Kappa
$\alpha$	Alpha
$\beta$ TD	$\beta$ tail domain
$\beta$	Beta
$\gamma$	Gamma
1089	Myoepithelial cell line

## **Chapter 1. Introduction**

## **1.1. Breast Cancer**

### **1.1.1. Epidemiology and risk factors**

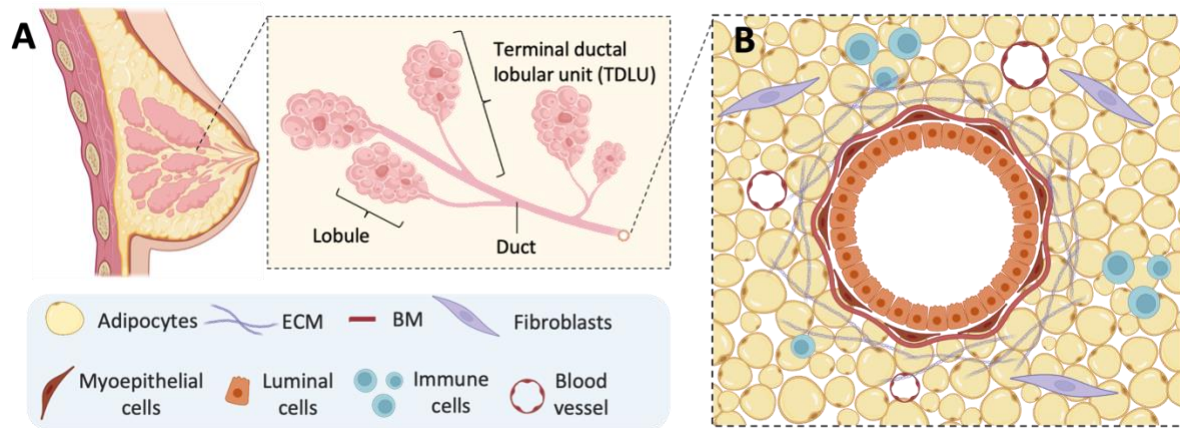
Breast cancer remains the most diagnosed cancer in women, with a yearly incidence of over 1.3 million, accounting for 23% of all malignancies in the Western world (1, 2). In the UK alone, approximately 55,000 women are diagnosed per year, equating to a 1 in 7 chance of a woman being diagnosed during their lifetime (1, 3). Breast cancer can also arise in men, with approximately 350 cases of male breast cancer diagnosed annually in the UK (3). Breast cancer remains the fourth most common cause of cancer death, with approximately 11,500 women who die from their disease each year. As with most cancers, breast cancer presents a multifactorial aetiology. As such, there are a variety of risk factors including age, weight, diet, alcohol consumption, family history and breast density (4).

Despite its high incidence, breast cancer has the second highest age-standardised ten-year net survival of the ten most common cancers in females (3, 5). It is predicted that around 3 in 4 women diagnosed with breast cancer survive their disease for ten years or more, with breast cancer survival doubling over the last 40 years in the UK (3, 5). This can be attributed to improvements in treatment and care, as well as earlier detection through the introduction of screening programmes. When diagnosed at its earliest stage, almost all women (98%) with breast cancer will survive their disease for five years or more, compared to only 25% of women whose disease is diagnosed at more advanced stages (3, 5).

### 1.1.2. Cellular architecture of the breast

The healthy breast consists of a branched epithelial ductal network that functions to drain milk to the nipple. Each breast contains 15-20 lobes that surround the nipple in a radial manner. These lobes consist of smaller sections, termed lobules, which contain secretory units called acini. The lobules, along with their associated ducts, form terminal ductal lobular units (TDLU), which make up the functional basic secretory unit of the breast (Fig 1.1) (6).

The breast duct consists of two differentiated cell types that are organised into a bilayer: an inner monolayer of luminal cells and a surrounding outer layer of myoepithelial cells. While luminal cells secrete milk proteins, myoepithelial cells deposit the basement membrane (BM), induce luminal cell polarity, and contract the ducts to eject milk during lactation. The BM, rich in laminin-I and collagen-IV, encircles the myoepithelial cells and separates the ducts from the adjacent stroma. The ductal network is surrounded by adipose tissue and other cell types in the breast, including adipocytes, blood vessels, immune cells, and fibroblasts (Fig 1.1) (6, 7).

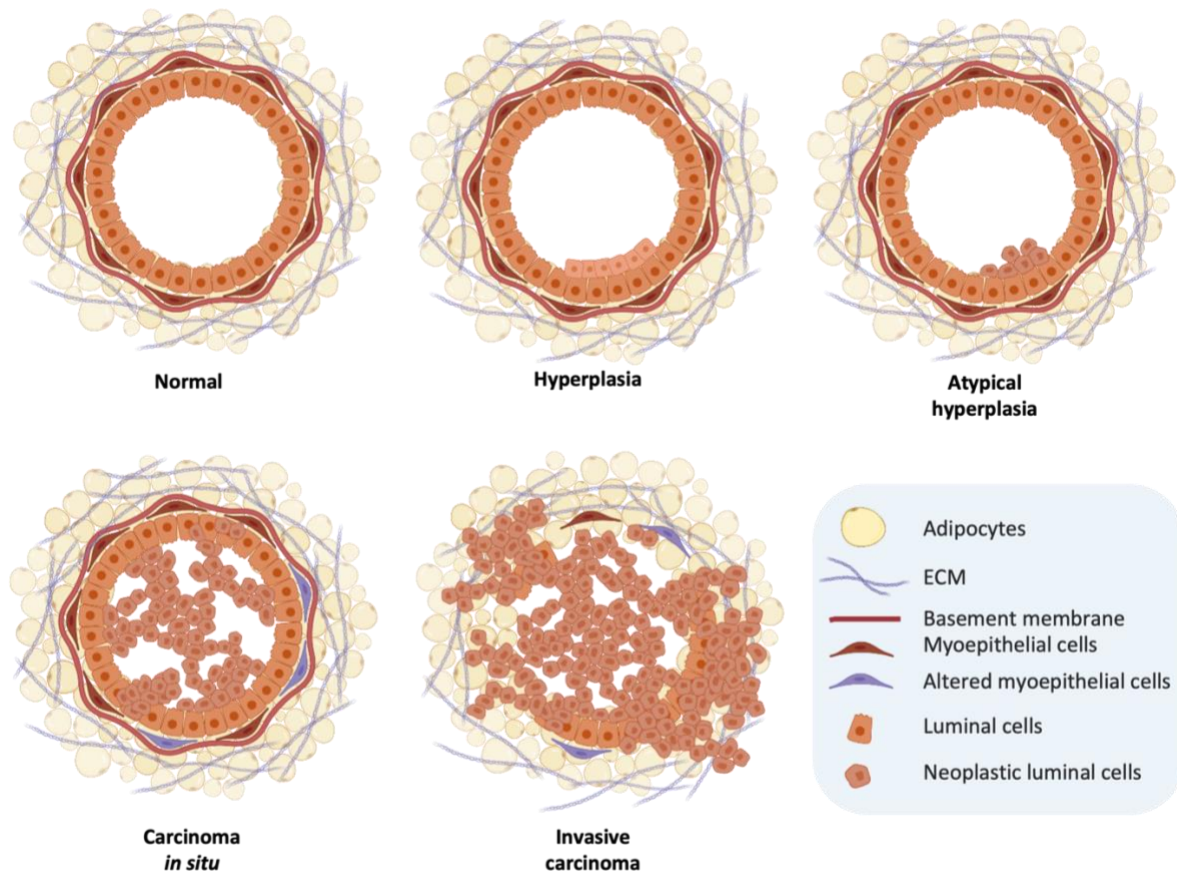


**Figure 1.1 Anatomy of the human breast.**

**(A)** Breast tissue contains epithelial ducts which connect secretory milk units, called lobules, to the nipple. **(B)** Ducts consist of inner luminal and outer myoepithelial cell layers, and are encircled by basement membrane (BM). The ductal network is surrounded by adipose tissue and other cell types, including immune cells, blood vessels and fibroblasts. Created with BioRender.

### 1.1.3. Breast cancer development and progression

Breast cancer progression follows a stepwise transition through pathologically defined stages (8-10). The classical model suggests that breast cancer first initiates as a benign epithelial lesion, which transitions into an overgrowth of luminal cells lining the ducts (ductal hyperplasia) or lobules (lobular hyperplasia) (Fig 1.2). At this stage, the hyperplasia is described as 'usual,' with cells maintaining original histological features. Further abnormal proliferation promotes progression to atypical hyperplasia, where cells become morphologically distinct from their healthy counterparts. From this stage, cells proliferate further to form carcinoma *in situ*, characterised by the luminal filling of neoplastic cells contained within a myoepithelial-BM barrier (Fig 1.2). Depending on the cells of origin, the *in situ* stage can be divided into either ductal carcinoma *in situ* (DCIS) or lobular carcinoma *in situ* (LCIS), accounting for 88% and 9% of *in situ* carcinomas, respectively (3, 11). While the *in situ* stage is not life threatening, non-obligatory progression of the disease can drive transition to invasive ductal carcinoma (IDC), which describes a group of malignant breast cancers. At this stage, the myoepithelial-BM barrier is lost, and cancer cells have invaded into the surrounding breast tissue (Fig 1.2) (8-10).



**Figure 1.2 Classical model of breast cancer progression.**

Neoplastic evolution initiates hyperplasia in the luminal epithelium. This progresses to atypical hyperplasia and can evolve to ductal (DCIS) or lobular (LCIS) carcinoma *in situ*. At this stage, neoplastic cells are still confined to the ductal space by an intact myoepithelial-basement membrane barrier. From carcinoma *in situ*, the disease can advance to invasive carcinoma, which is characterised by a breach in the basement membrane and invasion of cancer cells into the surrounding stroma. Created with BioRender.

#### 1.1.4. Classification of breast cancer

As a highly heterogeneous disease, the classification of breast cancer into clinically relevant subtypes is important to inform treatment strategies. For example, lesions can be categorised into grades based on the extent of abnormality and differentiation when compared to their healthy tissue counterparts. While low-grade lesions comprise of undifferentiated cells with small nuclei, high-grade lesions are characterised by highly atypical and differentiated cells (12, 13). A staging system, from 0 to 4, is also utilised to classify breast cancer based on the extent of invasion into the surrounding breast tissue (13). Stage 0 is used for DCIS, while stages 1 and 2, or 3 and 4, are indicative of invasion into the surrounding breast tissue, with or without metastasis. In addition to staging and grades, the subtyping of breast cancer has also shifted to include the characterisation of molecular profiles (14, 15). Expression of specific biomarkers, such as the hormone receptors oestrogen receptor (ER) and progesterone receptor (PR), and the human epidermal growth factor receptor 2 (HER2), are routinely evaluated for histological examination in core biopsies by immunohistochemistry and/or fluorescence *in situ* hybridisation. Based on the presence or absence of these markers, breast cancer can be broadly classed into five intrinsic subtypes; luminal A, luminal B, HER2-enriched, triple negative (TNBC) and basal (Table 1.1) (14, 15).

While the luminal subtypes share expression of ER and/or PR, they are further divided into subtype A and B depending on their extent of proliferation, determined by levels of the proliferation marker Ki67. While luminal A, which accounts for approximately 40% of breast cancers, is associated with a more favourable prognosis, luminal B corresponds to approximately 20% of cases, and tends to have a higher proliferation rate, correlating with a slightly worse prognosis (16, 17). The HER2-enriched subtype, accounting for 23-30% of cases, exhibits overexpression of HER2 and lacks expression of the hormone receptors. Finally, 15% of cases exhibit no expression of ER, PR or HER2 and can be grouped



into TNBC or basal subtypes. Despite similarities, TNBC and basal subtypes are not identical and are further categorised based on the expression of cytokeratins (CK) 5 and 6. Compared to the luminal subtypes, HER2-enriched, TNBC and basal subtypes correlate with disease that is more aggressive with a higher tendency to relapse (Table 1.1) (17, 18) .

**Table 1.1 Intrinsic classification of breast cancer subtypes**

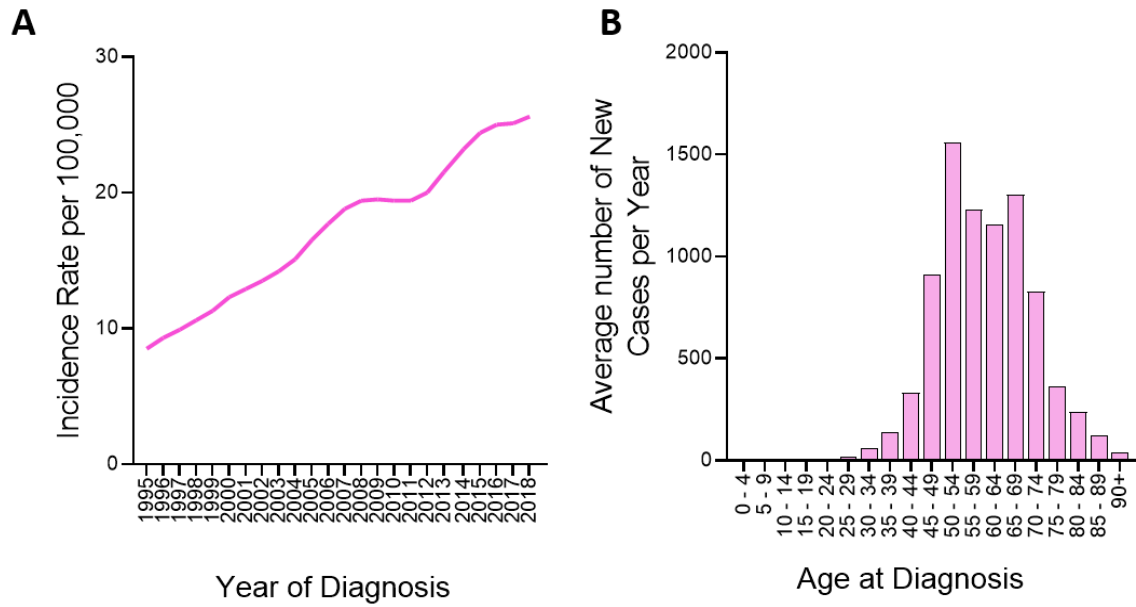
Subtype	Expression profile	Histological grade	Prognosis
Luminal A	ER <sup>+</sup> / PR <sup>+</sup> / HER2 <sup>-</sup> / Ki67 <sup>-</sup>	1/2	Good
Luminal B	ER <sup>+</sup> /PR <sup>+</sup> / HER2 <sup>-</sup> / Ki67 <sup>+</sup>	2/3	Intermediate
	ER <sup>+</sup> /PR <sup>+</sup> / HER2 <sup>+</sup> / Ki67 <sup>+</sup>		Poor
HER2-enriched	ER <sup>-</sup> / PR <sup>-</sup> / HER2 <sup>+</sup>	2/3	Poor
Triple negative	ER <sup>-</sup> / PR <sup>-</sup> / HER2 <sup>-</sup> / CK5/6 <sup>-</sup>	3	Poor
Basal	ER <sup>-</sup> / PR <sup>-</sup> / HER2 <sup>-</sup> / CK5/6 <sup>+</sup>	3	Poor

Table adapted from Dai *et al.* (18).

## 1.2. DCIS

### 1.2.1. Screening and current treatment guidelines

The aim of screening programmes is to prevent disease-specific mortality through detection and treatment of the disease at its earliest stage – making DCIS an ideal target in the prevention of subsequent IDC. In the UK, the NHS breast-screening programme invites women between the ages of 50 and 70 to mammographic screening every three years. The implementation of population-based screening has caused a 200% increase in the observed incidence of DCIS since the 1990s, with approximately 8,000 cases currently being diagnosed annually in the UK (Fig 1.3) (19). Upon detection through screening, the DCIS lesion is classified by a pathologist into low, intermediate, or high histological grades, based on the degree of cell differentiation and proliferation. Regardless of grade, current guidelines dictate surgical resection as the standard treatment for all women with a DCIS diagnosis. For limited DCIS, breast-conserving surgery is most common, and consists of a lumpectomy with adjuvant radiotherapy. With radiotherapy reducing risk of recurrence by up to 50% (20, 21), this treatment strategy has a recurrence risk of 12% (22, 23). For more extensive DCIS, women may undergo mastectomy, which has a 1-2% risk of recurrence (22, 23).



**Figure 1.3 Breast carcinoma *in situ* trends over time.**

**(A)** Carcinoma *in situ* incidence rate per 100,000 in the UK from 1995 until 2018. **(B)** Age-specific incidence of carcinoma *in situ* cases in the UK, 2018. Data taken from CRUK statistics (19).

Despite being a precursor lesion, epidemiological and observational studies estimate that between 30 to 50% of diagnosed DCIS will not progress to invasive disease within a woman's lifetime (24-30). For example, a study which followed 28 women with low-grade DCIS, who underwent biopsy only, showed that 11 (39%) cases developed invasive disease. While seven (25%) were diagnosed with IDC within 10 years, one (4%) was diagnosed within 12 years and the remaining three (11%) were diagnosed within 23 and 42 years (28). In a comparable study, consisting of 89 DCIS patients who underwent surgical resection or for whom surgery was delayed for at least 12 months, 29 (33%) women developed IDC after a median interval of 45 months (29). Grading of DCIS lesions within this cohort indicated that 3 of 17 (18%) low-grade cases developed to IDC at a median interval of 51 months, while 14 of 29 (48%) high-grade cases developed to IDC at 38 months (29). This was supported by the subsequent Forget-Me-Not 2 study of 311 eligible women, where 60 (19%) developed IDC, with the 10-year risk of invasive disease progression being 9%, 39% and 36% for low, intermediate and high-grade DCIS lesions, respectively (30).

While these studies suggest that high-grade DCIS lesions display quicker progression to invasive disease, they also highlight that all cases have the potential to progress, irrespective of grade. Thus, the DCIS grading system cannot be used to reliably differentiate between those that will progress to the invasive stage.

### **1.2.2. Overdiagnosis and overtreatment**

Despite a good prognosis and normal life-expectancy, a DCIS diagnosis labels women as those who are at risk of developing invasive cancer. As such, those diagnosed may overestimate the implications of their diagnosis and experience substantial levels of stress and anxiety. This raises critical questions

regarding how the communication of a DCIS diagnosis can be improved to mitigate labelling effects associated with subsequent psychological distress (31).

Concerns regarding the overtreatment of these women have also been raised as, if left untreated, most DCIS patients will not exhibit symptoms or die of their disease (23). As the incidence of early-stage breast cancer mortality has not changed since the introduction of screening, this further implies that the management of DCIS does not in fact reduce breast cancer mortality. Taken together, these issues highlight the critical need to identify robust markers, which predict the likelihood of invasive and recurrent disease following a DCIS diagnosis.

The clinical importance of tackling DCIS overtreatment was raised at the NIH State-of-Science conference on DCIS, which highlighted the need to improve our understanding of the molecular mechanisms underlying the transition of DCIS to IDC, to aid the development of prognostic markers for patient stratification (32). Supporting this, a £15 million Cancer Grand Challenge was awarded to PRECISION (PREvent ductal Carcinoma In Situ Invasive Overtreatment Now) in 2015, an international initiative that aims to establish markers to reduce the burden of unnecessary DCIS treatment (33).

### **1.2.3. Tools for predicting risk of recurrence**

To address the issues of overdiagnosis and overtreatment, prediction tools have been developed in an attempt to stratify patients and guide clinical decisions. While an increased tendency for high-grade DCIS to progress to IDC has been implied (34), others have demonstrated that grade is not significantly associated with increased risk of invasive disease (26, 35). In 1995, The Van Nuys Prognostic Index was developed for DCIS in attempt to predict recurrence. This classification system is based on the age of

the patient at diagnosis and includes the tumour features: size, margins, nuclear grade, and the presence of necrosis. Each of the features is assigned a score from 1 to 3, where 1 is the most clinically favourable and 3 constitutes the most unfavourable. A patient's DCIS lesion is then scored between 4 and 12, with a score of 12 indicative of most likely to recur (36). While low scoring DCIS lesions are typically present in older patients with small, low-grade lesions, high scoring DCIS lesions are more often present in younger patients with large, high-grade lesions (36). While the method can be helpful with patient assessment, it has been deemed as non-optimal, due to its inability to predict recurrence as either DCIS or invasive cancer (37, 38).

Another prediction tool is the Oncotype DCIS score, a multi-gene expression assay adapted from the original Oncotype test, which is used to predict chemotherapy benefit in ER+ IDC patients. The aim of the modified Oncotype DCIS score is to identify low-risk patients who could forego radiotherapy following surgery (39, 40). Based on the expression of seven cancer related (Ki67, CCNB1, MYBL2, AURKA, BIRC5, PR and GSTM1) and five reference genes, patients are grouped into three risk categories: high-risk (score  $\geq 55$ ), intermediate-risk (score = 39-54) and low-risk ( $<39$ ). Initial validation of the method in a cohort of 327 patients who underwent surgical excision without radiation, displayed a recurrence risk of 26%, 27% and 11%, for high, medium and low-risk patients, respectively (40). The method has since been validated in a larger cohort of 718 cases, where its score was also associated with recurrence. Like the initial validation study, a respective 10-year risk of 28%, 33% and 13% were reported, for high, medium and low-risk patients (41). While these studies indicate a predictive value for the Oncotype DCIS score, whether it can guide the need for radiotherapy, for which the test was initially intended, remains largely unexplored. Another limitation of this test is that selection of genes for the algorithm relied on studies of IDC, rather than DCIS. Currently, the NHS does not offer the test routinely.

In another study, a tumour-intrinsic biomarker panel, used to predict recurrence risk, was identified using a cohort of 324 patients who later developed ipsilateral recurrent disease (42). Those with DCIS lesions detected by palpation that were also p16, COX-2 and Ki67 triple positive, had a significantly higher 8-year risk of developing recurrent invasive cancer. This was attributed to p16 being indicative of a more aggressive phenotype, COX-2 expression associating with invasive potential, and Ki67 reflecting higher rates of proliferation (42). These findings have since been utilised in developing the DCISionRT test, another tool aimed at guiding radiotherapy decision making. However, while it seems promising, additional evaluation is required in clinical trials to investigate its robustness in guiding the use of radiotherapy post lumpectomy (24).

#### **1.2.4. Active surveillance trials for DCIS**

The long-term benefit of treating asymptomatic DCIS that may or may not progress to IDC is difficult to evaluate. Thus, to study the natural course of DCIS progression, long-term monitoring of patients who forgo surgery and/or radiotherapy is the only way to determine the requirement for treatment interventions following a DCIS diagnosis. However, with current strategies presenting as an effectively curative treatment, this makes recruiting women onto surveillance trials challenging. Despite this, there are four ongoing surveillance trials for DCIS: COMET (Comparison of Operation to Monitoring, with or without Endocrine Therapy, NCT02926911), LORD (LOW Risk DCIS, NCT02492607), LORETTA (Single-arm confirmatory trial of endocrine therapy alone for ER positive, low-Risk DCIS, JCOG1505) and LORIS (Surgery Versus Active Monitoring for LOW RISK DCIS, NCT02766881). These trials aim to compare conventional treatment with active surveillance, where the primary end point is the development of ipsilateral invasive cancer. While the results of these trials will enable the natural history of breast cancer to be followed, the study outcomes remain to be reported (43-46).



**Table 1.2 Summary of active surveillance trials**

<b>Study</b>	<b>Participant target</b>	<b>Participant age (years)</b>	<b>Design</b>	<b>Endocrine therapy</b>	<b>Screening</b>
COMET	1200	≥ 40	Randomised	Permitted	MMG
LORD	2500	≥ 45	Patient preference	-	MMG
LORETTA	340	40-75	Single arm	Permitted	MMG, US, MRI
LORIS	932	≥ 48	Randomised	-	MMG

MMG; mammography, US; ultrasound, MRI; magnetic resonance imaging. Table adapted from Casasent *et al.*, and Angarita *et al.* (24, 46).

### 1.2.5. Comparative genomics of DCIS and IDC

Previous efforts to identify predictive markers have focused primarily on comparing the tumour cells in DCIS with those of IDC (47-49). These studies have yielded limited information, reporting negligible differences in the signature of cancer cells throughout disease progression. While DCIS and IDC are heterogeneous within their own right, comparative genomic studies show that both disease stages share surprisingly similar molecular profiles (24). For example, copy number alterations show almost no difference between DCIS and IDC (24, 50), with the most common driver mutations in *TP53* and *PIK3CA*, already detected in DCIS (51, 52). Epigenetic alterations, such as methylation patterns, are also consistent between the two disease states (24, 50). For example, SOX-family transcription factors, whose genes are hyper-methylated in DCIS, are also downregulated in IDC (15, 24). *In situ* gene expression profiles of the malignant epithelial compartment also reveal extensive similarities in the transcriptomic signature between pre-invasive and invasive stages, demonstrating that cells from DCIS present similar profiles to their invasive counterparts (48). Taken together, these studies highlight that, although a precursor lesion, cancer cells from DCIS are genetically indistinguishable from their IDC counterparts.

Interestingly, while gene expression profiles across DCIS and IDC cluster together according to molecular subtype (53), a unique subset of DCIS and IDC lesions has been shown to cluster based on expression of microenvironment-related genes, such as matrix metalloproteinases (MMPs) and extracellular matrix (ECM) proteins (53, 54). This, combined with a lack of distinct genetic changes between DCIS and IDC, raises the question of whether the most critical changes for DCIS progression lie in the surrounding microenvironment. Studies that report the requirement of a tumour-permissive stroma to facilitate local invasion have since supported this hypothesis (53, 55-58).

### 1.3. The tumour microenvironment in DCIS progression

#### 1.3.1. Overview

The following sections will focus on discussing the contributions of each cell type within the DCIS microenvironment, with an emphasis on identifying prognostic markers associated with disease progression.

#### 1.3.2. Myoepithelial cells

In the healthy breast duct, myoepithelial cells are positioned between the luminal cells and the stroma, acting as a physical barrier against luminal invasion. While myoepithelial cells support the ejection of milk during lactation by contracting, myoepithelial cells also contribute to BM deposition through the secretion of laminin and collagen-IV proteins, to regulate the polarity of luminal epithelial cells (59-64). Their tumour-suppressive nature can be emphasised by *ex vivo* DCIS studies where myoepithelial cells actively restrain and restrict the escape of neoplastic luminal cells into the surrounding stroma (65). Myoepithelial cells can also induce cancer cell cycle arrest and halt invasion *in vitro*, potentially through expression of endogenous protease inhibitors (66). As loss of the myoepithelium is a key determinant in distinguishing between DCIS and IDC (Fig 1.2) (67), there is extensive interest in the understanding the fate of myoepithelial cells as well as dissecting their role in modulating disease progression (59, 68).

Surprisingly, transcriptomic analysis of breast tissue cell types with cell-specific markers, reveals that gene expression changes in the microenvironment are already present at the DCIS stage. Strikingly,

DCIS-associated myoepithelial cells show the most significant transcriptomic change when compared to their healthy counterparts (69). Supporting this, spatial analysis of matched primary DCIS tissue from women who later showed recurrent DCIS with invasive disease, revealed integrity of the myoepithelium to be key in predicting disease recurrence. Counterintuitively, this study reported that discontinuous myoepithelium was more likely in the lesions of non-progressive cases, whilst lesions with more continuous myoepithelium were at higher risk of developing invasive recurrence (58). An interpretation of this finding is that DCIS-associated myoepithelial cells may become altered to lose their tumour-suppressive phenotype, implying their role as drivers of disease progression.

*In vitro* assays support an altered role for the DCIS myoepithelium, where reduced secretion of BM proteins, such as laminin-I, alters their ability to correctly polarise luminal cells (70). Tumour-promoting effects of the DCIS myoepithelium have also been displayed, where myoepithelial cells enhance cancer cell proliferation and migration through increased expression of the C-X-C motif chemokines ligands CXCL12 and CXCL14 (69). Increased secretion of myoepithelial transforming growth factor beta (TGFβ) can also drive upregulation of epithelial-mesenchymal transition (EMT)-related genes, such as FOXC2, SNAIL1, ZEB1/2 and VIM1 in the MCF10DCIS.com cell line model (71). Finally, co-injection of tumour-associated myoepithelial cells, together with cancer cells, can potentiate tumour growth in the xenograft models, further supporting their tumour promoting role (71, 72).

### **1.3.3. Adipocytes**

The contribution of the adipose microenvironment has received little attention until recently, where evidence now highlights their additional roles as cancer-associated adipocytes (CAAs). Tumour cell-

adipocyte crosstalk are reciprocal, with DCIS cells able to block adipogenic differentiation to induce expression of cancer-associated fibroblast (CAF) markers (73). CAAs then facilitate the survival of cancer cells, and enhance tumour growth in xenograft models (73). This is dependent on adipokine secretion of leptin, chemokine (C-C motif) ligand 5 (CCL5) and hepatocyte growth factor (HGF), as well as proteases, including MMP11, which support the cancer cell invasion and dissemination into the stroma (74, 75). Adipocytes also promote neoplastic cells to adopt a stem cell phenotype (76, 77), where adipokines drive disease progression through SRC/SOX2 activation (76).

Adipocytes can also disrupt the tumour-suppressive myoepithelium by altering expression of the BM and integrins (75). In mouse models of obesity, an inverse correlation is observed between body mass index and number of myoepithelial cells in mammary fat pads (78). Elevated levels of leptin stimulate secretion of pro-inflammatory cytokines (TNF $\alpha$ , IL-1 $\beta$ , COX-2), which mediate chronic inflammation (79). Crown-like macrophage structures as a result of chronic inflammation are also present in the breast, and can be characterised by encircling TNF $\alpha$ -producing pro-inflammatory macrophages around necrotic adipocytes by (80). Interestingly, exposure of MCF10A cells to macrophage conditioned medium drives IKK $\epsilon$ /TBK1 dependent acquisition of malignant properties, and can be abrogated by inhibition of this pathway, resulting in delayed tumour formation in mouse models of obesity-associated breast cancer (81). While crown-like structures in healthy breast can be associated with increased breast cancer risk (80), their prognostic significance in DCIS is unclear (82, 83). Pathologic assessment of DCIS tissue with matched long-term follow up revealed that adipocyte features could be used to predict recurrent invasive disease (84). Specifically, adipocyte size associated significantly with recurrent ipsilateral IDC, with a significant odds ratio of 2.75 (84). When adipocyte size was combined with high COX-2 expression in DCIS cells, an additional marker identified (85), risk of subsequent ipsilateral IDC was significantly higher (84).

### 1.3.4. Immune cells

In invasive breast cancer, enhanced levels of tumour infiltrating lymphocytes (TILs) are associated with a more favourable prognosis and better treatment response. This is supported by TNBC trials, which demonstrate a synergistic effect between TIL density and chemotherapy response (86, 87). However, the significance of TILs in the context of DCIS is less clear, with contradictory results displayed (88-91). Interestingly, when TILs in contact with, or within one lymphocyte cell thickness away from the BM, are quantified, their density is indicative of a shorter time to DCIS recurrence (89). Thus, whilst TILs may be associated with enhanced immune-mediated tumour cell killing in the context of IDC, increased infiltration in DCIS could indicate early inflammation and remodelling of the surrounding stroma to facilitate invasion. Characterisation of infiltrating immune cell subpopulations identify higher infiltration of FOXP3+ T regulatory cells in DCIS associated with invasion, compared to pure DCIS (92), with shorter recurrence-free survival (88). Additionally, higher stromal expression of programmed cell death ligand (PDL-1), as well as low human leukocyte antigen D-related receptor (HLA-DR) are both implicated as prognostic markers of recurrence. Taken together, this highlights that even during the early stages of breast cancer progression, an enhanced immunosuppressive milieu can be established to drive invasion (88).

In addition to lymphocytes, several myeloid cell types are also associated with DCIS progression (93, 94). For example, HER2/NF- $\kappa$ B mediated cancer cell signalling can promote CCL2 dependent recruitment of CD206<sup>hi</sup> macrophages, which results in increased Wnt-1 secretion and subsequent loss of cancer cell E-cadherin (95). The prognostic relevance of macrophages can be highlighted by their increased abundance in DCIS, in high- versus low- grade cases (96). Interestingly, DCIS with high macrophage density is associated with a poorer prognosis, with the presence of peri-vascular macrophages indicative of both DCIS and ipsilateral invasive recurrence (94).

### 1.3.5. Vasculature

To facilitate tumour growth, increased blood vessel formation is required to supply oxygen and nutrition to fast-proliferating cancer cells. In breast cancer, vascular endothelial growth factor (VEGF) expression and microvessel density are indicators of poor prognosis (97, 98). In DCIS, two distinct patterns of blood vessel formation have been identified. The first pattern is a diffuse stromal pattern, where there is an increase in blood vessels in the surrounding tissue. The second pattern is a periductal pattern, where blood vessels surround the ducts affected by DCIS (97, 99, 100). The presence of the periductal pattern has been associated with DCIS that has an invasive component, suggesting that the formation of blood vessels around the ducts may facilitate local invasion of cancer cells (101). However, conflicting results have been observed upon transcriptomic analysis of angiogenic markers, displaying increased expression of pro-angiogenic/lymphangiogenic factors (VEGF-C, FLT-4 and VEGFR2) in pure DCIS, versus DCIS with coexisting invasion (102). This could be due to distinct differences in the vascular phenotype with distinct pre-existing angiogenic patterns at the DCIS stage, before it evolves to the invasive stage (102). The idea that DCIS and invasive disease can have different angiogenic phenotypes is supported by the correlation between stromal vascularity and the perivascular expression of a pro-angiogenic regulator, thymidine phosphorylase (103), in DCIS patients and not patients with IDC.

Studies using a series of benign myoepithelial tumour cell lines (HMS-1-6) highlight that these cells can inhibit tumour-induced angiogenesis in xenograft models. This may be attributed to their low levels of pro-angiogenic factors and high expression of angiogenic inhibitors (104). However, conflicting results were reported when analysing gene expression and DNA methylation in different breast cell populations. Specifically, myoepithelial cells from DCIS exhibit a down-regulation of anti-angiogenic thrombospondin-1 (TSP-1) and up-regulation of pro-angiogenic CXCL12, compared to their

normal counterparts (69, 105). Additionally, an up-regulation of the pro-angiogenic urokinase system has been described in myoepithelial cells associated with DCIS (106). These findings suggest that myoepithelial cells may have the ability to switch from an anti-angiogenic state to a pro-angiogenic state, potentially influencing the angiogenic microenvironment. The phenotype is possibly dependent on the stage of progression, with angiogenic markers holding promise as a potential means to identify patients at higher risk of DCIS progression.

### 1.3.6. Fibroblasts

CAFs make up a critical component of the tumour microenvironment, where they mediate matrix deposition and remodelling, cancer cell growth, and immune cell recruitment (107, 108). In breast cancer, increased CAF density is associated with an aggressive phenotype, partly due to their ability to induce stemness and chemoresistance via activation of PI3K/AKT and Wnt pathways in tumour cells (109). CAFs can also facilitate a milieu associated with immunosuppression by promoting differentiation of FOXP3<sup>+</sup> T regulatory lymphocytes (110). While the roles of IDC-associated CAFs are better defined, it remains unclear whether they possess the same functions in the early stages of breast cancer. Interestingly, comparison of tissue from DCIS and IDC cohorts reveals that, aside from myoepithelial loss, an increase in proliferative CAFs is also a distinctive feature (58).

While heterotypic co-culture of luminal, myoepithelial and fibroblast populations, isolated from the healthy breast, lead to luminal-myoepithelial co-unit formation, co-culture with tumour-associated fibroblasts prevents this phenotype, indicative of a disruption to luminal-myoepithelial crosstalk (111). *In vitro* studies demonstrate that fibroblasts promote DCIS invasion through activation of NF- $\kappa$ B and COX-2, and up-regulation of MMPs 9 and 14 (112). Additionally, fibroblast-secreted IL-6 can



support tumour cell invasion through the ECM via paracrine-mediated mechanisms (113). Moreover, murine CAFs isolated from invasive carcinoma express higher levels of Cxcl1 compared to those isolated from slow-growing non-invasive lesions. This enhanced Cxcl1 expression promotes tumour cell invasion through activation of Cxcr2-dependent signalling pathways involving of Nf-kb, Akt, Stat3 and Mapk (114). *In vivo* studies have supported the tumour-promoting role of fibroblasts, where co-injection of tumour cells combined with fibroblasts leads to increased invasiveness, compared to injection of tumour cells alone (112, 115). Patient follow-up studies have shown that upregulation of the fibroblast marker platelet derived growth factor receptor  $\beta$  (PDGFR $\beta$ ), regulated by the Notch pathway, is associated with increased risk of DCIS recurrence (116). The loss of stromal caveolin-1, a characteristic feature of breast cancer CAFs, is also associated with DCIS progression to IDC (117).

### 1.3.7. Matrix

The ECM is a complex network of macromolecules that modulates cell behaviour through its biochemical and biomechanical properties. While the transition of DCIS to IDC can be partly defined by BM loss, little is known about how cancer cells may alter BM assembly in DCIS. In the MCF10DCIS.com cell line model, increased expression of the filipodia-inducing protein, myosin-10 (MYO10), correlates with prominent BM deposition and an anti-invasive phenotype (118). Interestingly, depletion of MYO10 in DCIS xenograft models leads to defective BM assembly and cancer cell dispersal, suggesting that inadequate BM assembly, rather than BM degradation, may facilitate DCIS progression (118). However, it is unclear whether the increased cancer cell dissemination in MYO10-deficient tumours is directly due to a compromised BM or a gain in invasive capacity (118). Further dysregulation of the ECM, demonstrated by increased collagen deposition in the peritumoural regions, is also crucial in shaping both tumour growth and metastatic dissemination (119, 120). The mammary gland becomes progressively stiffer and collagen-rich in breast cancer, with

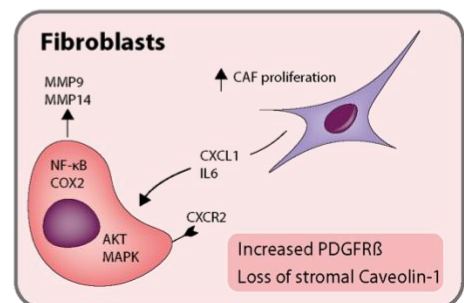
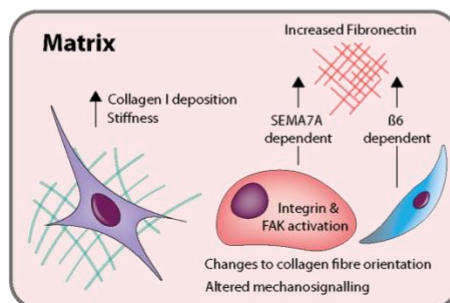
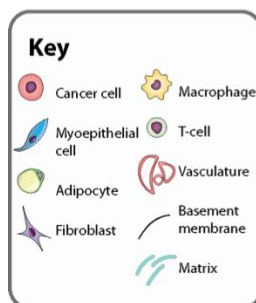
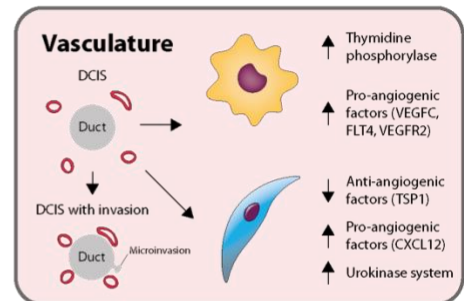
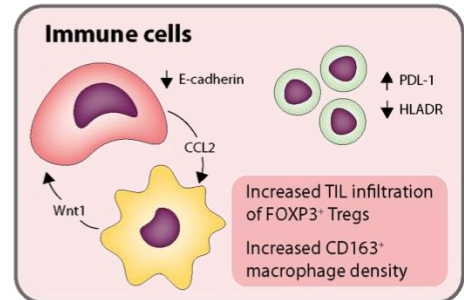
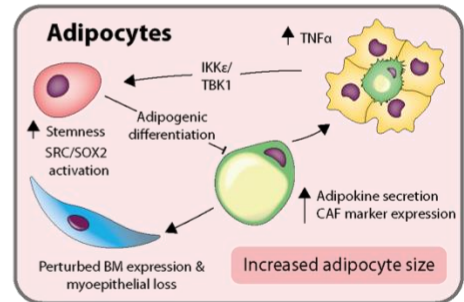
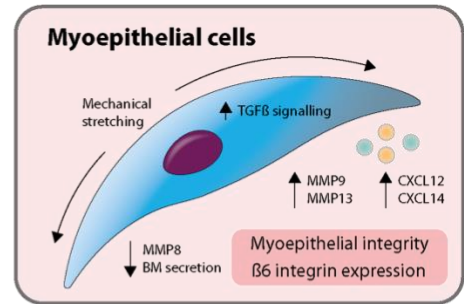
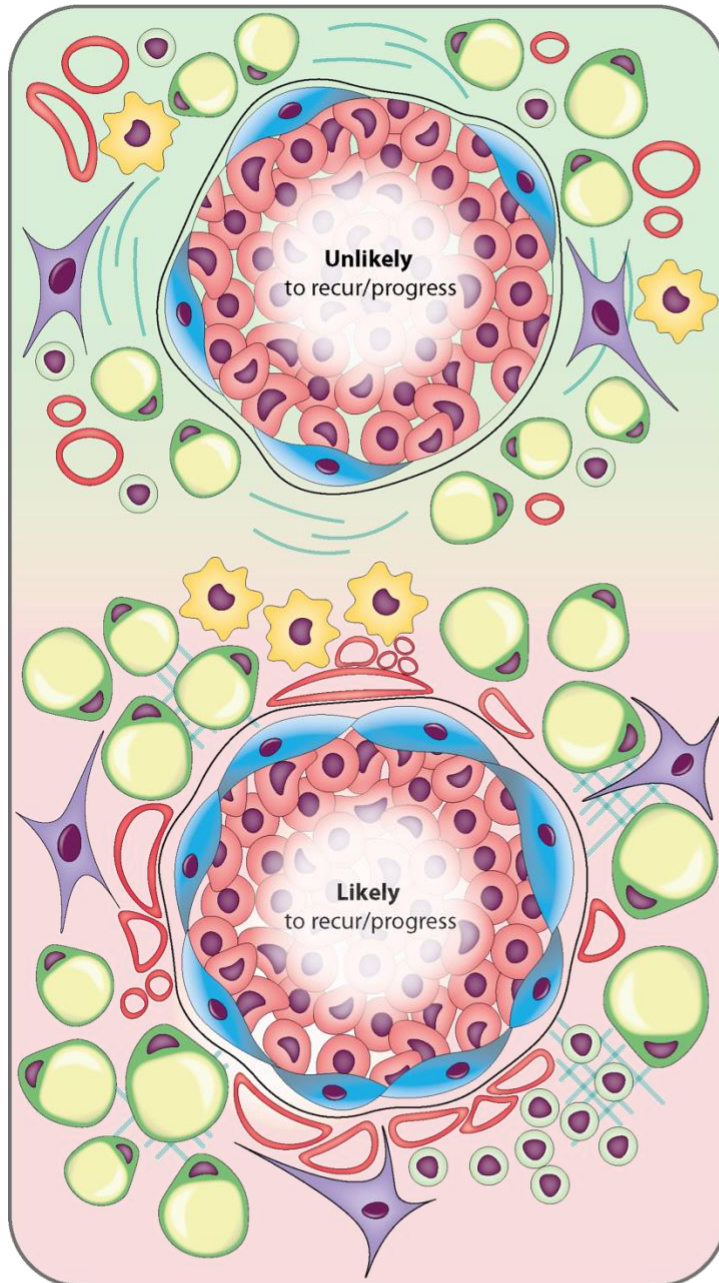
enhanced changes to matrix remodelling attributed predominantly to CAFs. Indeed, tissue stiffness correlates with cancer grade, and the presence of thick aligned collagen fibres at the invasive front of tumours is associated with worse patient outcome (121). The orientation of collagen fibres is also altered during disease, progressing from being adjacent to the tumour margin in DCIS to a perpendicular orientation in IDC (121-124). These changes enable the formation of tracks to facilitate cancer cell migration into the surrounding stroma (122). While perpendicular collagen fibres are more frequently present in DCIS displaying poor prognostic markers, such as comedo necrosis and HER2 positivity, changes to these orientation patterns are cannot predict likelihood of recurrence (125). Temporal profiling of the breast cancer matrisome, which consists of ECM and ECM-associated genes, identifies CAF-secreted collagen-XII as a fundamental component in collagen-I organisation, with increased secretion associated with poor outcome (126). Although collagen-XII can create a more pro-invasive microenvironment, further investigation is needed to determine its potential as a marker for DCIS stratification (126).

Other matrix proteins, such as fibronectin, are also implicated in promoting early-stage breast cancer progression. Deposition of periductal fibronectin is increased in DCIS compared to healthy ducts and is thought to support a pro-invasive niche (127). In addition to CAFs (128), fibronectin can also be deposited by other cell types. While cancer cells deposit fibronectin in a SEMA7A-mediated manner (129), DCIS-associated myoepithelial cells increase its expression upon mechanical stretching (127). In addition to the mechanical influence of the matrix on cancer cells, the collagen-dense microenvironment can provide anchorage to infiltrating immune cells, such as tumour-associated neutrophils, which may further facilitate disease progression (130). Mechanosensors, such as integrins, enable cancer cells to respond to changes in the mechanical properties and facilitate transition to an invasive phenotype (131). Indeed, increased integrin  $\beta 1$  activation and FAK<sup>Y397</sup> phosphorylation have been associated with invasion in study with MCF10A mammary spheroids (132).

While a progressive increase in the levels and co-localisation of integrin  $\beta 1$  and  $\text{pho-FAK}^{\text{Y397}}$  was observed across normal, DCIS and IDC clinical samples (132), protein scoring is required to confirm their association with disease recurrence.

### **1.3.8. Potential tumour-extrinsic markers of DCIS progression**

As the role of the tumour microenvironment becomes increasingly recognised, it is likely that tumour-extrinsic markers will be necessary (Fig. 1.4, Table 1.3, Table 1.4), along with conventional histopathological factors, to guide treatment plans for DCIS. To effectively translate potential prognostic markers into clinical practice, further understanding of the microenvironmental drivers that underpin DCIS progression will be key.



**Figure 1.4 Role of the DCIS microenvironment in facilitating DCIS progression.**

The schematic illustrates cell type-specific molecular mechanisms, that could serve as potential stratification markers associated with indolent DCIS (Unlikely) or DCIS progression (Likely). Mechanical stretching of the myoepithelium in DCIS leads to the upregulation of integrin  $\beta_6$ , activation of TGF $\beta$  signalling and changes in expression of basement membrane (BM) proteins, proteases and chemokines. Cancer cells can inhibit adipocytic differentiation to increase adipokine secretion and their expression of cancer associated fibroblast (CAF) markers. This interaction may promote stem-like characteristics in cancer cells through SRC/SOX-2 pathway activation. Adipocytes also influence myoepithelial cell fate, altering BM protein expression. Inflammatory crosstalk in the breast adipose tissue stimulates TNF $\alpha$  secretion by inflammatory macrophages, which may promote malignant properties of DCIS cells through IKK $\epsilon$ /TBK1 pathways. Establishment of an immunosuppressive stroma, characterised by increased presence of FOXP3<sup>+</sup> tumour infiltrating lymphocytes (TILs), and respective high and low expression of PDL-1 and HLADR, may contribute to DCIS progression. Cancer cell-derived CCL2 increases macrophage recruitment to drive increased Wnt-1 secretion and loss of E-cadherin in cancer cells. Fibroblast-cancer cell mediated crosstalk triggers activation of NF- $\kappa$ B and COX-2, and up-regulation of MMPs 9 and 14. Increased secretion of CXCL1 and IL-6 by fibroblasts can drive activation of AKT/MAPK pathways in DCIS cells. Alterations to collagen-I deposition, fibre orientation and matrix stiffness can occur due to fibroblast activation. Increased deposition of fibronectin by DCIS cells, dependent on SEMA7A signalling, and by myoepithelial cells through mechanotransduction, may also occur. The altered matrix drives changes in downstream mechanosignalling, particularly through integrin and FAK activation in DCIS cells. Different vascular patterns can be observed in DCIS, with a periductal pattern can be observed in DCIS with an invasive component. Interestingly, pure DCIS shows an upregulation of VEGF related proteins, suggesting a pre-existing angiogenic state before invasive progression. Myoepithelial cells in DCIS display reduced expression of anti-angiogenic factor TSP-1, and increased expression of pro-angiogenic factors, as well as upregulation of the urokinase system. The features of the tumour microenvironment highlighted in the associated boxes may indicate the likelihood of DCIS progression and invasive recurrence. Figure designed by S. Gibson and R. Roozitalab and constructed by R. Roozitalab using Illustrator Software.

**Table 1.3 Microenvironmental markers based on recurrence data**

Marker	Microenvironmental compartment	Prognostic relevance	References
<b>Established from patient cohorts with recurrence data</b>			
Myoepithelial integrity	Myoepithelial	Discontinuous myoepithelium may be present in the DCIS lesions of non-progressing patients, whilst more continuous myoepithelium may be at higher risk of invasive recurrence	(58)
Integrin $\beta 6$ expression	Myoepithelial	Increased myoepithelial integrin $\beta 6$ expression is associated with high-risk DCIS	(72)
Adipocyte size	Adipocyte	Large breast adipocyte size associated with ipsilateral invasive recurrence	(84)
Density of 'touching' TILs	Immune cells	Increased density of touching TILs is predictive of a shorter recurrence-free interval in DCIS	(89)
CD163 <sup>+</sup> macrophage density	Macrophages	Presence of CD163 <sup>+</sup> peri-vascular macrophages associated with increased likelihood of DCIS and ipsilateral invasive recurrence	(94)
PDGFR $\beta$ expression	Fibroblasts	Increased expression of PDGFR $\beta$ associated with increased risk of recurrence	(116)
Caveolin-1	Fibroblasts	Loss of stromal caveolin-1 may be associated with progression to IDC	(117)

**Table 1.4 Microenvironmental markers based on comparing disease states**

<b>Marker</b>	<b>Microenvironmental compartment</b>	<b>Prognostic relevance</b>	<b>References</b>
<b>Established from comparison between pure DCIS and DCIS with coexisting invasion or IDC</b>			
Periductal vasculature	Vasculature	Periductal pattern may be associated with DCIS with coexisting invasion	(101)
Stiffness and Collagen alignment	Matrix	Increased collagen-I deposition and perpendicular fibre orientation may be associated with invasion	(121-125)

## 1.4. Metzincins

### 1.4.1. Overview

In mammals, proteases are categorised into five distinct classes: cysteine, serine, metallo, aspartic and threonine. The metzincins belong to the metalloproteinase class and are named due to their conserved methionine residue within their active site. They are further divided into subfamilies consisting of MMPs, adamalysins, astacins, pappalysins and serralysins (133).

Metzincins play fundamental roles in modulating various biological processes due to their ability to degrade the ECM and facilitate invasion and metastasis (133, 134). However, recent studies have shown that these proteases cleave a range of additional substrates besides the matrix, to support the early stages of cancer progression (135, 136). Additionally, metzincins can possess both cancer-progressive and suppressive functions depending on context, ECM composition, and presence of other proteases within the tumour microenvironment (137-139). It is increasingly clear that metzincins are not simply degraders of the matrix, but possess more nuanced functions, acting as master orchestrators of cell signalling through the proteolytic processing of bio-active molecules at site-specific cleavage sites (139).



### 1.4.2. Classification and structure of metzincins

#### MMP

The MMPs comprise a family of 23 zinc-dependent endopeptidases, which can cleave virtually any component of the ECM. Initially, MMPs were categorised into subtypes based on their substrate specificity into collagenases, gelatinases, stromelysins and matrilysins. However, as the list has grown, along with an improved understanding of each MMP, they are now grouped into eight structural classes: five secreted and three membrane bound-MMPs (MT-MMPs) (Fig 1.5) (136, 137).

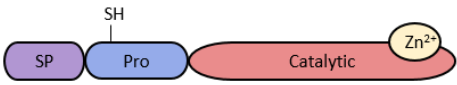
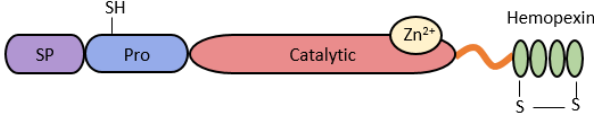
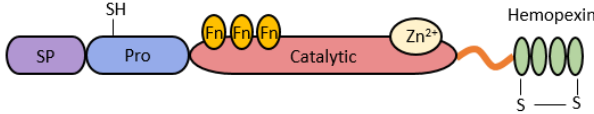
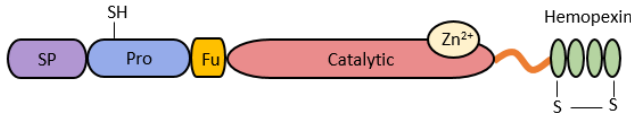
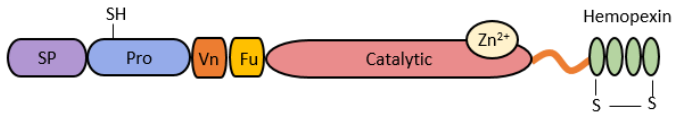
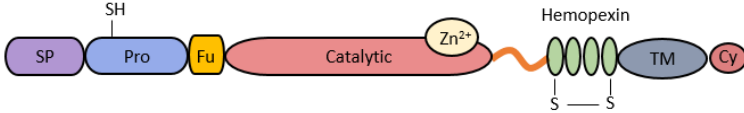
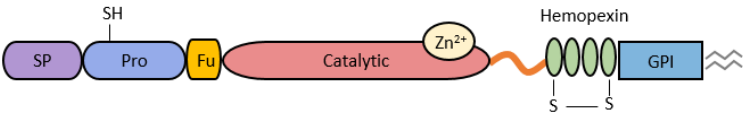
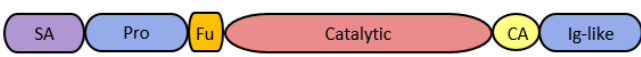
The standard structure of MMPs consists of a signal peptide (SP), a zymogenic pro-peptide (Pro), a catalytic domain, a linker and a hemopexin domain. However, while all MMPs possess a catalytic domain, individual MMPs display variation within their C-termini. For example, whereas minimal domain MMPs lack the linker and hemopexin domains completely, type II transmembrane MMPs possess a cysteine array (CA), and an immunoglobulin-like (Ig-like) domain instead. To enable anchorage to the membrane, other MT-MMPs contain either glycosylphosphatidylinositol (GPI) linkers or transmembrane (TM) and cytosolic domains. Some MMPs contain fibronectin type-II-related domains (Fn), that are required for collagen binding, whereas others contain a vitronectin-like domain (Vn). Further differences include the presence of a furin-recognition site (Fu) between pro and catalytic domains in furin-activated MMPs, which enable binding by intracellular serine proteinases (Fig 1.5) (136, 137).

## ADAM

The a disintegrin and metalloproteinases (ADAMs) are a family of 21 membrane-bound proteases, known for facilitating the shedding of cytokines and growth factors to mediate cellular signalling (140, 141). Like the MMPs, ADAMs contain a signal peptide (SP), a zymogenic pro-peptide (Pro), and a catalytic domain. They also contain a unique disintegrin domain, required for integrin binding. Additional domains include a cysteine-rich domain, an epidermal growth factor (EGF)-like domain, as well as a transmembrane and cytoplasmic domain to enable membrane binding (Fig 1.6) (140-142). While all ADAMs possess the metalloproteinase domain, 8 out of 21 ADAMs lack site residues required for catalysis, rendering them catalytically inactive (Fig 1.6) (143).

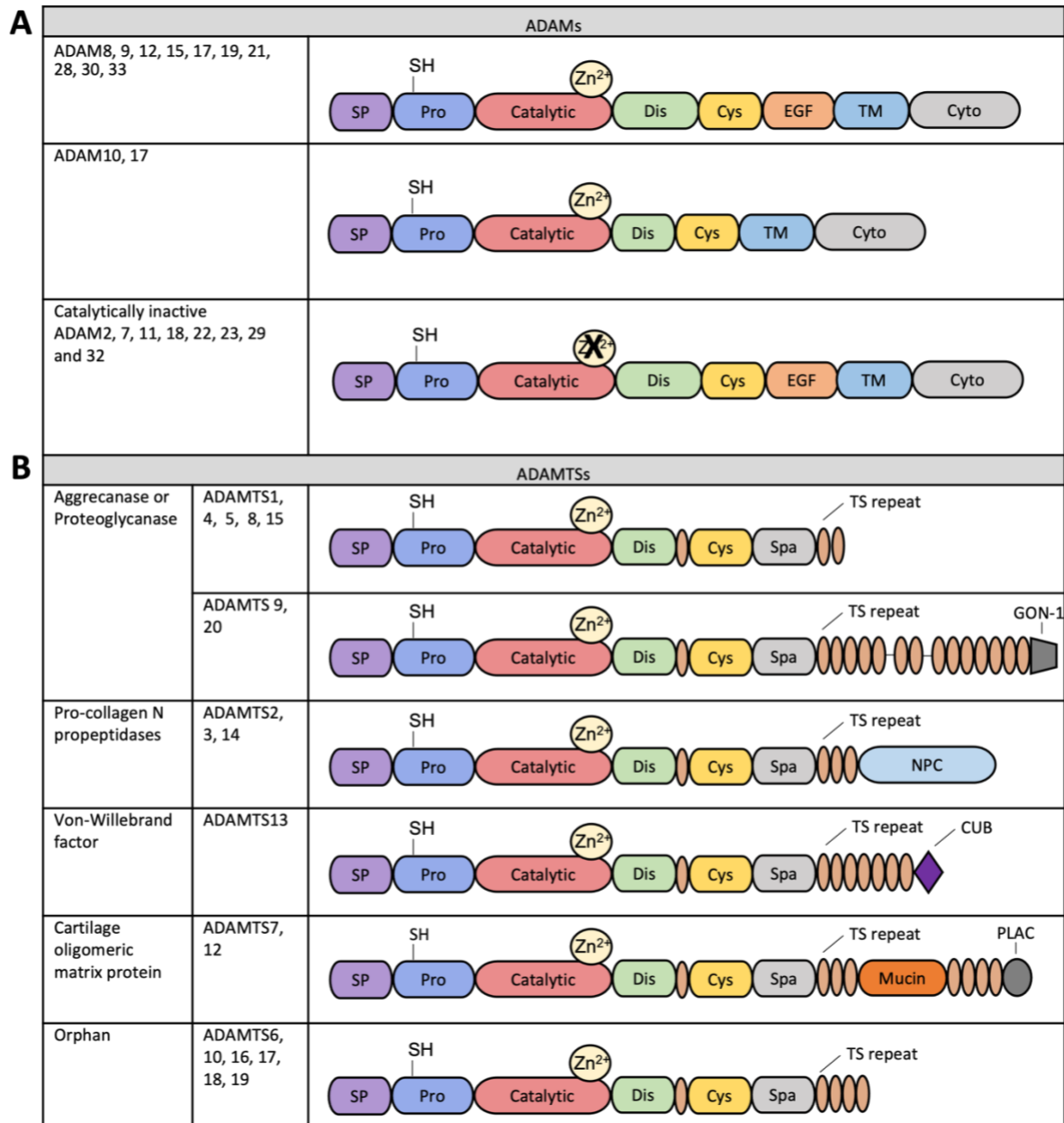
## ADAMTS

In contrast to the ADAMs, the a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS) subfamily are secreted enzymes, comprising 19 extracellular proteases. They are categorised into six groups based on their structure and functions, namely the aggrecanases or proteoglycans, the pro-collagen N-propeptidases, the cartilage oligomeric matrix protein-cleaving enzymes, the von-Willebrand Factor proteinase and the orphan enzymes. Their N-terminal structures are like those of ADAMs, although they lack transmembrane and cytoplasmic domains and can thus be distinguished structurally. They comprise a signal peptide (SP), a zymogenic pro-peptide (Pro), a catalytic domain, a disintegrin-like domain, a thrombospondin repeat (TS) and a cysteine-rich domain (Cys) followed by a spacer region (Spa). The C-terminus downstream of the thrombospondin sequence repeat displays the most variation between ADAMTS enzymes (Fig 1.6), enabling individual ADAMTS enzymes to carry out distinct functions (Fig 1.6). For example, ADAMTS2, 3 and 14 contain procollagen pro-peptidase domains (NPC), which gives rise to their roles in collagen processing (140).

Secreted MMPs		
Minimal-domain	Matrilysin (MMP7) Matrilysin 2 (MMP26)	
Simple hemopexin-domain containing	Collagenase 1 (MMP1) Collagenase 2 (MMP8) Collagenase 3 (MMP13) Stromelysin 1 (MMP3) Stromelysin 2 (MMP10) Metalloelastase (MMP12) RASI-1 (MMP19) Enamelysin (MMP20) MMP27	
Gelatin-binding	Gelatinase A (MMP2) Gelatinase B (MMP9)	
Furin-activated	Stromelysin 3 (MMP11) Epilysin (MMP28)	
Vitronectin-like insert	MMP21	
Membrane-bound MMPs (MT-MMPs)		
Transmembrane	MMP14 (MT1-MMP) MMP15 (MT2-MMP) MMP16 (MT3-MMP) MMP24 (MT5-MMP)	
GPI-anchored	MMP17 (MT4-MMP) MMP25 (MT6-MMP)	
Type II transmembrane	MMP23	

**Figure 1.5 Classification and structure of MMPs.**

Based on their structure, the MMPs can be grouped into eight groups. Minimal domain MMPs contain a signal peptide (SP), a pro-peptide (Pro) with a zinc interacting thiol (SH) group that maintains protease inactivity and a catalytic domain with a zinc binding site (Zn). In addition to these domains, simple hemopexin containing MMPs have also have a hemopexin-like domain, connected by a variable linker region, which enables interactions with substrates, cell surface molecules and tissue inhibitors of metalloproteinases. The two gelatin-binding MMPs contain inserts that resemble collagen-binding type II repeats of fibronectin (Fn) within their catalytic domain and furin-activated MMPs have a furin-like serine protease (Fu) motif between their propeptide and catalytic domain that allows intracellular activation by these proteinases. In addition to a Fu motif, MMP21 also has a vitronectin-like insert (Vn). Membrane-type MMPs can either be anchored to the membrane with a glycosylphosphatidylinositol (GPI) linker, or by transmembrane (TM) domains. MMP23 has a signal anchor (SA), a cysteine array (CA) domain and an immunoglobulin-like (Ig-like) domain. Adapted from Egeblad *et al.* (136).



**Figure 1.6 Classification and structure of ADAM and ADAMTS proteases.**

(A) The typical ADAMs structure contains a signal peptide (SP), a pro-peptide (Pro), a catalytic domain, a disintegrin-like domain (Dis) and a cysteine rich domain (Cys). Due to their membrane-bound nature, ADAMs also have an epidermal growth factor-like domain (EGF), with the exception of ADAM10 and 17, a transmembrane (TM) and a cytoplasmic tail (Cyto). (B) Based on their structure and function, the ADAMTSs can be grouped into six groups. All ADAMTSs comprise a signal peptide (SP), a pro-peptide (Pro) with a zinc interacting thiol (SH) group that, a catalytic domain with a zinc binding site (Zn), a disintegrin-like domain (Dis), a central thrombospondin repeat (TS), a cysteine rich domain (Cys), a spacer region (Spa) and a variable number of thrombospondin repeats (TS). ADAMTSs display variation in their C termini. For example, ADAMTS2, 3 and 14 contain a pro-collagen propeptidase domain (NPC), whereas ADAMTS13 contains a CUB domain, consisting of 10 conserved cysteine residues. ADAMTS7 and 12 contain a protease and lacunin (PLAC) domain, that contains 6 conserved cysteine residues. Adapted from Zhong *et al.*, Kelwick *et al.* (144, 145).

### 1.4.3. Regulation and activation of metzincins

To avoid unwanted enzymatic activity, the metzincins are meticulously controlled by multiple layers of regulation (137, 146). In addition to tightly regulated transcription, metzincins are synthesised as inactive zymogens, where a thiol (SH) in the pro-peptide domain shields the zinc ion in the active site to prevent substrate access. As such, their activation requires disruption and/or removal of the pro-domain, which occurs upon direct proteolytic cleavage or conformational changes during interactions with ECM or cell surface molecules. For most MMPs, extracellular activation occurs upon cleavage by already activated MMPs, or by serine proteases that cleave the bonds within the pro-peptide domain. There are exceptions to this, such as MMP2, which requires a multistep activation process utilising both MT-MMPs and tissue inhibitors of metalloproteinases (TIMP) 2 (146, 147). Some MMPs, such as 11, 28 and MT-MMPs can also be activated prior to their secretion by intracellular furin-like serine proteinases (146, 148).

Upon activation, the proteolytic activity of metzincins can be further controlled by the endogenous inhibitors, TIMPs and  $\alpha$ 2-macroglobulin, to negatively regulate metzincin activity (146). The TIMP family of endogenous metalloproteinase inhibitors consist of four TIMPs, accordingly named TIMP1 through to 4. While the N-terminal domain of TIMPs is responsible for their inhibiting function via binding to the metzincin active site, the C-terminal domain varies between TIMPs, giving rise to their distinct individual properties. For example, while the C-terminal domain supports binding to cell surfaces for TIMP2, it instead supports ECM binding for TIMP3 (149, 150).  $\alpha$ 2-macroglobulin, secreted mainly by hepatocytes, is another inhibitor of metzincins, present predominantly in the plasma. It exerts an inhibitory effect via its cleavable bait region, which when proteolytically cleaved, drives a conformational change to enable physical trapping of the protease (151).

#### 1.4.4. Role of metzincins in cancer

##### MMP

MMPs are often upregulated in cancer, where their expression is associated with poor prognosis. They are key mediators in creating a tumour-permissive microenvironment, due to their ability to degrade and remodel the tumour-adjacent stroma. They also possess a large repertoire of substrates in addition to structural components of the ECM, such as growth factors and cell surface proteins. Whereas some MMPs may be expressed exclusively by cancer cells, other MMPs are often secreted by cells of the tumour microenvironment (136).

An example of a role beyond ECM proteolysis is activation of TGF $\beta$  by MMP9, where docking of MMP9 to the cell surface receptor, CD44, mediates MMP9-induced TGF $\beta$  activation (136). Another example includes cleavage of E-cadherin by MMP3, which results in loss of cell-cell adhesions and subsequent nuclear translocation of  $\beta$ -catenin to promote EMT (152). MMPs have also been implicated in inducing chemoresistance through the cleavage of ligands and receptors that transduce pro-apoptotic signals. This has been shown to be mediated by MMP7, which can cleave Fas ligand from the surface of cancer cells treated with doxorubicin, resulting in the abrogation of apoptosis and reduced chemosensitivity (153). MMP7 has since shown potential to serve as a biomarker for chemotherapy resistance in non-small cell lung cancer and colorectal cancer (154, 155).

In addition to exerting effects through proteolysis, an increasing number of studies have demonstrated that MMPs also possess non-proteolytic functions (156, 157). This typically relies on localisation of MMPs to the cell surface, which in the case of MT-MMPs, is mediated through the presence of a transmembrane domain or GPI linker. For secreted MMPs lacking these structures,

localisation is instead mediated through their binding to cell surface receptors, including integrins and receptor tyrosine kinases (RTKs). MMP-induced signalling through receptors can occur upon interaction with the MMP hemopexin domain, to drive activation of downstream signalling pathways associated proliferation, migration and mediating angiogenesis (157). Examples include MMPs 2 and 9, which have been shown to enhance angiogenesis and activation of MAPK, through integrin  $\alpha V\beta 3$  and ERBB receptors, respectively (158, 159).

While MMPs are typically known for their tumour-promoting roles, several MMPs have also been implicated as tumour suppressors in the context of breast cancer (160-162). For example, downregulation of MMP8 in breast cancer cells has been shown to increase metastatic potential (161), with MMP8 loss observed in DCIS cases with an invasive component compared to pure DCIS (160). Interestingly, overexpression of MMP8 in DCIS-associated myoepithelial cells induces tumour-suppressive effects by promoting cellular adhesion, which reduces their ability to migrate. This is thought to be mediated via its ability to alter MMP9 activity and reduce MMP9-mediated cleavage of integrin  $\alpha 6\beta 4$  structures in hemidesmosomes. MMP12 has also been implicated as a suppressive mediator of tumour growth and angiogenesis (162), due to its ability to cleave plasminogen and generate angiostatin, which blocks endothelial cell proliferation and migration (163).

## ADAM

ADAMs are known for their ability to process and shed the ectodomains of cytokines, membrane-anchored growth factors and growth factor receptors, to modulate cellular functions such as cell-cell adhesion and cellular signalling (164). In cancer, ADAMs 9, 10, 12, 15 and 17 are implicated as promoters of malignant progression. Indeed, pharmacological inhibition of ADAM17 decreases breast cancer cell growth *in vitro* (165), while genetic knockdown of ADAMs 9 and 15 reduces tumour growth *in vivo* (140). Out of the ADAM-activated growth factors, ERBB receptor ligands are amongst the best studied. ADAMs 10 and 17 are able to cleave ligands such as TGF $\alpha$ , EGF, HB-EGF, and heregulins to activate MAPK, PI3K/AKT and JAK/STAT pathways (165-167), while ADAM15 mediates E-cadherin shedding to increase cell migration (168).

While protease activity is the best-defined function of ADAMs, their second defining feature is their ability to bind integrin receptors. Under certain conditions, ADAMs are able to promote cell migration and invasion through integrin activation or integrin-mediated localisation of ADAMs to the cell surface (169). This has been demonstrated in the context of ADAM9, where its binding to  $\alpha 2\beta 1$ ,  $\alpha 6\beta 1$  or  $\alpha 6\beta 4$  induces an invasive phenotype (169, 170).

## ADAMTS

Despite improvements in our understanding of ADAMs, the secreted ADAMTS subfamily have yet to be studied in the same detail. However, many of the ADAMTS family members are proposed to be tumour-suppressive, reflected by their epigenetic silencing or mutational inactivation in cancer (171). For example changes to methylation of the *ADAMTS1* gene have been shown to support gastric cancer progression (171), while its knockdown stimulates migration and invasion of breast cancer cells (172). Expression of ADAMTS proteases can also deactivate key proliferation and survival pathways, such as



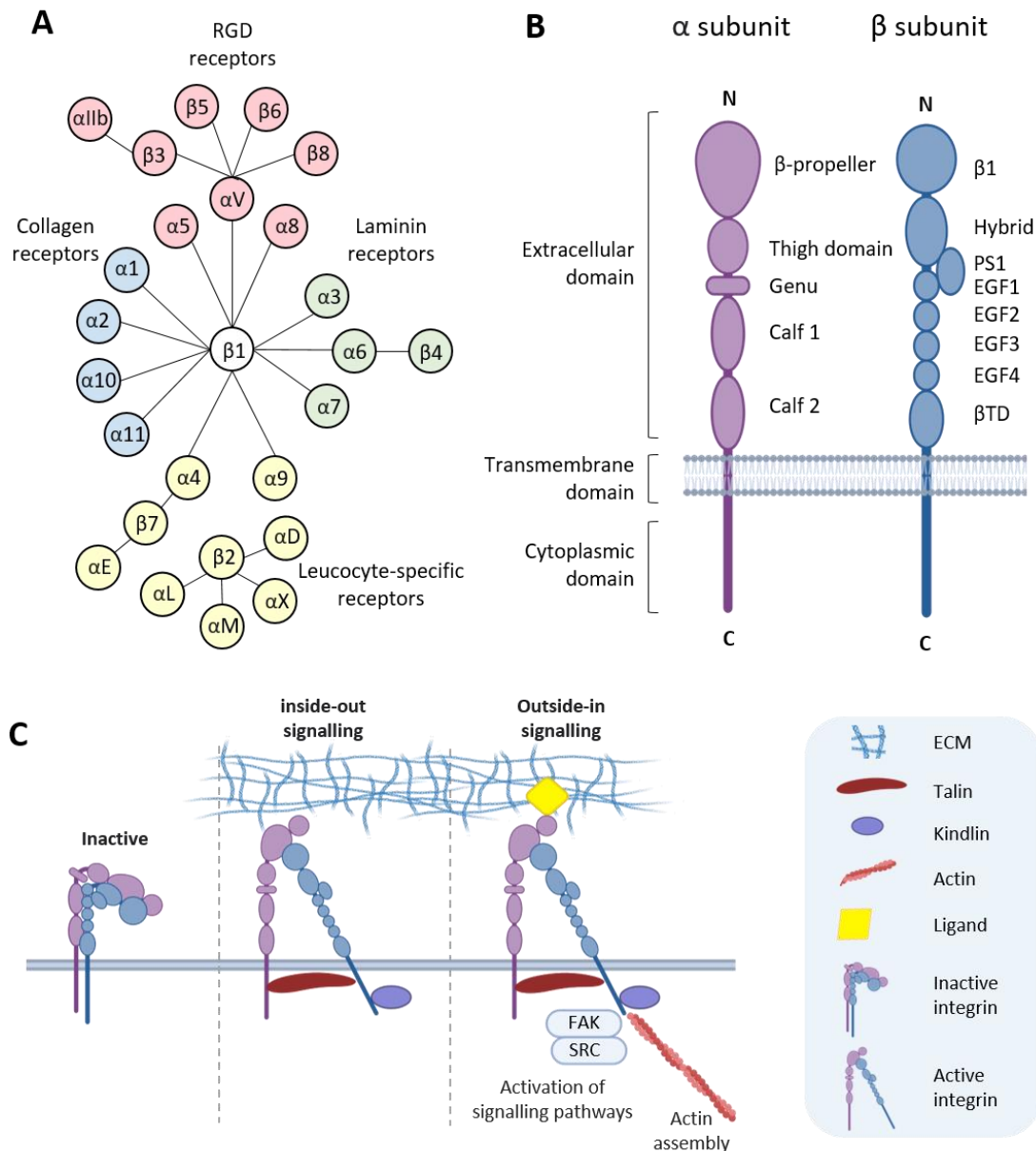
inhibition of AKT signalling by ADAMTS9 (173), as well as ERK signalling by ADAMTS8 (174), 12 (175) and 15 (176). ADAMTS members have also been implicated as anti-angiogenic modulators (135, 177, 178), where ADAMTS1 can cleave thrombospondins 1 and 2 to generate anti-angiogenic fragments (178), as well as degrade angiogenic factors VEGF and FGF (179). In breast cancer, ADAMTS 1, 3, 4, 8, 9, 10 and 18 are reportedly downregulated, whereas only ADAMTS4, 5, 14 and 20 are upregulated (180). However, their relevance in either suppressing or promoting breast cancer progression remains to be elucidated.

## 1.5. Integrins

### 1.5.1. Overview

Integrins are bidirectional signalling receptors that facilitate signals between intracellular and extracellular environments to mediate cell-cell and cell-matrix interactions (181-184). Importantly, they function as mechanotransducers, using mechanical forces to stimulate biochemical signals. Integrins are heterodimeric transmembrane proteins that consist of an  $\alpha$  and  $\beta$  subunit. In humans, there are 18  $\alpha$  and 8  $\beta$  subunits, which can associate to form 24 different variations of integrin receptors, each with different binding specificity and distribution. Based on their ligands, integrins can be grouped into collagen, RGD (Arg-Gly-Asp) and laminin receptors, or according to their expression on leukocytes, where they are referred to as leukocyte-specific receptors (Fig 1.7A).

Integrins are comprised of three domains: extracellular, transmembrane, and cytoplasmic (Fig 1.7B). The extracellular domain is largely responsible for ligand binding and is comprised of several subdomains. For the  $\alpha$  subunit, the N-terminal head consists of a  $\beta$  propeller, a thigh domain and two calf domains. Instead, the  $\beta$  subunit is made up of a hybrid domain that connects the I-like domain ( $\beta 1$ ) to a plexin-semaphorin-integrin domain, four epidermal growth factor domains, and a membrane proximal  $\beta$  tail domain ( $\beta$ TD). The transmembrane domains of both subunits span the membrane once, and form glycine-glycine interactions to enable their close association. The cytoplasmic domain recruits scaffold proteins, which associate with the adaptor and actin cytoskeleton proteins for intracellular signal transduction (Fig 1.7B) (181-184).



**Figure 1.7 Overview of integrin signalling.**

**(A)** 18  $\alpha$  and 8  $\beta$  subunits can associate to form 24 combinations of integrin receptors. Based on their ligands, integrins are classified into collagen (blue), RGD (red), laminin (green) or leucocyte-specific receptors (yellow).

**(B)** Integrins contain three domains; extracellular, transmembrane and cytoplasmic. The  $\alpha$  subunit consists of a  $\beta$ -propeller head, a thigh domain, and two calf domains. The  $\beta$  subunit is made up of a I-like domain ( $\beta 1$ ), a hybrid domain, a plexin-semaphorin integrin domain, four epidermal growth factor domains and a proximal  $\beta$  tail domain ( $\beta$ TM). Both subunits also contain transmembrane and cytoplasmic domains.

**(C)** When inactive, integrins exhibit a bent conformation. In their activated state, integrins adopt an extended conformation and can signal 'inside-out' and 'outside-in.' During inside-out signalling, talin and kindlin bind to the cytoplasmic domain of the integrins, disrupting association of the transmembrane domains. During outside-in signalling, ligand binding promotes recruitment of adhesome to activate downstream signalling pathways and assembly of the cytoskeleton. Partly created with BioRender.

### 1.5.2. Integrin signalling in cancer

Integrins communicate in both directions, relaying signals between the extracellular (termed outside-in signalling) and intracellular environments (termed inside-out signalling) (185). In their inactive form, integrins are unbound to a ligand and exhibit a bent conformation through tight association between the transmembrane and cytoplasmic domains of each subunit. During inside-out signalling, binding of proteins, talin and kindlin, to the cytoplasmic domain of the integrin, drives separation of the transmembrane and cytoplasmic domains. Disruption of this association causes integrins to adopt an active conformation via extension of their extracellular domains and exposure of binding pockets toward the extracellular space (Fig 1.7C) (182, 186).

During outside-in signalling, binding of integrins to their respective ligands promotes recruitment of the adhesome: a complex association of signalling, scaffolding and cytoskeletal proteins that regulate aspects of cell fate. Engagement of the adhesome drives activation of FAK and SRC family kinases, which drives cooperation with RTKs to promote signalling through cascades including RAS-ERK and PI3K/AKT (185, 187). The GTPases; Cdc42, Rac and Rho, are also activated to coordinate actin filament polymerisation, to regulate formation of filopodia, lamellipodia and focal adhesions (Fig 1.7C) (182, 186). Integrins also bind directly to growth factors, with the best example being integrin  $\alpha\text{v}\beta\text{6}$ , which plays a crucial role in TGF $\beta$  activation (182).

Integrin signalling has been implicated in nearly every step of cancer progression, where its altered expression can facilitate initiation and proliferation, migration and invasion, metastatic colonisation, and therapeutic resistance (184, 186, 188, 189). In breast malignancies, histological grade correlates with overexpression of FAK (190), while SRC activation promotes cancer stemness and increases

metastatic potential (191). With most solid tumours derived from epithelial origin, integrins expressed by epithelial cells, such as  $\alpha\text{v}\beta\text{5}$ ,  $\alpha\text{2}\beta\text{1}$ ,  $\alpha\text{3}\beta\text{1}$ ,  $\alpha\text{6}\beta\text{1}$  and  $\alpha\text{6}\beta\text{4}$ , are often altered (183). These integrins are typically involved in adhesion to the BM but may also contribute to the migration, invasion, and survival of tumour cells. Integrins usually detected at low or undetectable levels in adult epithelium, most notably  $\alpha\text{v}\beta\text{3}$ ,  $\alpha\text{v}\beta\text{6}$  and  $\alpha\text{5}\beta\text{1}$  can also be upregulated to support cancer progression (183). Interestingly, tumour-suppressive integrins, such as integrin  $\alpha\text{2}\beta\text{1}$  can also be lost during malignant progression, as observed in breast cancer (192).

Activation of pro-tumorigenic integrins can drive upregulation of MMPs to facilitate initial proteolysis of the BM (182, 193). For example, integrin  $\alpha\text{v}\beta\text{6}$  promotes expression of MMPs 2, 3 and 9 across several cancers (194-196), while integrin  $\alpha\text{v}\beta\text{3}$  upregulates MMP2 in invasive breast cancer cells upon RGD binding (197). In addition to increasing MMP expression, integrins are also able to localise secreted MMPs to the cell surface to enhance pericellular proteolysis. This has been demonstrated for several MMPs, including binding of MMP2 to integrin  $\alpha\text{v}\beta\text{3}$  (198), and MMP9 to integrins  $\alpha\text{v}\beta\text{6}$  or  $\alpha\text{3}\beta\text{1}$  (199, 200). In breast cancer, activated integrin  $\alpha\text{v}\beta\text{3}$  can cooperate with MMP9, through its hemopexin domain to promote ECM degradation and cell migration (201).

While these studies focus on altered integrin expression within cancer cells, such changes also occur in the tumour microenvironment. Indeed, CAFs can utilise integrin-mediated contractile forces to remodel and stiffen the surrounding ECM (107, 128, 184, 202). For example, integrin-dependent activation of TGF $\beta$  can induce CAFs to upregulate lysyl oxidase enzymes (184, 203, 204) and matrix proteins, such as collagen and fibronectin (184).

### 1.5.3. Integrin $\alpha\beta 6$

Integrin  $\alpha\beta 6$  is an epithelial integrin expressed at minimal levels in healthy adult tissue. Its expression is upregulated during embryogenic development, wound healing and in cancer, where it can regulate cellular functions such as proliferation, migration, and matrix remodelling (205). Upregulation of integrin  $\alpha\beta 6$  is observed in up to one third of solid malignancies, where its expression is associated with increased invasiveness and poor prognosis (72, 206-209). Its ligands consist of proteins containing the RGD motif, such as fibronectin, TGF $\beta$ , and latency associated peptide (LAP) (205).

Expression of myoepithelial integrin  $\alpha\beta 6$  has been reported as a nascent way to stratify high-risk DCIS patients (72, 206). Immunohistochemical tissue analysis of archival samples (n = 532), showed that myoepithelial expression of integrin  $\alpha\beta 6$  was significantly higher in DCIS with associated invasion compared to pure DCIS. Subsequent analysis of myoepithelial integrin  $\alpha\beta 6$  as a predictor of recurrence was confirmed in a DCIS cohort from the UK/ANZ study (210), where integrin  $\alpha\beta 6$ -positive cases displayed a median recurrence time of 2.3 years compared to integrin  $\alpha\beta 6$ -negative cases, which showed a median of 11.4 years (72, 206). Interestingly, conditioned media from myoepithelial cells overexpressing integrin  $\beta 6$ , was shown to mediate invasion of breast cancer cell lines through integrin  $\beta 6$ /TGF $\beta$ -dependent MMP9 secretion (72). Initial upregulation of integrin  $\alpha\beta 6$  may be due to mechanical alterations exerted through myoepithelial stretching during DCIS progression (127), where increased DCIS duct size triggers myoepithelial cells to alter integrin-dependent expression of BM degrading MMPs (72, 127, 160).

## 1.6. TGF $\beta$ signalling

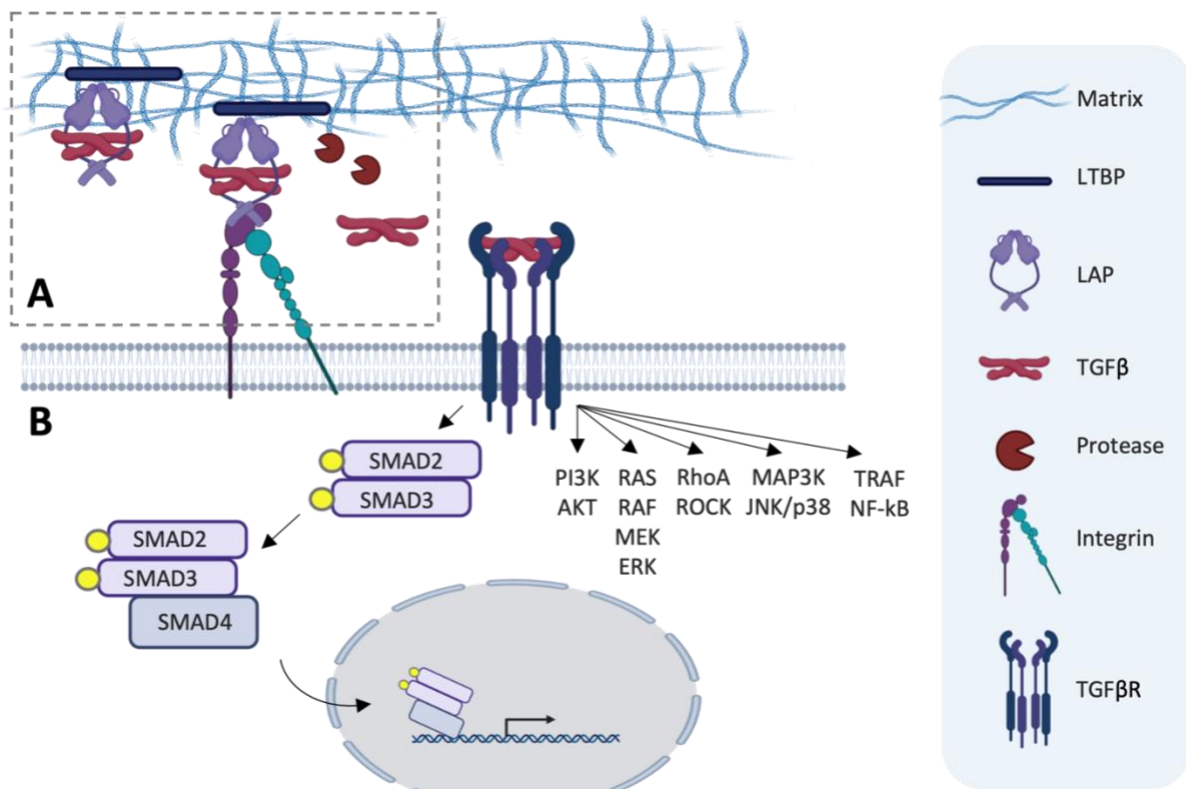
### 1.6.1. TGF $\beta$ signalling overview

The TGF $\beta$  family are multifunctional cytokines, comprising three isoforms – TGF $\beta$ 1, TGF $\beta$ 2 and TGF $\beta$ 3, which act as important modulators of various cellular processes. The three TGF $\beta$  isoforms are synthesised as homodimeric pro-proteins, interacting with LAP prior to secretion. LAP shields the receptor-interacting epitopes, maintaining TGF $\beta$  in its latent form, and can together be referred to as the small latent complex (SLC). TGF $\beta$ -bound LAP can also interact with latent TGF $\beta$  binding protein (LTBP), to form the large latent complex (LLC). Following secretion, LTBP of the LLC facilitates the localisation of TGF $\beta$  to the ECM, through interactions with various proteins such as fibronectin and fibrillin (211, 212). TGF $\beta$  becomes activated through its dissociation from LAP, via integrin-dependent or -independent mechanisms. In integrin-dependent contexts, interaction with integrins  $\alpha\beta$ 6 or  $\alpha\beta$ 8, are required to induce a tension-driven conformational change, via simultaneous binding to LAP and to the cytoskeleton, via the  $\beta$  subunit. For integrin-independent activation of TGF $\beta$ , mechanisms include cleavage of LAP by MMPs 2, 9, 13 and 14 (213), and TSP-1 binding, which disrupts interactions between LAP and TGF $\beta$  (Fig 1.8A) (211, 213).

Upon TGF $\beta$  activation, canonical signalling is initiated via binding to type II TGF $\beta$  receptor (TGF $\beta$ RII), a dimeric transmembrane protein with a constitutively active serine/threonine kinase domain. TGF $\beta$ -bound TGF $\beta$ RII can then recruit type I TGF $\beta$ R receptor (TGF $\beta$ RI), a dimeric transmembrane protein containing an inactive kinase domain (211, 214, 215). The formation of this heterotetrameric receptor complex initiates activation of TGF $\beta$ RI via phosphorylation of its kinase domain, disrupting its interaction with FKBP12, an endogenous TGF $\beta$  signalling inhibitor. The functional TGF $\beta$ R complex then recruits and phosphorylates receptor regulated proteins, SMAD2 and SMAD3, which form subsequent

heteromeric complexes with SMAD4 for nuclear localisation. In the nucleus, they can associate with other DNA-binding transcription factors to regulate transcription of target genes containing SMAD-binding elements. TGF $\beta$  is also able to regulate pathways independently of SMADs, through activation of p38, JNK, RAS-ERK, NF- $\kappa$ B and Rho GTPase cascades (Fig 1.8B) (211, 214, 215).





**Figure 1.8 Overview of TGFβ activation and signalling.**

**(A)** TGFβ is secreted in a latent form composed of a TGFβ dimer associated with latency association peptide (LAP) to form the small latent complex (SLC). The SLC can be secreted as such or can form a covalent bond with the latent TGFβ binding protein (LTBP), which mediates its association with the extracellular matrix (ECM). For TGFβ to become activated, it must be liberated from LAP. This can be mediated through integrin-dependent or -independent mechanisms. Integrin independent mechanisms include degradation of LAP by proteases. **(B)** Binding of TGFβ to TGFβRII drives activation of TGFβRI. Formation of this heterotetrameric receptor complex leads to phosphorylation of SMAD2 and SMAD3. Phosphorylated SMAD2 and SMAD3 form a complex with co-Smad, SMAD4, which leads to nuclear translocation and transcription of target genes which possess smad-binding elements. Non-canonical TGFβ signalling can also occur, independently of SMADs, through JNK, Ras-ERK, NF-κB, PI3K-Akt and Rho GTPase pathways. Created with BioRender.

### 1.6.2. Relevance of TGF $\beta$ signalling in cancer

Depending on the context, the TGF $\beta$  signalling pathway possesses a dual function, and can act as both a tumour-suppressive and tumour-promoting pathway (211). For example, during the early stages of cancer development, TGF $\beta$  can suppress proliferation and promote apoptosis, while in the later stages, it can adopt a tumour-promoting function to regulate proliferation, invasion, and metastasis (211).

The role of TGF $\beta$  as a tumour suppressor can be demonstrated by disruption to its signalling pathway in several cancers. For example, inactivating mutations and loss of heterozygosity in *TGFBR1* and *SMAD4* are commonly observed in colon cancer (216), coupled with decreased SMAD4 expression reported in others such as pancreatic, and head and neck cancer (211). In some contexts, TGF $\beta$  stimulation can inhibit cell proliferation through repressed MYC expression (211, 217), or via upregulation of cyclin-dependent kinase inhibitors, p21 and p15 (211, 218). The suppressive effects of TGF $\beta$  during breast cancer initiation can be observed following TGF $\beta$  inhibition, which increases proliferation of the cancer stem cell compartment. Supporting this, stimulation with TGF $\beta$  reduces mammosphere formation, and results in the loss of stem cell-like properties (211, 219).

Interestingly, while inactivating mutations in the TGF $\beta$  signalling cascade may evade tumour-suppressive effects, these mutations are absent from most cancers that retain functional signalling components. Indicative of its pleiotropic role, increased TGF $\beta$  signalling has been shown to correlate with cancer progression (220). In breast cancer, elevated levels of plasma TGF $\beta$  correlate with poor outcome (220, 221), with higher levels of TGF $\beta$  present in lymph node metastases compared to the primary tumour (220, 222). At the molecular level, TGF $\beta$  can support breast cancer progression

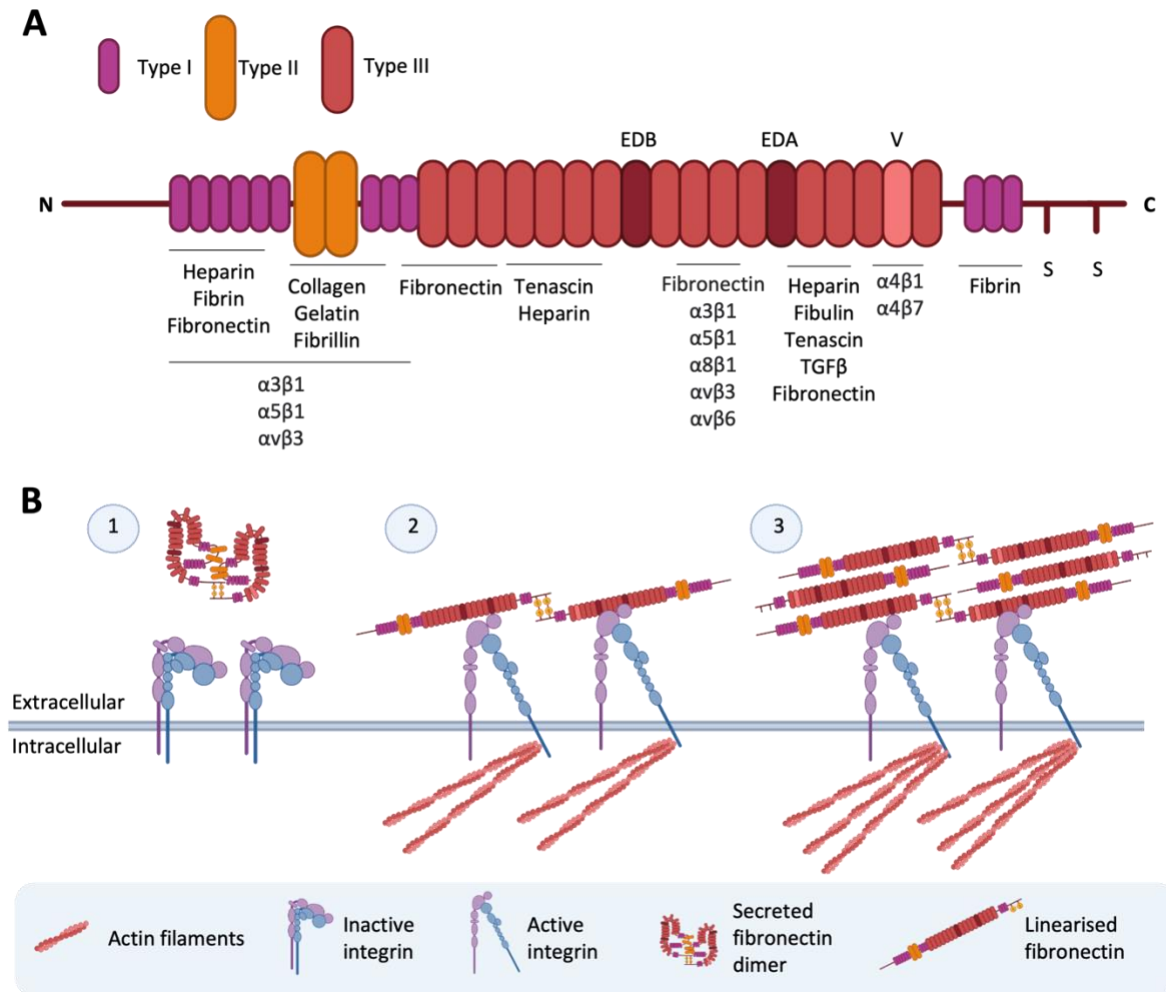
throughout the metastatic cascade, from facilitating local invasion to supporting metastatic growth. For example, TGF $\beta$  induces EMT in mammary epithelial cells, by which they lose their epithelial characteristics and adopt mesenchymal properties to enable local invasion. This process disrupts the formation of epithelial cell-cell junctions and reorganises the cytoskeleton to promote cellular motility (211, 223, 224). TGF $\beta$  has also been implicated in supporting the pro-angiogenic niche, where its high levels are associated with increased microvessel density. High levels of TGF $\beta$  induce the expression of key angiogenic factors such as VEGF and connective tissue growth factor (CTGF) in both cancer cells and surrounding fibroblasts (211, 225). TGF $\beta$  in the surrounding tumour microenvironment can also drive the transition of fibroblasts to CAFs, causing them to remodel the matrix through increased deposition, contraction, and protease secretion (107). CAFs can then upregulate and secrete additional TGF $\beta$ , creating a positive feedback loop. As an immunosuppressive cytokine, high levels of TGF $\beta$  can also dampen immune-mediated tumour cell killing (226). For example, TGF $\beta$  signalling can drive expression of FOXP3 through SMAD-regulated pathways, to promote Treg cell expansion (227). Differentiation and proliferation of CD4 $^{+}$  and CD8 $^{+}$  T cells also becomes blocked, due to TGF $\beta$ -mediated repression of IL-2, IFN $\gamma$  and granzyme B (227-229).

## 1.7. Fibronectin

### 1.7.1. Overview and structure

The ECM is a multifaceted network of proteins essential for regulating cellular behaviour through exerting biochemical and biomechanical cues. Profound perturbations in the ECM architecture are observed in cancer. Thus, there is a growing interest in understanding how the matrix may give rise to dysregulated activation of signalling pathways associated with tumour progression (120, 230, 231).

Fibronectin is a glycoprotein present in the ECM and has diverse roles in inducing cellular behaviours such as proliferation, differentiation, adhesion and migration. It typically exists as a dimer, with two nearly identical 250 kDa monomers, covalently linked by disulfide bonds at the C-terminal end (Fig 1.9A). Each monomer is made up of three types of repeating units including, 12 type I, 2 type II, and 15 type III repeats. Through alternative splicing at three regions of the pre-mRNA, 20 different isoforms of fibronectin can be generated. As such, the diversity in the functional domains of fibronectin allows for multiple binding partners (232). The first and second splicing regions consist of inclusion or exclusion of type III repeats between positions 11 and 12 (extradomain A;EDA) and/or position 7 and 8 extradomain B;EDB), while the third region consists of one type III repeat known as the variable (V) region (Fig 1.9A).



**Figure 1.9 Structure of fibronectin and overview of fibronectin fibrillogenesis.**

**(A)** The fibronectin monomer is composed of a series of type I, type II and type III sequences, which give rise to a variety of binding sites for other matrix proteins, growth factors and cell receptors. Fibronectin can undergo splicing in three regions, where type III repeats are included or excluded at extradomain A (EDA), extradomain B (EDB), and/or variable (V) region. **(B) 1)** Fibronectin is secreted as a dimer, with two fibronectin monomers linked by disulfide bonds. **2)** Fibronectin matrix assembly is initiated upon binding to integrins, such as integrin  $\alpha 5\beta 1$ , which activates intracellular signalling pathways to induce polymerisation of the actin cytoskeleton. Increased tension upon cell-mediated contraction induces fibronectin to adopt a linearised confirmation. **3)** Exposure of fibronectin self-associating domains enables the formation of fibronectin fibrils. Simplified from Spada *et al.* (231). Created with BioRender.

### 1.7.2. Function of fibronectin and relevance in cancer

Fibronectin fibrillogenesis is a dynamic, cell type-dependent process by which a fibrillar matrix is assembled (Fig 1.9B). In summary, binding of compact fibronectin dimers to the classical fibronectin receptor, integrin  $\alpha 5\beta 1$ , triggers its activation and the recruitment of cytoplasmic partners to drive Rho-mediated fibre formation. Increased cellular contractility applies strain on the fibronectin molecule, inducing its conformation to a stretched state (Fig 1.9B). Creation of this polymerised network enables the incorporation of other matrix proteins, where it can control the stability of collagen, gelatin, tenascin and fibrillin. The scaffold also acts as a growth factor reserve, regulating the bioavailability of growth factors such as TGF $\beta$ , PDGF, FGF and VEGF (233, 234).

While stromal fibronectin is absent from the healthy adult breast, its deposition is upregulated in cancer and associated with an invasive and metastatic phenotype (235, 236). Remodelling of the matrix is a key hallmark of CAFs in facilitating invasion (107, 237). Indeed, integrin  $\alpha 5\beta 1$ -mediated fibronectin alignment has been shown to promote directional cancer cell migration (202), while integrin  $\alpha 5\beta 1$  expression is associated with tumour metastasis (238). Supporting this, culture of tumour cells on fibroblast-derived matrices can also promote invasion via integrin  $\alpha v\beta 6$  binding and subsequent TGF $\beta$  activation (239). In addition to its roles as a physical scaffold, fibronectin can also mediate intracellular signalling through integrin-mediated outside-in signalling. For example, stimulation with soluble fibronectin can activate signalling through the FAK/SRC/PI3K axis to promote secretion of MMPs and TGF $\beta$  (240, 241).

## 1.8. Hypothesis and project aims

While mammographic screening successfully detects breast cancer at an earlier, more treatable stage, issues in the overtreatment of women with DCIS, whose disease may have remained indolent, have emerged. With a lack of tools to differentiate between low and high-risk patients, an improved understanding of the mechanisms that drive disease progression are required to identify prognostic markers. However, with a lack of distinct changes in cancer cells across DCIS and IDC, it is suggested that the most critical changes that underpin progression, lie in the surrounding microenvironment.

With the loss of the myoepithelium presenting as a key step in the transition of DCIS to IDC, there is extensive interest in understanding its role in modulating disease progression. As DCIS-associated myoepithelial cells become altered compared to their naturally tumour suppressive counterparts, we hypothesise that they may become tumour-promoting in DCIS to drive the invasive progression of neoplastic luminal cells. However, with a lack of myoepithelial-containing DCIS models, dissecting the myoepithelial contributions presents as a challenge. Thus, an aim of this study is to develop a physiologically relevant model of DCIS progression, with the incorporation of myoepithelial cell and matrix compartments. By using a model whereby expression of integrin  $\beta 6$  can be induced, this will enable further investigation into myoepithelial-associated tumour-promoting mechanisms, with a focus on matrix remodelling processes. Determining critical changes in the protease expression profile, together with identifying key proteolytic products, will allow for biological mechanisms to be explored. To confirm the translational relevance of mechanistic insights, key findings will also be validated in clinical cohorts with the aim of identifying candidate markers that could aid patient stratification in DCIS.

Thus, the objectives of this study are to:

(1) Develop a genetically tractable physiomimetic *in vitro* model of DCIS progression. By establishing a model whereby luminal and myoepithelial cell compartments are combined in a physiologically relevant matrix, myoepithelial-matrix interactions can be dissected and validated in a robust and high-throughput manner.

(2) Identify protease-substrate pairs that may mediate myoepithelial-led invasion. This will allow for key myoepithelial mechanisms that promote matrix remodelling and the rewiring of subsequent signalling pathways, through proteolytic processing, to be elucidated.

(3) Assess the translatable relevance of key discoveries by validating findings in relevant *in vivo* mouse models, such as the intraductal xenograft model, and in tissue sections from a clinical DCIS cohort.

Taken together, these objectives will allow for the role of the myoepithelial cell, with a focus on tumour-promoting mechanisms, be determined in physiologically relevant models of DCIS progression. A greater understanding of these processes will pave the way towards identifying robust markers for risk stratification in DCIS.



## **Chapter 2. Methods**

## **2.1. Cell culture**

### **2.1.1. Cell culture conditions**

#### **2.1.1.1. Cell lines**

The 1089 myoepithelial cell line was provided by Dr. M. Allen (72). 1089 cells were originally isolated from reduction mammoplasty, using integrin  $\beta$ 4 magnetic beads as described, and immortalised via expression of both the catalytic subunit of human telomerase (hTERT) and simian virus 40 large-tumour antigen (SV40 Tag) (111, 242). The non-tumorigenic HB2 luminal cell line was provided by Prof. V. Spiers (243, 244). HB2 cells were originally isolated from human lactation milk as described (243, 244), and immortalised via expression of SV40 Tag. Human embryonic kidney (HEK293T) cells were used to generate lentiviral particles.

#### **2.1.1.2. Culture of cell lines**

HB2 luminal cells were cultured in Dulbecco's modified eagle medium (DMEM) (Gibco, 11550356), supplemented with 10% foetal bovine serum (FBS) (Gibco, 11550356), 0.5  $\mu$ g/ml hydrocortisone (Sigma-Aldrich, H0888) and 5  $\mu$ g/ml insulin (Sigma-Aldrich, I9278). 1089 myoepithelial cells were cultured in Ham's Nutrient mixture F-12 (F12) medium (Sigma, N6658), supplemented with 10% FBS, 0.5  $\mu$ g/ml hydrocortisone, 10 ng/ml epidermal growth factor (EGF) (Sigma-Aldrich, E9644) and 5  $\mu$ g/ml insulin. HEK293T cells were cultured in Dulbecco's modified eagle/Nutrient mixture F-12 (DMEM/F12) medium (Sigma-Aldrich, D8437) supplemented with 10% FBS. All cells were incubated in a 5% CO<sub>2</sub> humidified atmosphere at 37 °C.

To passage cells, medium was removed from flasks without disturbing the cell layer. To detach cells, medium was replaced with 2.5%/1% trypsin/ethylenediaminetetraacetic acid (EDTA) (Sigma-Aldrich, 59418C) and cells incubated at 37°C for approximately 5 min. Once detached, full medium was added to quench the effects of trypsin and the cell suspension was centrifuged at 240 x g for 3 min to pellet cells. Cells were resuspended in fresh full medium and counted using a haemocytometer prior to seeding. Unless stated otherwise, seeding densities for 1089 cells were as follows:  $5 \times 10^4$  cells/well for a 6-well plate,  $2.5 \times 10^4$  cells/well for a 12-well plate,  $1.25 \times 10^4$  cells/well for a 24-well plate for 48 h experimental time points. Densities were adjusted accordingly depending on desired experimental culture times.

For cell storage, cell pellets were resuspended in 1 ml 10% dimethyl sulfoxide (DMSO) (Fisher Scientific, D/4120/PB08) in FBS aliquots in cryovials. These were frozen slowly to -80 °C and transferred to liquid nitrogen for long-term storage. Upon cell recovery, cell-containing cryovials were thawed in a 37 °C water bath and cells were transferred into falcons to be centrifuged at 240 x g for 3 min. Supernatant containing DMSO was removed, and cell pellets were resuspended in respective fresh full medium prior to culture.

#### **2.1.1.3. Primary cells**

All primary cells were isolated by Dr. J. Gomm and I. Goulding at the Breast Cancer Now Cell Bank. In summary, fresh human breast tissue was digested in 1 mg/ml collagenase 1A (Sigma, C2674) and 0.5 mg/ml hyaluronidase (Sigma, H3506) prior to further digestion to single-cell suspensions with a 0.05%/0.02% trypsin/EDTA solution (Hyclone, SV30031) containing 0.4 mg/ml DNase (Sigma, 10104159001) for 10 min at 37 °C. Cells were isolated by fluorescence-activated cell sorting (FACS) as

described, using anti-EpCAM-FITC (BD Biosciences, 347197) or anti-CD10-APC (BD Biosciences, 332777) antibodies, for luminal and myoepithelial populations, respectively (245-247).

#### **2.1.1.4. Culture of primary cells**

Both primary luminal and myoepithelial cells were cultured on 0.08 mg/ml collagen I (Corning Life Sciences, 354236) coated 6-well plates, seeded at approximately  $1 \times 10^5$  cells/well. Myoepithelial cells were cultured in Human mammary epithelial cell (HuMEC) medium (Gibco, 12752010), supplemented with 50 µg/ml bovine pituitary extract (BPE) (Invitrogen, 13028014), 0.5 µg/ml hydrocortisone, 10 ng/ml EGF and 5 µg/ml insulin. Luminal cells were cultured in DMEM/F12 medium supplemented with 10% FBS, 0.5 µg/ml hydrocortisone, 10 µg/ml apotransferrin (Sigma-Aldrich, T1147), 10 ng/ml EGF and 5 µg/ml insulin. All cells were incubated in a 5% CO<sub>2</sub> humidified atmosphere at 37 °C.

When passaging cells, medium was removed from flasks without disturbing the cell layer. To detach cells, 0.25%/0.1% trypsin/EDTA, diluted in phosphate buffered saline (PBS), was added and incubated at 37°C. Once detached, full DMEM/F12 medium was added to quench the effects of trypsin and the cell suspension was centrifuged at 240 x g for 3 min to pellet cells. Cells were resuspended in respective fresh full medium and counted using a haemocytometer prior to seeding.

### 2.1.2. 3D cell culture models

#### 2.1.2.1. Primary duct model

Ductal cultures with primary cells were performed as previously described (245). Luminal and myoepithelial cells were seeded in collagen-I coated 6-well plates at a density of  $1 \times 10^5$  cells/well 24 h prior to lentiviral transduction with HER2 pINDUCER or ITGB6 pINDUCER constructs. 48 h post transduction, primary luminal and myoepithelial cells were combined, in a 1:1 ratio, to give a final concentration of  $4.8 \times 10^5$  cells/ml. A 4 mg/ml collagen-I gel mix was prepared, consisting of 4 mg/ml collagen type I (Corning Life Sciences, 354249) and 25 mM HEPES (pH 7.0), prepared in complete DMEM/F12 medium, adjusted to neutral pH with 1N NaOH (Table 2.1). Equal amounts of 4 mg/ml collagen-I gel mix was added to the cell suspension, to give a final concentration of 2 mg/ml (Table 2.1). Once gels had polymerised and set for 30 min at 37 °C, they were overlaid with complete luminal culture medium. Ductal structures were formed by 14 days, after which cultures were treated with 1 µg/ml doxycycline to induce transgene expression alongside indicated treatments (Table 2.5).

**Table 2.1 Collagen-I gel recipe for primary duct model**

	Volume ( $\mu$ l)
Combined 1° Lumi and 1° Myo suspension ( $4.8 \times 10^5$ cells/ml)	250
Full DMEM/F12 medium	132.5
1 M HEPES (pH 7.5)	5
Rat tail collagen type I (10 mg/ml)	107.5
1N NaOH	4

**2.1.2.2. Cell line spheroid model**

HB2 and 1089 cells were combined in a 2:1 ratio, at a concentration of  $3.3 \times 10^4$  cells/ml in 0.24% methylcellulose (Sigma, M0512) prepared in DMEM/F12 culture medium (Table 2.2). Using a multichannel pipette, 20  $\mu$ l hanging droplets (containing approx. 500 cells/droplet) of methylcellulose cell suspension was pipetted onto the underside of a 100 mm dish lid and incubated overnight at 37 °C to form spheres. 10 ml of H<sub>2</sub>O was added to the lower dish to create a humid chamber and prevent droplets from drying out. Spheres were collected using wide-bore tips, spun at 300 x g with brake off for 4 min prior to removing the methylcellulose-containing supernatant. Spheres were washed with 1 ml full 1089 medium to remove traces of methylcellulose and spun again. Medium was removed and spheres were resuspended in collagen gels, consisting of 4 mg/ml collagen type I and 25 mM HEPES (pH 7.0), prepared in complete F12 medium, adjusted to neutral pH with 1N NaOH (Table 2.3). 50  $\mu$ l sphere-gel mixture, containing approximately 6 spheres, was added per well of a 96-well plate prior to gel polymerisation at 37 °C for 30 min with plate inversion every 30 sec for the initial 5 min. Gels were overlaid with indicated drug treatments and incubated for 4 days (Table 2.5).

**Table 2.2 Corresponding methylcellulose and cell suspension volumes**

Number of spheroids (approx.)	1.2% Methylcellulose ( $\mu$ l)	Cell suspension (concentration of $3.3 \times 10^4$ / ml) ( $\mu$ l)		Total Volume ( $\mu$ l)
		HB2	1089	
50	200	533	266	1000
100	400	1066	533	2000
150	600	1600	800	3000

**Table 2.3 Collagen-I gel recipe for HB2/1089 cell line model**

	Volume ( $\mu$ l)
Complete F12 medium	267
1 M HEPES (pH 7.5)	10
Rat tail collagen type I (10 mg/ml)	215
1N NaOH	8



### 2.1.2.3. Primary myoepithelial spheroid model

Luminal and myoepithelial cells were seeded in collagen coated 6-well plates at a density of  $1 \times 10^5$  cells/well, 24 h prior to lentiviral transduction with fluorescent histone and ITGB6 pINDUCER constructs on consecutive days.

HB2 monoculture spheres were made using the methylcellulose hanging drop method described above, to contain approximately 500 cells/droplet (Table 2.2). Prior to the resuspension of collected HB2 spheroids in collagen gels, primary myoepithelial cells were first added to the neutralised collagen gel mixture, prepared in complete HuMEC medium, to give approximately 8000 myoepithelial cells per 50  $\mu$ l gel (Table 2.4). Upon sphere resuspension, 50  $\mu$ l sphere and primary myoepithelial cell gel mixture, containing approximately 6 spheres, was added per well of a 96-well plate prior to gel polymerisation at 37 °C for 30 min with plate inversion every 30 sec for the initial 5 min. Gels were overlaid with indicated drug treatments in complete HuMEC medium and incubated for 3 days (Table 2.5).

**Table 2.4 Collagen-I gel recipe for HB2/1° Myo model**

	Volume (μl)
Complete HuMEC medium	132.5
1° Myo suspension (8 x 10 <sup>4</sup> / ml)	132.5
1 M HEPES (pH 7.5)	10
Rat tail collagen type I (10 mg/ml)	215
1N NaOH	8

#### 2.1.2.4. Spheroid model quantification and image analysis

At the assay endpoint, spheres were imaged using an Axiovert 135 (Carl Zeiss MicroImaging LLC) at 10X magnification. To quantify spheres objectively, the following were determined: 1) the percentage of invasive area, 2) the number of projections, 3) the projection length ( $\mu\text{m}$ ) and 4) the central sphere area ( $\mu\text{m}^2$ ). To determine the percentage of invasive area, the following equation was used:

$$\% \text{ Invasive Area} = \frac{\text{Total area} - \text{Central area}}{\text{Central area}} \times 100$$

To determine the projection length, the distance from the margin of the core sphere to the leading end of the projection was measured. All sphere quantification was performed using Fiji Software version 2.0.

### 2.1.3. Transient cell transfections

#### 2.1.3.1. siRNA transfection

1089 cells were seeded into 6-well plates at a density of  $8 \times 10^4$  cells/well, 24 h prior to siRNA transfection. In one tube, smart pool siRNA (Horizon Discovery) was diluted in 125  $\mu$ l optiMEM (Gibco, 31985062) to give a final concentration of 20 nM. In a second tube, 3.75  $\mu$ l Lipofectamine 3000 Reagent (ThermoFisher, L3000001) was added to 125  $\mu$ l optiMEM, as indicated by manufacturer's guidelines. Diluted siRNA was then added to the Lipofectamine-optiMEM mix and incubated at room temperature (RT) for 15 min. Following incubation, the siRNA-Lipofectamine mix was added dropwise to the cells. 24 h post transfection, medium was changed. Sufficient knockdown of siRNA targets was assessed at 72 h post transfection by quantitative polymerase chain reaction (qPCR) or western blot analysis. All siRNA and their respective catalogue identifiers are listed in Table 2.5.

**Table 2.5 siRNA targets and catalogue identifiers**

<b>Target</b>	<b>Catalogue ID</b>	<b>Target</b>	<b>Catalogue ID</b>
Non-targeting control	D-001206-14	MYF6	M-009758-00
ADAMTS3	M-005773-00	MYOCD	M-015905-01
ARNT2	M-008920-00	NEUROD1	M-008667-01
ASCL1	M-008307-01	NEUROG3	M-010040-01
CBFB	M-011602-00	NFAT5	M-009618-01
CREB1	M-003619-01	NKX2-1	M-019105-01
EBF2	M-013842-01	NKX2-5	M-019795-02
EOMES	M-017483-00	NOTCH1	M-007771-02
EP300	M-003486-04	NOTCH4	M-011883-01
FN1	M-009853-01	PAX6	M-011098-01
FOXA1	M-010319-01	POU4F1	M-016061-01
FOXA2	M-010089-01	PPARGC1A	M-005111-01
GATA2	M-009024-00	PRDM1	M-009322-02
GATA6	M-008351-01	RBPJ	M-007772-00
HDAC4	M-003497-03	REST	M-006466-02
HIF1A	M-004018-05	SIM1	M-009791-00
HNF1A	M-008215-01	SMAD4	M-003902-01
HNF1B	M-009721-01	SMARCA4	M-010431-00
HTT	M-003737-02	SMARCA5	M-011478-00
ITGA5	M-008003-02	SOX3	M-012143-00
JUNB	M-003269-01	SP1	M-026959-00
KLF4	M-005089-03	SP3	M-023096-02
KMT2D	M-004828-02	SRF	M-009800-02
LEF1	M-015396-00	STAT3	M-003544-02
LHX1	M-012121-01	STAT5A	M-005169-02
MEF2A	M-009362-00	STAT5B	M-010539-02
MEF2C	M-009455-00	STAT6	M-006690-01
MKL1	M-015434-01	TCF7L1	M-014703-01
MMP13	M-005955-01	ZNF217	M-004987-00
MYB	M-003910-00		

### 2.1.3.2. DNA transfection

1089 cells were seeded into 6-well plates at a density of  $8 \times 10^4$  cells/well, 24 h prior to DNA transfection. In one tube, 2.5  $\mu$ g DNA (pLV CMV MMP13-HA plasmid) was diluted in 125  $\mu$ l optiMEM with P300 Reagent (2  $\mu$ l/ $\mu$ g DNA) (ThermoFisher, L3000001). In a second tube, 3.75  $\mu$ l Lipofectamine 3000 Reagent was added to 125  $\mu$ l optiMEM, as indicated by manufacturer's guidelines. Diluted DNA and P300 Reagent was then added to the Lipofectamine-optiMEM mix, vortexed, and incubated at RT for 15 min. Following incubation, the DNA-Lipofectamine mix was added dropwise to the cells and medium was changed 24 h post transfection. Sufficient overexpression of MMP13 was assessed at 72 h post transfection by western blot analysis.

### 2.1.3.3. SMAD-SMAD binding element luciferase reporter assay

1089 cells were seeded into 24-well plates at a density of  $2.5 \times 10^4$  cells/well 24 h prior to transfection. Cells were co-transfected with a renilla luciferase control plasmid (pRL, Promega, E2231/E2261) and a reporter plasmid encoding SMAD binding element (SBE)-firefly luciferase, gifted by Dr. A. Cameron, using Lipofectamine LTX/Plus reagent (Invitrogen, 15388-100). In one tube, 170 ng renilla luciferase plasmid was mixed with 330 ng firefly luciferase reporter plasmid in 25  $\mu$ l optiMEM. 0.5  $\mu$ l Plus reagent (ThermoFisher, A126) was added to this solution and mixed. In a second tube, 2  $\mu$ l Lipofectamine LTX reagent (ThermoFisher, A126) was added to 25  $\mu$ l optiMEM and mixed. The plasmid and Plus reagent solution was then added to the Lipofectamine LTX solution and incubated for 20 min at RT. 50  $\mu$ l transfection mix was then added dropwise to each well. The following day, medium was changed. For the TGF $\beta$  stimulated conditions, cells were serum-starved for 24 h prior to stimulation with 5 ng/ml TGF $\beta$  (Peprotech, 100-21) for 24 h (Table 2.6).

To assess the activity of renilla and firefly reporters, the Dual-Glo luciferase assay system (Promega, E2940) was used, as per manufacturer's instructions. In brief, medium was removed from cells and 70  $\mu$ l 1X PBS was added to each well. 70  $\mu$ l Dual-Glo luciferase reagent was then added to each well and incubated on a shaker at 100 rpm for 10 min at RT. 120  $\mu$ l of sample from each well was then transferred to a 96-well white flat bottom plate (Corning, 353296) and firefly luminescence was measured using a FLUOstar Optima plate reader (BMG Labtech). The following protocol was used: 2 sec pre-read delay, 10 sec measurement period, detector gain: 4095, no emission filter applied. Once firefly luminescence was measured, 60  $\mu$ l of Stop & Glo reagent was added to each well and plates were incubated on a shaker at 100 rpm for 10 min at RT. Renilla luminescence was then measured using the same protocol described for firefly luminescence. For the analysis, firefly luminescence was normalised to renilla luminescence, to control for differences in transfection efficiency, and expressed as fold change relative to the untreated control condition.

#### **2.1.4. Lentiviral production and cell transduction**

HEK293T cells were seeded in 60 mm dishes to reach 90% confluency at the time of transfection. For transfection, 3.25  $\mu$ g pCMVR8.2 (Addgene, 12263) and 1.7  $\mu$ g pMD2.G (Addgene, 12259) packaging plasmids were combined with 5  $\mu$ g of lentiviral transfer plasmid in 470  $\mu$ l of OptiMEM with 30  $\mu$ l FuGENE transfection reagent (Promega, E2311). Plasmids and FuGENE were mixed by pipetting and incubated for 10 min at RT before being added dropwise to a 60 mm dish of HEK293T cells. Medium was changed 24 h post transfection and virus-containing supernatant was collected an additional 24 h later. Viral supernatant was stored at -80 °C. Transfer plasmids used were H2B-RFP (Addgene, 26001), H2B-GFP (Addgene, 25999), HER2 pINDUCER, ITGB6 pINDUCER, pLV CMV ADAMTS3-WT and pLV CMV ADAMTS3-E399A.

24 h prior to viral infection, HB2 or 1089 cells were seeded in 6-well plates at a density of  $5 \times 10^4$  cells/well, while primary cells were seeded in collagen-I coated 6-well plates at a density of  $1 \times 10^5$  cells/well. Viral supernatant was added dropwise to the cells and incubated for 24 h. To improve transduction efficiency in primary cells, 20 mU/ml of neuraminidase (Sigma, 9001676) was added to 1 ml of viral supernatant and incubated at 37 °C for 45 min prior to transduction (248). Medium was changed 24 h post transduction and cells were selected using the appropriate antibiotic of 10 µg/ml puromycin (Sigma, P9620), or 10 µg/ml blasticidin (Sigma, SBR00022) for a minimum of 72 h.

### **2.1.5. Fibronectin cleavage validation**

10 mm coverslips were coated with 2 mg rhodamine-conjugated fibronectin (Cytoskeleton, FNR01) in a 24-well plate format. To coat coverslips, 500 µl of 4 mg/ml rhodamine-fibronectin was added to each well and incubated at 37 °C for 30 min. Excess solution was removed and plates were air dried for 30 min at RT. 1089 cells were seeded at a density of  $2 \times 10^4$  cells/well for 24 h prior to fixation with 4% formaldehyde (Fisher, 11586711) for 5 min at RT. Coverslips were mounted on slides using ProLong Gold Antifade Mountant with DAPI (ThermoFisher, P36931) to visualise nuclei.



**Table 2.6 Cell culture treatment conditions**

<b>Treatment</b>	<b>Concentration</b>	<b>Duration</b>
<b>2D</b>		
Doxycycline ( $\beta 6$ induction)	1 $\mu\text{g/ml}$	48 h
TGF $\beta$	5 ng/ml	24 h
SB435142	10 $\mu\text{M}$	48 h
SGCCBP30	1 $\mu\text{M}$	48 h
<b>3D</b>		
<b>HB2/1089 spheres</b>		
Doxycycline ( $\beta 6$ induction)	1 $\mu\text{g/ml}$	4 d
GM6001 (MMPi)	10 $\mu\text{M}$	4 d
DB04760 (MMP13i)	1 $\mu\text{M}$	4 d
SB435142 (TGF $\beta$ Ri)	10 $\mu\text{M}$	4 d
SGCCBP30 (EP300i)	1 $\mu\text{M}$	4 d
RGD peptide	10 $\mu\text{M}$	4 d
<b>HB2/1° myo spheres</b>		
Doxycycline ( $\beta 6$ induction)	1 $\mu\text{g/ml}$	3 d
GM6001 (MMPi)	10 $\mu\text{M}$	3 d
DB04760 (MMP13i)	1 $\mu\text{M}$	3 d
<b>Duct model</b>		
Doxycycline ( $\beta 6$ induction)	1 $\mu\text{g/ml}$	Between d14-21
DB04760 (MMP13i)	1 $\mu\text{M}$	Between d14-21
SGCCBP30 (EP300i)	1 $\mu\text{M}$	Between d14-21

## 2.2. Molecular cloning

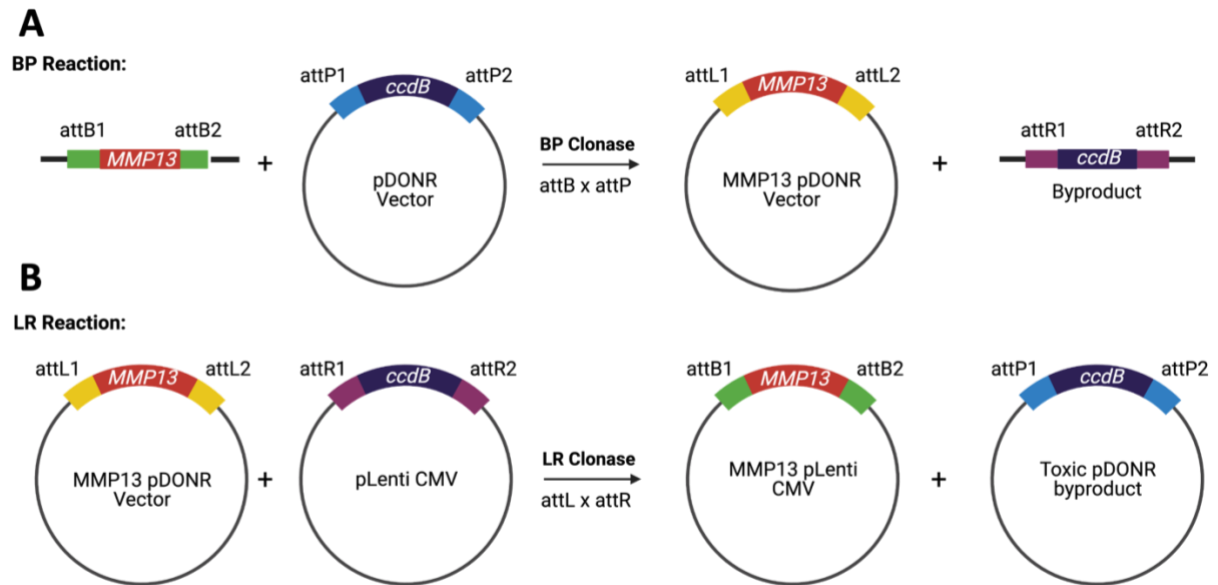
### 2.2.1. HER2 and ITGB6 expression plasmids

Inducible HER2 and ITGB6 plasmids were previously constructed by Dr. E. Carter, by cloning the *HER2* or *ITGB6* open reading frame into pINDUCER21 (Addgene, 46948) (249), using Gateway recombination described below (Fig 2.1).

### 2.2.2. MMP13 expression plasmid

#### 2.2.2.1. Gateway cloning overview

Gateway cloning is based on the integration and excision reactions utilised by lambda phage during bacterial infection (250). These reactions involve the recombination of attachment sites from the bacteria (attB) and phage (attP), which result in the integration of phage sequences into the bacterial genome. This forms two new recombination sites, attL and attR (Fig 2.1A). The attL and attR sites can subsequently recombine, leading to the excision of the phage sequences from the bacterial genome and the regeneration of attB and attP sites (Fig 2.1B). To ensure the selection of recombinant plasmids during gateway cloning, the pDONR and destination plasmids that are utilised contain the *ccdB* gene. This codes for the lethal enzyme that inhibits DNA gyrase to induce cell death. As bacterial cells containing non-recombinant vectors would die, this technology ensures that propagated bacterial colonies are from bacterial cells containing recombinant vectors (250).



**Figure 2.1** Schematic summarising gateway cloning strategy.

**(A)** The amplified attB flanked MMP13 sequence is combined with a pDONR vector that contains attP sites. The BP reaction takes place between the attB and the attP sites to generate the MMP13 pDONR entry clone containing attL sites. The *ccdB* gene, encoding a lethal enzyme is excised from the donor vector. **(B)** The LR reaction takes place between the attL sites located either side of the MMP13 sequence within the MMP13 pDONR vector, and the attR sites located in the pLV CMV destination plasmid. The reaction generates the MMP13 pLV CMV expression plasmid and a toxic pDONR byproduct containing the *ccdB* gene.

### 2.2.2.2. Amplification of MMP13 and addition of HA-Tag sequence

The Q5 High-Fidelity DNA Polymerase kit (New England BioLabs, M0491) was used to amplify the MMP13 sequence from 1089 myoepithelial cell mRNA, and subsequently add the HA-Tag sequence and attB sites, over two PCR reactions.(Table 2.7). For the first reaction, primers including the HA-Tag sequence were used to introduce the HA-Tag onto the C-terminal of the MMP13 sequence (Table 2.7). For the second reaction, attB primers were used to add attB sites to each end of the MMP13-HA sequence (Table 2.7). For each PCR reaction, 200 mM of dNTPs, 0.5 mM of each primer, 0.02 U/ml of Q5 DNA Polymerase, and 1 µg mRNA were diluted in 1X Q5 Reaction Buffer (New England BioLabs, B9027S) and topped up with nuclease-free water to a total reaction volume of 50 µl. Thermocycling conditions were as follows: Initial denaturation at 98 °C for 30 sec, annealing at 72 °C for 10 sec followed by extension at 70 °C degrees for 30 sec, for 35 cycles, and a final extension step at 72 °C for 2 min. PCR products were run on a 1% agarose gel, in Tris acetate-EDTA (TAE) buffer, to confirm successful PCR amplification of MMP13, with the presence of a band at approx. 1400 bp equating to the length of the MMP13 sequence (Fig 2.2A,B). The PCR product was excised from the DNA gel and purified using the Monarch DNA Gel Extraction Kit (New England BioLabs, T1020), following manufacturer's instructions. In brief, 4 volumes of Monarch Gel Dissolving Buffer were used per 100 mg agarose and incubated at 50 °C for 10 min until the excised gel was completely dissolved. DNA was recovered using a DNA column and washed with Monarch DNA Wash Buffer prior to elution in a volume of 6 µl.

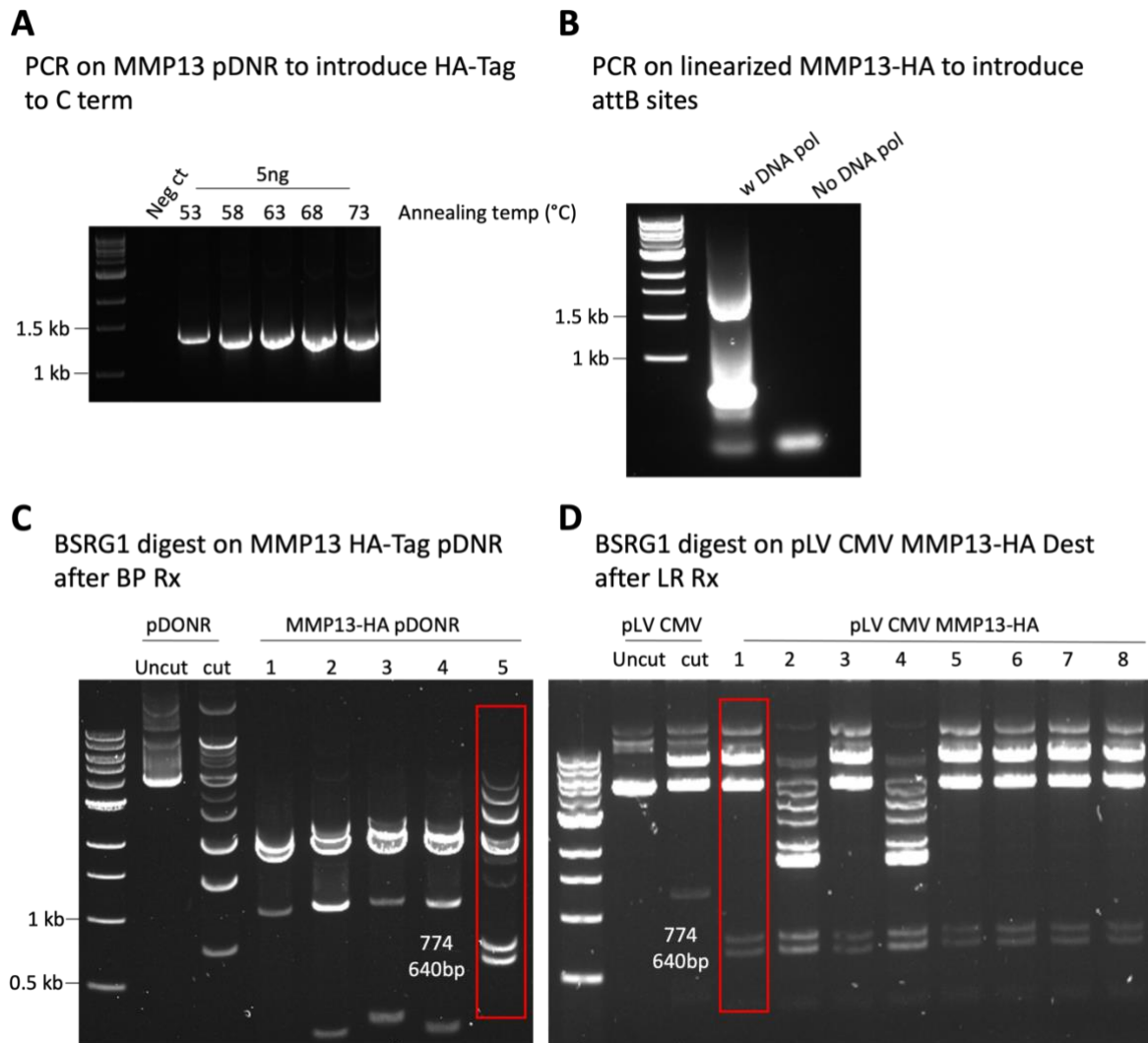
Table 2.7 MMP13 gateway cloning primers

	Gene / attB / Glycine Linker / HA-Tag / Stop Codon
MMP13-HA	Fw: GGGGACAAGTTTGTACAAAAAAGCAGGCTTC-ATGCATCCAGGGGTCTGGCTGCC Rv: TTA-AGCGTAATCTGGAACATCGTATGGGTA-CCC-ACACCACAAAATGGAATTTC
HA-attB	Fw: GGGGACAAGTTTGTACAAAAAAGCAGGCTTC-ATGCATCCAGGGGTCTGGCTGCC Rv: GGGGACCACTTTGTACAAGAAAGCTGGGTC-TTA-AGCGTAATCTGGAACATCGTATGGGTA

### 2.2.2.3. Gateway cloning of pLV CMV MMP13-HA Tag plasmid

For the BP reaction (Fig 2.1A), the Gateway BP Clonase II Enzyme mix (ThermoFisher, 11789100) was used. In summary, 100 fmol of attB-MMP13-HA-attB PCR product and 300 ng of pDONR vector (ThermoFisher, 12535035) was added to 4 µl of 5X BP Clonase reaction buffer and topped up to a volume of 16 µl in Tris-EDTA (TE) buffer (pH 8.0). 4 µl of BP Clonase enzyme mix was added and the reaction was incubated at 25 °C for 1 h. 4 µg Proteinase K solution (New England Biolabs, P8107S) was then added and incubated at 37 °C for 10 min to terminate the reaction. Competent 10-beta E. coli (New England BioLabs, C3109) were transformed, and Zeocin-resistant clones were selected for plasmid uptake and effective recombination (see 2.2.4 bacterial transformation step). Following plasmid extraction, a BSRG1 diagnostic digest (New England BioLabs, R3575) was performed by combining 1 µg of DNA with 20 units of BSRG1 (New England BioLabs, R3575) in rCutSmart Buffer (New England BioLabs, B6004) for 15 min at 37 °C. Products were run on a 1% agarose gel (in TAE buffer) to confirm successful incorporation of the MMP13-HA sequence into the pDONR vector, via the presence of two closely migrating bands at approx. 640 and 770 bp (Fig 2.2C).

For the LR reaction (Fig 2.1B), the Gateway LR Clonase II Enzyme mix (ThermoFisher, 11791020) was used. In summary, 100 ng of MMP13 pDONR was mixed with 150 ng of pLV CMV DEST vector (Addgene, 17452) and topped up with TE buffer (pH 8.0) to a volume of 8 µl. 2 µl of LR Clonase II enzyme mix was then added and the reaction was incubated at 25 °C for 1 h. 2 µg Proteinase K solution was then added and incubated at 37 °C for 10 min to terminate the reaction. Competent 10-beta E. coli were then transformed, and ampicillin-resistant clones were selected for (see 2.2.4 bacterial transformation step). To confirm successful cloning of the MMP13-HA tag sequence in to the pLV CMV DEST plasmid, a BSRG1 diagnostic digest was performed as above (Fig 2.2D).



**Figure 2.2 Gateway cloning and construction of pLV CMV MMP13 plasmid.**

**(A)** Amplification of MMP13 to introduce HA-Tag to C terminus by PCR. Presence of MMP13 sequence is detected by a band at approx. 1.4kb. **(B)** PCR amplification of MMP13-HA Tag sequence to introduce attB sites. **(C)** DNA products following BSRG1 diagnostic digest of MMP13-HA pDONR (after BP reaction). Presence of MMP13 sequence is detected by 2 bands at 774 and 640 bp. **(D)** DNA products following BSRG1 diagnostic digest of pLV CMV MMP13-HA destination plasmid (after LR reaction). Presence of MMP13 sequence is detected by 2 bands at 774 and 640 bp. Red boxes indicate plasmids taken forward for subsequent steps.

### 2.2.3. ADAMTS3 expression plasmids

The pLV CMV ADAMTS3-WT and ADAMTS3-E399A plasmids were designed and manufactured using Vector Design Studio via [VectorBuilder](https://en.vectorbuilder.com/) (<https://en.vectorbuilder.com/>). For the catalytically inactive ADAMTS3 mutant, the Uniprot database was used to identify the essential amino acid, required for proteolytic activity, within the ADAMTS3 active site. The glutamic acid (E) amino acid at position 399 was converted into alanine (A), with codon GAA converted to GCA.

### 2.2.4. Bacterial transformation and plasmid extraction

For the bacterial transformation of constructs, competent 10-beta E. or Stbl3 E. coli cells were used. For each transformation reaction, approximately 5 ng of plasmid DNA was added to 50 µl competent bacteria and incubated on ice for 30 min. Following this, the mixture was heat shocked at exactly 42 °C for 30 sec and placed back on ice for 5 min. 950 µl of Stable Outgrowth Medium (New England Biolabs, B9035S) was then added to the mixture and incubated at 37 °C for 1 h with agitation at 250 rpm. After this, 50 µl of the mix was spread onto 3% Luria Bertani (LB) agar (Sigma-Aldrich, L3147) plates containing appropriate antibiotic and incubated overnight at 37 °C. Plasmids constructed by VectorBuilder, were provided as bacterial stabs, and streaked to enable growth of single colonies. Antibiotics for selection included 400 µg/ml zeocin (Fisher, R25001) and 50 µg/ml ampicillin (ThermoFisher, 11593027). Colonies were picked using a P2 tip and propagated in 5 ml of LB broth containing appropriate antibiotic at 37 °C for 16-20 h at 250 rpm. Plasmids were extracted from bacterial cultures using the Monarch plasmid miniprep kit (New England Biolabs, T1010), following manufacturer's guidelines. Between 1-5 ml of bacterial culture was pelleted by centrifugation prior to resuspension with Plasmid Resuspension Buffer and subsequent lysis with Plasmid Lysis Buffer.



Following neutralisation, plasmids were recovered using Plasmid Miniprep Columns and washed with ethanol (EtOH)-based Plasmid Wash Buffers prior to elution of plasmids in a 30 µl volume.

## 2.3. Chromatin immunoprecipitation (ChIP)

### 2.3.1. Chromatin isolation

$7 \times 10^5$  1089 cells were seeded into 150 mm dishes per condition to allow for 48 h of treatment with indicated agents and for cells to achieve confluency prior to harvesting (Table 2.6). For the ChIP workflow, the SimpleChIP Plus Sonication Chromatin IP Kit (Cell Signalling, 56383) was used, following manufacturer's guidelines. In brief, to crosslink proteins at the time of harvest, 1% formaldehyde was added directly to the medium-containing dishes and incubated for 10 min at RT. To quench the fixation reaction, 1X glycine (Cell Signalling, 7005) was added for 5 min at RT, before washing cells twice in ice-cold 1X PBS. To collect the chromatin, cells were scraped in PBS containing 1X protease cocktail inhibitor (Cell Signalling, 7012). Cells were then centrifuged at  $1,000 \times g$  for 5 min at  $4^\circ\text{C}$  and the supernatant discarded.

The pellet was resuspended in ChIP lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.0 plus 1X protease inhibitor cocktail) using a syringe with a 27-gauge needle and incubated on ice for 30 min. Samples were transferred to 15 ml TPX tubes with beads (Diagenode, C01020031) and a Bioruptor pico sonicator (Diagenode, B01060010) was used to fragment the chromatin for 5 cycles (30 seconds ON and 30 seconds OFF). Lysates were then clarified by centrifugation at  $21,000 \times g$  for 10 min at  $4^\circ\text{C}$  and the supernatant collected. Following DNA extraction using the DNA Purification Kit (Cell Signaling, 14209), a 10  $\mu\text{l}$  sample was then removed to examine successful fragmentation, defined by the presence of DNA fragments between 200 to 1000bp, on a 1% agarose gel (in TAE buffer).

### 2.3.2. Immunoprecipitation

Upon satisfactory fragmentation of chromatin, 5 µg of sonicated, cross-linked chromatin per condition was diluted in a 1:4 ratio in 1X ChIP buffer (Cell Signalling, 7008S) plus 1X protease inhibitor cocktail (Cell Signalling, 7012) and subjected to immunoprecipitation with 2.5 µg of H3K27ac antibody (Abcam, 4729) overnight at 4 °C with 250 rpm rotation. 30 µl of ChIP-Grade Protein G Magnetic Beads (Cell Signaling, 9006) were then added to each IP reaction and incubated for 2 h at 4 °C with rotation. To clear non-specific binding, the bound beads were subjected to a series of low and high salt washes (Cell Signalling, 7008), before elution of antibody-protein-DNA complexes in 1X ChIP Elution Buffer (Cell Signalling, 7009) at 65 °C for 30 min with gentle vortexing at 1,200 rpm. The enriched chromatin samples were then incubated at 65 °C for 2 h with 1X Proteinase K (Cell Signaling, 10012) to reverse cross-links prior to purification of DNA using the DNA Purification Kit (Cell Signaling, 14209) as per manufacturer's guidelines. DNA was eluted in 50 µl of 1X ChIP Elution Buffer.

### 2.3.3. ChIP-qPCR

2 µl of extracted DNA and 0.25 µM of forward and reverse primers were used per ChIP-qPCR reaction using the Luna Universal qPCR Master Mix (New England BioLabs, M3003), following manufacturer's guidelines. Reactions were carried out on a Step One Plus Instrument (Applied Biosystems, 4376600) with cycling conditions as follows: Initial denaturation at 95 °C for 1 min, denaturation at 95 °C for 15 sec, annealing at 60 °C for 15 sec and extension at 72 °C for 30 sec, for 43 cycles. Genomic MMP13 primer sequences are detailed in Table 2.8. For the analysis, data were presented as a percent of the total input chromatin using the equation below, where Ct is the average threshold cycle of the PCR reaction (251).

$$\text{Percent Input} = 2\% \times 2^{(\text{Ct of 2\% input} - \text{Ct of IP sample})}$$

**Table 2.8 Genomic chIP qPCR primers**

Target	Sequences	Product Length (bp)
MMP13	Fw: CACAGGCCACTTGAGAGGTT	101
	Rv: AAGGTTGGTGGTGAAAGTGAGA	

## **2.4. Transcriptomic analysis of cells**

### **2.4.1. PCR**

#### **2.4.1.1. RNA extraction**

RNA was harvested from cells using the Monarch Total RNA Miniprep Kit (New England BioLabs, T2010), following manufacturer's guidelines. In summary, cells were lysed in 300 µl of RNA Lysis Buffer (New England BioLabs, T2012) then combined with equal parts of 100% EtOH to precipitate RNA. The RNA was added to an RNA extraction column and washed three times with EtOH-based RNA Priming and Wash buffers (New England BioLabs, T2013, T2014) prior to elution in 30 µl nuclease free water. The concentration of extracted RNA was determined using a Nanodrop Spectrophotometer (Thermo Scientific, ND-ONE-W).

#### **2.4.1.2. cDNA synthesis**

Using between 500 ng - 1 µg of extracted RNA as a template, cDNA was synthesised by reverse transcription using the Luna Script RT SuperMix Kit (New England BioLabs, E3010), following manufacturer's guidelines. For each reaction, RNA was combined with 4 µl of 5X LunaScript RT SuperMix and topped up to 20 µl with nuclease free water. The reaction was performed with the following conditions: primer annealing for 2 min at 25 °C, cDNA synthesis for 10 min at 55 °C and heat inactivation for 1 min at 95 °C.

#### 2.4.1.3. Quantitative PCR (qPCR)

The Luna Universal qPCR Master Mix (New England BioLabs, M3003) was used to perform qPCR, following manufacturer's guidelines. For each reaction, 50 ng of cDNA, 5 µl of 2X Luna Universal qPCR Master Mix (New England BioLabs, M3003) and 0.25 µM of each forward and reverse primer was combined with nuclease free water to a total volume of 10 µl. Reactions were carried out on a Step One Plus Instrument (Applied Biosystems) with cycling conditions as follows: Initial denaturation at 95 °C for 1 min, denaturation at 95 °C for 15 sec, annealing at 60 °C for 15 sec and extension at 72 °C for 30 sec, for 43 cycles. End products were run on a 1% agarose gel (in TAE) to confirm primer-dependent amplification of intended targets. All primer sequences are detailed in Table 2.9. Relative gene expression was calculated using the  $\Delta C_t$  and  $2^{-\Delta\Delta C_t}$  method (252), normalising to a ACTB reference.

**Table 2.9 List of qPCR primers**

<b>Target</b>	<b>Sequences</b>	<b>Product length (bp)</b>
ACTB	Fw: AGAGCTACGAGCTGCCTGAC	184
	Rv: AGCACTGTGTTGGCGTACAG	
ADAMTS3	Fw: CTCGACTCGTGATGGTTCTCC	97
	Rv: ATTTCTTCATTACCAGCTTGTCCA	
EP300	Fw: GCAGTGTGCCAAACCAGATG	105
	Rv: GCAGTGTGCCAAACCAGATG	
FN1	Fw: CCGGGACTCAATCCAAATGC	150
	Rv: TCCAGGAACCCTGAACTGTA	
ITGB6	Fw: AGCCTCTCAGTGTAGGCAGA	91
	Rv: GCACCACCTGGTCTCAACTT	
MMP13	Fw: AACATCCAAAAACGCCAGAC	155
	Rv: GGAAGTTCTGGCCAAAATGA	

## **2.4.2. RNAseq**

### **2.4.2.1. RNAseq workflow and differential expression analysis**

RNA was extracted as previously described (see 2.4.1.1 RNA extraction). For 1089 samples, RNA from two biological replicates were isolated. For primary myoepithelial cells, RNA was extracted from two independent reduction mamoplasty donors. All subsequent RNA sequencing steps were performed by Dr. C. Mein at Barts Genome Centre. RNA quality was determined using the Agilent Bioanalyser 2100, and an RNA integrity number (RIN) was generated based on the ratio of 28S:18S ribosomal RNA. For the mRNA library preparation, mRNA was isolated from total RNA and primed using AMPure XP beads to enable first and subsequent second strand cDNA synthesis. Double-stranded cDNA was then purified prior to end prep of the cDNA library. Adaptor ligation was performed, and ligation was purified using AMPure XP beads. The library quality was then assessed using a Bioanalyser (Agilent High Sensitivity Chip). RNA sequencing was performed using the Illumina NextSeq 500 platform. Reads were aligned using STAR (version 2.7.3a) against GENCODE Homo Sapiens Hg38.

Normalisation of counts and differential expression analysis was performed using DESeq2 (253) by E. Tomas Bort. All RNAseq data generated during the study are publicly available and have been deposited in Gene Expression Omnibus with the accession code GSE224401.

### **2.4.2.2. Gene set enrichment analysis and upstream regulator analysis**

The resulting matrix was imputed into the gene set enrichment analysis (GSEA) Broad Institute to examine the Hallmark gene set (254) and generate normalised enrichment scores (NES). To perform



upstream regulator analysis, Ingenuity Pathway Analysis software (QIAGEN) was used to determine Activation Z scores of predicted transcriptional regulators, as described (255).

## 2.5. Proteomic analysis of cells

### 2.5.1. Immunofluorescence

#### 2.5.1.1. 2D Immunofluorescence

Cells cultured on 13 mm glass coverslips were fixed in 4% formaldehyde for 10 min, and subsequently washed twice with 1X PBS to remove residual formaldehyde. Cells were permeabilised in 0.1% Triton-X100-PBS (Alfa Aesar, A16046) for 5 min and blocked in 5% bovine serum albumin (BSA) (Sigma-Aldrich, A8022) for 30 min, prior to incubation with primary antibody diluted in 5% BSA-PBS for 1 h at RT, with cells washed twice with 1X PBS between steps. Samples were then subsequently incubated with species-appropriate fluorescent secondary antibodies diluted in 5% BSA-PBS for 1 h at RT, washed twice with PBS, before being mounted using Molecular Probes ProLong Gold Antifade Mountant with 4,6-diamidino-2-phenylindole (DAPI) (Life Technologies, P36931). Samples were air-dried overnight and imaged using a LSM710 Zeiss confocal microscope (Carl Zeiss MicroImaging LLC). All antibodies used for immunofluorescence are listed in Table 2.10.

To quantify intensity of nuclear SMAD4 staining, image colour channels were split. The DAPI channel was then converted to a binary image and the nuclear compartment outlined. These outlines were then overlaid to the channel representing SMAD4 immunofluorescence, to exclude the measurement of any cytoplasmic staining. The mean fluorescence intensity per nucleus was then calculated. All nuclear intensity quantification was performed using Fiji Software version 2.0.

### 2.5.1.2. 3D Immunofluorescence

For immunofluorescence of spheres, collagen gels were fixed in 4% formaldehyde for 30 min, permeabilised in 0.1% Triton-X100-PBS and blocked in 5% BSA-PBS, each for 1 h, prior to incubation with primary antibody diluted in 5% BSA-PBS for 48 h in 4 °C with agitation. Gels were then incubated with species-appropriate fluorescent secondary antibodies diluted in 5% BSA-PBS for 2 h at RT. Gels were washed twice in 1X PBS between steps, prior to mounting on slides with mowiol. To prevent gels from being compressed, electrical tape was used to make a platform on the slide, with the gel placed inside and a coverslip on top. Immunofluorescent images were acquired using a LSM710 Zeiss confocal microscope and reconstructed as max intensity Z-stacks. All antibodies used for immunofluorescence are listed in Table 2.10.

**Table 2.10 Antibodies for immunofluorescence**

<b>Protein Target</b>	<b>Cat no. and supplier</b>	<b>Host species</b>	<b>Dilution</b>
Active integrin $\alpha 5$ (SNAKA1)	#MABT201 (Sigma-Aldrich)	Mouse, monoclonal	1:100
Cleaved collagen	#0217-050 (Immunoglobine)	Rabbit, polyclonal	1:100
Collagen IV	#6586 (Abcam)	Rabbit, polyclonal	1:200
Fibronectin	#F0791 (Sigma)	Mouse, monoclonal	1:100
Laminin	#L9393 (Sigma)	Rabbit, polyclonal	1:200
pFAK (Y397)	#3283 (Cell Signalling)	Rabbit, polyclonal	1:100
SMAD4	#SC7966 (Santa Cruz)	Mouse, monoclonal	1:200
Anti-mouse 647	#A32787 (Invitrogen)	Donkey, polyclonal	1:500
Anti-rabbit 405	#A31556 (Invitrogen)	Goat, polyclonal	1:500
Anti-rabbit 647	#A32795 (Invitrogen)	Donkey, polyclonal	1:500

## **2.5.2. Western blotting**

### **2.5.2.1. Protein extraction**

Prior to protein extraction, cells were washed twice with ice-cold 1X PBS to remove residual medium. To extract protein, cells were lysed with NP-40 lysis buffer (50mM Tris pH 7.5, 150mM NaCl, 1% NP40), supplemented with phosphatase and protease inhibitor cocktails (EMD Millipore, 539131 and 524625) and scraped using a cell scraper. For cells harvested from a 6-well plate, 150  $\mu$ l lysis buffer was used. The lysate was collected and centrifuged for 10 min at 20,000 x g at 4 °C to pellet cell debris. The protein supernatant was collected and quantified using the Bio-Rad DC Protein Assay Kit (BioRad, Solution A 5000113, solution B 5000114 and solution S 5000115), following manufacturer's guidelines. Briefly, 20  $\mu$ l of reagent S was added to 980  $\mu$ l of reagent A to make the working solution. 25  $\mu$ l working solution was then added to 5  $\mu$ l of the relevant protein standard or sample. 200  $\mu$ l of reagent B was then added and incubated for 5 min at RT. Absorbance was measured at 620 nm, using a 96-well microplate reader (Infinite F50, Magellen software).

### **2.5.2.2. Western Blot**

Prior to loading, 30  $\mu$ g protein samples were prepared with sample buffer to a final concentration of 60 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 0.5 M dithiothreitol (DTT) and 0.01% bromophenol blue and incubated for 5 min at 95 °C to denature proteins. Samples were run on a 10% SDS-PAGE gel and electrophoresed for 1.5 h at 120 V in running buffer (25 mM Tris-HCl, 192 mM glycine and 0.1% SDS) using an XCell II Blot Module (Bio-Rad). Proteins were then transferred onto a nitrocellulose membrane at 25 V for 2 h in transfer buffer (12 mM Tris at pH 8.3, 96 mM glycine and 20% methanol). For detection of high molecular weight protein EP300, NuPAGE 3-8% Tris-Acetate Gels (ThermoFisher,

EA0375PK2) were used with NuPAGE Tris-Acetate SDS Running Buffer (ThermoFisher, LA0041) to separate proteins, and transferred onto membranes using NuPAGE Transfer Buffer (ThermoFisher, NP0006). To confirm efficient transfer, membranes were incubated in Ponceau S solution (Sigma, P7170), prior to blocking in 5% milk in Tris buffered saline (TBS) with 0.5% Tween 20 (TBST) (Severn Biotech, 20630110) for 1 h.

Membranes were incubated with primary antibodies in 5% milk in TBST at 4 °C overnight. To remove unbound antibodies, membranes were washed three times in TBST for 10 min per wash, prior to incubation in species-appropriate horse radish peroxidase (HRP)-conjugated secondary antibodies for 1 h at RT. Excess secondary antibody was removed by washing in TBST three times for 10 min per wash, prior to protein detection. Bands were developed using Immobilon Forte Western HRP Substrate (Merck Millipore, WBLUF0) and visualised using an Amersham Imager 600 (GE Healthcare). All antibodies used for western blotting are listed in Table 2.11.

**Table 2.11 Antibody for western blotting**

<b>Protein Target</b>	<b>Cat no. and supplier</b>	<b>Host species</b>	<b>Dilution</b>
EP300	#86377 (Cell signalling)	Rabbit, monoclonal	1:1000
Fibronectin	#F0791 (Sigma)	Mouse, monoclonal	1:1000
GAPDH	#Mab374 (EMD Millipore)	Mouse, monoclonal	1:2000
HA-Tag	#3724 (Cell signalling)	Rabbit, polyclonal	1:1000
HSC70	#SC7298 (Santa Cruz)	Mouse, monoclonal	1:2000
Integrin $\beta$ 6	#SC6632 (Santa Cruz)	Goat, polyclonal	1:1000
pSMAD2/3 (Ser423/425)	#SC11769 (Santa Cruz)	Goat, polyclonal	1:1000
SMAD2	#3103 (Cell signalling)	Mouse, monoclonal	1:1000
SMAD4	#SC7966 (Santa Cruz)	Mouse, monoclonal	1:1000
Anti-mouse HRP	#P0447 (Dako)	Goat, polyclonal	1:2000
Anti-rabbit HRP	#P0448 (Dako)	Goat, polyclonal	1:2000
Anti-goat HRP	#P0449 (Dako)	Rabbit, polyclonal	1:2000

### 2.5.3. Degradomic workflow

#### 2.5.3.1. Preparation of media and lysates

To identify candidate ADAMTS3 substrates, the terminus amine isotopic labelling of substrates (TAILS) workflow was used to analyse changes in the abundance of cleavage events following ADAMTS3 knockdown (Fig 2.3A) (256). In this method, N-termini are isotopically labelled prior to the depletion of internal tryptic peptides, allowing for the simultaneous detection of both natural and protease generated neo-N-termini in a given sample (Fig 2.3A) (256).

7 x 10<sup>5</sup> primary myoepithelial cells were cultured in DMEM/F12 medium, supplemented with EGF and insulin, and without phenol red and BPE, for 48 h with indicated siRNA treatments, prior to the collection of supernatant. Next, 1X phenylmethanesulfonylfluoride (PMSF) and 2 mM EDTA were added to the collected medium, followed by centrifugation at 5,000 x g for 30 min at 4 °C to remove cell debris. Medium was sterile filtered through a 0.22 µm filter, concentrated and buffer exchanged to 4 M Guanidine hydrochloride (GuHCl) (Sigma-Aldrich, G3272), 250 mM HEPES buffer pH 7.8, using Amicon Ultra-0.5 3kDa cut-off centrifugal filter units (Sigma-Aldrich, UFC900324), following manufacturer's instructions. The protein concentrations of the samples were determined by the Bio-Rad DC Protein Assay Kit following manufacturer's guidelines.

For lysates (Fig 2.3B), cells were washed with PBS, prior to scraping with lysis buffer (4 M GuHCl, 250 mM HEPES pH 7.8, 2 mM EDTA, 1X protease inhibitor cocktail), and the volume was adjusted empirically to have a final concentration of between (1 µg/µl and 5 µg/µl). Lysates were spun at 17,000 x g for 10 min. Protein concentration of media and lysate samples were determined by Bio-Rad DC Protein Assay Kit, following manufacturer's guidelines.



### 2.5.3.2. TAILS workflow

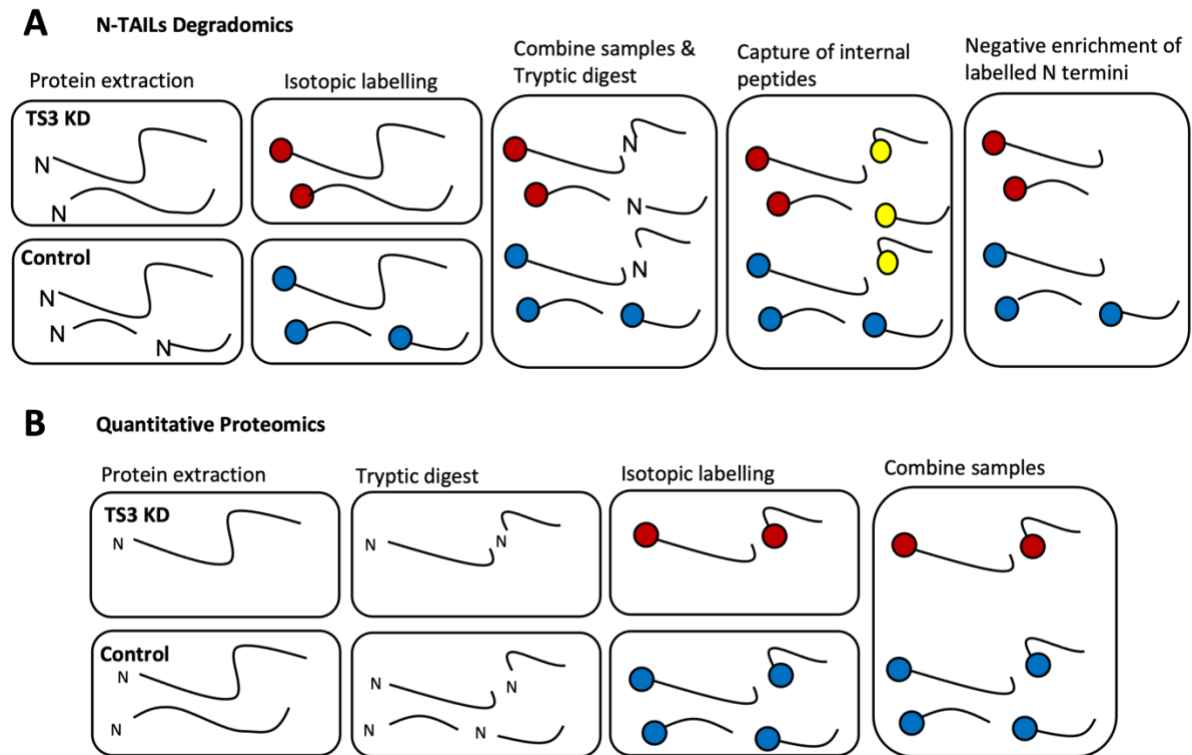
The tandem mass tag (TMT)-TAILS degradomics workflow was performed by Dr. E. Madzharova at Technical University of Denmark, as previously described (257). Three non-targeting control (NTC) and three ADAMTS3 knockdown (TS3<sup>KD</sup>) samples (50 µg total protein per sample) were subjected to TAILS. Initially, all samples were diluted to a concentration of 2.5 M GuHCl and 250 mM HEPES pH 7.8. Samples were denatured at 65 °C for 15 min, the cysteine residues were reduced with 3.5 mM tris (2-carboxyethyl) phosphine (TCEP) for 45 min at 65 °C, and then alkylated with 5 mM of chloroacetamide (CAA) for 30 min at 65 °C. Next, all samples were labelled with TMT tags at 1:4 protein:TMT ratio (TMT18plex Kit; Thermo Fisher Scientific, Waltham, MA) for 1.5 h at room RT, and subsequently quenched by adding 100 mM NH<sub>4</sub>HCO<sub>3</sub> for 30 min at RT. We then combined and precipitated the labelled samples with seven sample volumes of ice-cold acetone and one sample volume of ice-cold methanol for 2 h at -80 °C. The pooled samples were centrifuged at 4700 x g at 4 °C for 30 min, subsequently washed with 5 ml ice-cold methanol and centrifuged again. We air-dried the pellet, resuspended in 100 mM NaOH, adjusted to 1 mg/ml in 100 mM HEPES, pH 7.8, and digested with trypsin (Trypsin Gold, Promega, V5280) at a ratio of 1:50 enzyme:protein for 16 h at 37 °C at 350 rpm. Following the trypsin digestion, we took out 10% of the sample and stored it at -20 °C as a preTAILS sample. The remaining sample was incubated with a hyperbranched aldehyde-derivatized polyglycerol (HPG-ALD) polymer at a 4-fold excess in the presence of 50 mM sodium cyanoborohydride (NaBH<sub>3</sub>CN) at pH 6-7 for 16 h at 37 °C. We recovered the unbound peptides by filtration using 30 kDa Amicon Ultra-0.5 ml Centrifugal Filter Units at 10,000 x g for 10 min at RT. The polymer was washed with 30 µl 100 mM NH<sub>4</sub>HCO<sub>3</sub> and centrifuged again. The flow-throughs of the enriched N-terminal peptides (TAILS samples) were combined and stored at -20 °C.

The preTAILS and TAILS samples were analysed on an Orbitrap Exploris™ 480 mass spectrometer, interfaced with the FAIMS, and coupled with an EASY-nLC™ 1200 System. The samples were separated on an EASY-Spray™ HPLC 15 cm column (3 µm particle size, 75 µm diameter) using 140 min linear gradient from 6-60% solvent B (80% acetonitrile; 0.1% formic acid) at a constant flow of 4 µl/min. A full MS scan ( $m/z$  375–1500) was acquired at a resolution of 120,000 with a normalized AGC target of 300% and auto fill time. High-energy collision-induced MS/MS spectra were recorded in data-dependent acquisition mode using normalised collision energy of 32%. The fragmentation was performed at resolution 45,000 with a normalised AGC target of 100% and auto injection time, using a precursor isolation window of 0.7  $m/z$ .

#### **2.5.3.3. Data processing and pathway enrichment analyses**

TAILS analyses were performed by Prof. U. auf dem Keller at Technical University of Denmark. The preTAILS and TAILS samples were analysed using Proteome Discover 3.0 software. The data was searched against the human UniProt database (taxid: 9606), using the following parameters: semi-Trypsin\_R for enzyme specificity, allowing one missed cleavages; carbamidomethyl (C) and TMTpro(K) as fixed modifications, and oxidation (M), deamidation(N), TMTpro (N-Term), pyroQ(N-term) and acetyl (N-Term) as variable modifications; precursor mass error tolerance set to 10 ppm, and fragment mass error set to 0.02 Da. The proteins and peptides false discovery rate (FDR) was set to 1% for high confidence, and 5% for medium confidence. The data was normalised to 'Total Peptide Amount', excluding TMTpro N-terminally labelled peptides. Protein quantification was done using all peptides, protein ratio was estimated with 'Protein Abundance Based', and the proteins differential abundance was determined using ANOVA (Individual Proteins).

Enrichment analysis of candidate ADAMTS3 candidate substrates was performed using Search tool for the retrieval of interacting genes/proteins (STRING) Software (<https://string-db.org>) (258), while enrichment of quantitative proteomics data was performed using the GSEA Broad Institute to examine enriched pathways within the Reactome database (254).



**Figure 2.3 Schematic summarising proteomic workflow.**

**(A)** Conditioned medium was extracted from control and ADAMTS3<sup>KD</sup> cells prior to isotopic labelling of natural and neo-N termini. Isotopically labelled samples were combined and digested with trypsin. Unlabelled N-termini after the tryptic digest were removed using a hyperbranched polymer to enable negative enrichment of isotopically labelled N-termini. Enriched N-termini peptides were detected by liquid chromatography mass spectrometry. **(B)** Proteins were extracted from control and ADAMTS3<sup>KD</sup> cells and were subjected to tryptic digest. Digested peptides for each condition were isotopically labelled and combined prior to peptide detection by liquid chromatography mass spectrometry.

## **2.6. Clinical breast tissue**

### **2.6.1. Ethical approval and cohort data**

Human breast samples were obtained from surgical specimens from women undergoing breast surgery at Barts Health NHS Trust London. Written consent was obtained and ethical approval was issued by the Cambridge Central Research Ethics Committee (REC) with ethics approval number 21/EE/0072. Patient data are described in Tables 2.11 and 2.12. Patient data are listed in Table 2.12 and 2.13.

Table 2.12 Age at surgery and diagnosis of patients

	Normal	DCIS	IDC
<b>Cohort</b>	10	10	10
<b>Age at Surgery</b>	35 (20-52)	N/A	N/A
<b>Age at Diagnosis</b>	N/A	55 (43-60)	57 (46-67)

Table 2.13 DCIS and IDC grade

Grade	DCIS	IDC
<b>Low</b>	1	2
<b>Low/Inter</b>	0	1
<b>Inter</b>	3	2
<b>Inter/High</b>	3	0
<b>High</b>	3	5

### 2.6.2. RNAscope

All RNAscope assays were performed following supplier guidelines (ACD, Biotechne) using the RNAscope 2.5 HD Reagent Kit-RED (Biotechne, 322350) and RNAscope 2.5 HD Detection Reagents-RED (Biotechne, 322360). Sections were cut at 5  $\mu$ m thickness onto Superfrost Plus Slides (Fisher, Scientific, 12-550-15). To deparaffinise sections, slides were baked at 60 °C for 1 h prior to incubation in 100% xylene (Fisher Scientific, X/2050) (2 x 5 min) and 100% EtOH (Fisher Scientific, E/0665DF) (2 x 1 min). Sections were then incubated in RNAscope hydrogen peroxide buffer (ACD, 322335) for 10 min at RT prior to target retrieval (ACD, 322000) using the hot plate method for 15 min at 98-102 °C, with distilled water washes between pretreatments. Sections were treated with Protease plus reagent (ACD, 322331) for 30 min at 40 °C (in a HybEZ oven) prior to incubation with PPIB positive (Biotechne, 313908), DapB negative (Biotechne, 310043) and MMP13 target (Biotechne, 482371) RNAscope probes (2 h at 40 °C). Signals were then amplified using amplification probes: AMP1 (30 min at 40 °C), AMP2 (15 min at 40 °C), AMP3 (30 min at 40 °C), AMP4 (15 min at 40 °C), AMP5 (45 min at RT) and the alkaline phosphatase labelled probe, AMP6 (15 min at RT), with wash buffer washes between incubations. Slides were then incubated with Fast Red for 10 min at RT and counterstained with 50% Gill's haematoxylin (Sigma-Aldrich, GH132) for 3 min at RT, followed by mounting in Vectamount mounting medium (Vector Labs, H500060).

#### 2.6.2.1. Quantification of MMP13 positive ducts

Slides were imaged on the 210 NanoZoomer (Hamamatsu, C13239). For quantification, scans were uploaded into QuPath software as 'Brightfield other' images. Channel 1 was used to detect haematoxylin staining while channel 2 was set to detect red staining (MMP13). Cell segmentation was performed using 'cell detection' while chromagen detection was performed using 'subcellular

detection' with thresholds set according to staining intensity. For DCIS, 12 DCIS ducts per patient were annotated and the number of MMP13 positive cells per duct were recorded.

### **2.6.3. Immunohistochemistry**

Formalin fixed, paraffin embedded sections were dewaxed in xylene (2 x 5 min) and rehydrated through an EtOH series of 100% (2 x 2 min), 80% (1 x 2 min), 70% (1 x 2 min), 50% (1 x 2 min) and distilled water (1 x 2 min) prior to treatment with 2% hydrogen peroxide (Sigma, H1009) in methanol (1 x 10 min) to quench endogenous peroxidase activity. Enzymatic antigen retrieval to unmask epitopes was carried out using Pepsin Reagent (Sigma, R2283) for 12 min at 37 °C. Samples were then blocked in 2.5% horse or goat serum, depending on secondary antibody species, in 5% BSA (Sigma, A8022) in PBS for 30 min at RT. Sections were subsequently incubated in primary antibody overnight at 4 °C, followed by PBS washes to remove any unbound primary antibody. Species appropriate biotinylated horse or goat anti-mouse/ anti-rabbit secondary antibody (Vectastain, PK4002, PK4001) was then added to samples for 30 min at RT. All antibodies, as well as corresponding dilutions are listed in Table 2.14. Signal was developed using ABC reagent to add an avidin-biotin layer, and then 3,3'-diaminobenzidine (DAB) (Vectastain, SK4100) to view positive staining. Sections were then counterstained with Mayer's haematoxylin (Sigma-Aldrich, MHS1) and dehydrated through graded EtOH series prior to mounting with distyrene-tricresyl (DPX) (Sigma, 06522).

#### **2.6.3.1. Quantification of integrin $\beta$ 6 positive ducts**

Slides were imaged on the 210 NanoZoomer (Hamamatsu, C13239). Scans were uploaded into QuPath software as 'Brightfield DAB' images. Channel 1 was used to detect haematoxylin staining while channel 2 was set to detect DAB staining (integrin  $\beta$ 6). Cell segmentation was performed using 'cell



detection' while chromagen detection was performed using 'subcellular detection' with thresholds set according to staining intensity. For DCIS, 12 ducts per patient were annotated and the number of integrin  $\beta 6$  positive cells per duct were recorded.

**Table 2.14 Antibodies for immunohistochemistry**

Protein target	Cat no. and supplier	Host species	Dilution
Integrin $\beta$ 6	#407317 (Calbiochem)	Mouse, monoclonal	1:750
$\alpha$ SMA	#M0851 (Dako)	Mouse, monoclonal	1:500

## 2.7. Animal experiments

### 2.7.1. Animals

Animal experiments were performed by Dr. A. Agnoletto at The École polytechnique fédérale de Lausanne. NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ (NSG) mice were purchased from Charles River. Animal experiments were performed in accordance with protocols approved by the Service de la Consommation et des Affaires Vétérinaires of Canton de Vaud, Switzerland (VD 1865.5).

### 2.7.2. Intraductal xenograft

Three 8 to 12-week-old NSG female mice were anaesthetised by intraperitoneal injection with 10 mg/kg xylazine and 75 mg/kg ketamine (Graeb). After shaving and disinfecting the area around the nipple, the intraductal injection was performed by injecting 500,000 MCF7 cells (resuspended in 10 µl of PBS) into the cleaved 3rd and 4th teats with a blunt end Hamilton syringe (HAMI80508, specifications: 50 µl 705 N, gauge 30/13 mm/pst3), as previously described (259). MCF7 cells were grown intraductally for 12 weeks to form mammary tumours. At the end of the period, mice were euthanised by CO<sub>2</sub> inhalation. Engrafted mammary glands were harvested and fixed in 4% formaldehyde for 2 h, prior to dehydration and paraffin embedding.

## 2.8. Statistical analysis

Statistical significance was determined by unpaired two-tailed T-test or one-way ANOVA with Mann–Whitney  $U$  or Kruskal Wallis post-test, depending on normality of data points, using Prism (GraphPad Software).  $p < 0.05$  was considered significant. Sphere quantification data were presented as superplots, with different colours represented as different biological replicates and averages of each experimental replicate indicated as larger-sized points.

## **Chapter 3. Integrin $\beta 6$ expression drives myoepithelial-led invasion**

### 3.1. Introduction

Transcriptomic analysis across different cell compartments within the DCIS microenvironment reveals that compared to their normal counterparts, DCIS-associated myoepithelial cells show the most significant change (54, 69). For example, during DCIS, myoepithelial cells have been shown to become altered through their mechanical stretching, to drive upregulation of integrin  $\beta 6$ , where its expression correlates with high-risk DCIS and invasive recurrence (72). These data highlight the critical need to deepen our understanding of the DCIS-associated myoepithelium, in attempt to identify key regulatory mechanisms of DCIS progression in the breast.

While there is an abundance of cell lines representing different breast cancer subtypes, suitable models to study the early stages of breast cancer progression are scarce. In addition, most cell culture systems utilise 2D approaches that lack a physiologically relevant microenvironment. As such, there is a need to develop more complex 3D culture systems to interrogate the roles of the myoepithelium and ECM in a physiomimetic context. One existing model is the non-tumorigenic MCF10A cell line, which forms hollow, polarised spheroids similar to the breast duct morphology, when cultured in Matrigel (260, 261). However, a limitation of this model is the lack of a distinct luminal/myoepithelial bilayer structure. Instead, MCF10A cells adopt both epithelial and basal characteristics under certain culture conditions (262, 263), making luminal and myoepithelial-specific mechanisms difficult to dissect. Arguments against the use of Matrigel when modelling early-stage breast cancer have also been raised, due to their incorrect desmoplastic representation of the DCIS-surrounding breast tissue (243). While the main constituent of the breast stroma is collagen-I, Matrigel is instead comprised of myriad growth factors and BM components (264). Thus, there is a need for incorporating a more physiologically relevant matrix, such as collagen-I into new models of DCIS going forwards.

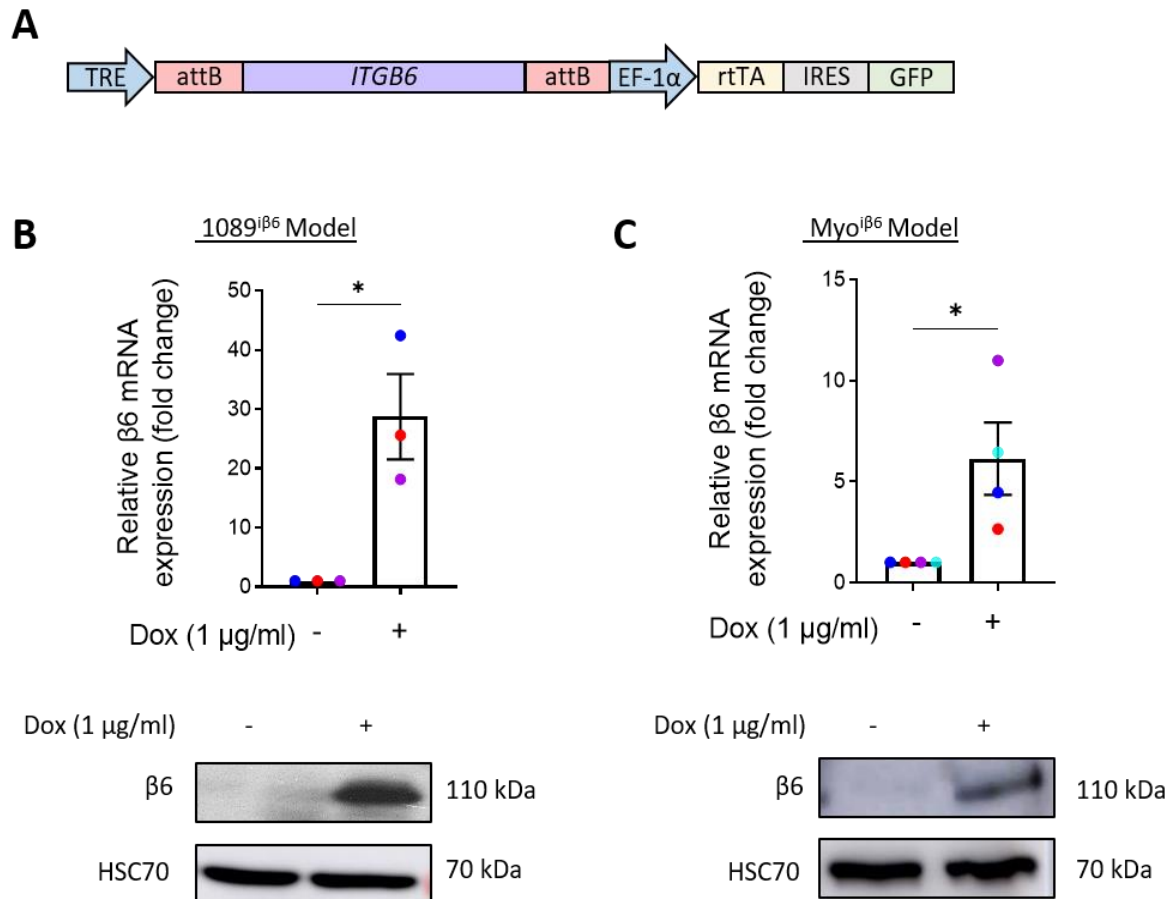
In this chapter, 3D *in vitro* models, incorporating both luminal and myoepithelial cells into physiologically relevant collagen-I gels are presented. Using these models, we further assess the role of myoepithelial integrin  $\beta 6$  in promoting DCIS progression.

## 3.2. Generation and characterisation of integrin $\beta 6$ -inducible DCIS models

### 3.2.1. Generation of integrin $\beta 6$ -inducible myoepithelial cells

Myoepithelial expression of integrin  $\beta 6$ , a marker of high-risk DCIS, correlates with disease progression and invasive disease (72). To investigate the biological consequences of myoepithelial integrin  $\beta 6$  expression *in vitro*, the *ITGB6* open reading frame was cloned into a doxycycline-inducible (pINDUCER21) lentiviral construct by Dr. E. Carter (Fig 3.1A) (249). We utilised two myoepithelial models: 1) the immortalised 1089 myoepithelial cell line and 2) primary, reduction mammoplasty-derived, myoepithelial cells. To generate integrin  $\beta 6$ -inducible cells, the *ITGB6* pINDUCER21 construct was integrated into the genome by lentiviral transduction, creating integrin  $\beta 6$ -inducible 1089 ( $1089^{\text{i}\beta 6}$ ) and primary cells ( $\text{Myo}^{\text{i}\beta 6}$ ). Sufficient overexpression of integrin  $\beta 6$  was validated by qPCR and western blot to assess integrin  $\beta 6$  expression at the RNA and protein level, respectively. For both the  $1089^{\text{i}\beta 6}$  and  $\text{Myo}^{\text{i}\beta 6}$  myoepithelial models, expression of integrin  $\beta 6$  was upregulated significantly upon 48 h treatment with 1  $\mu\text{g}/\text{ml}$  doxycycline (Fig 3.1B,C).



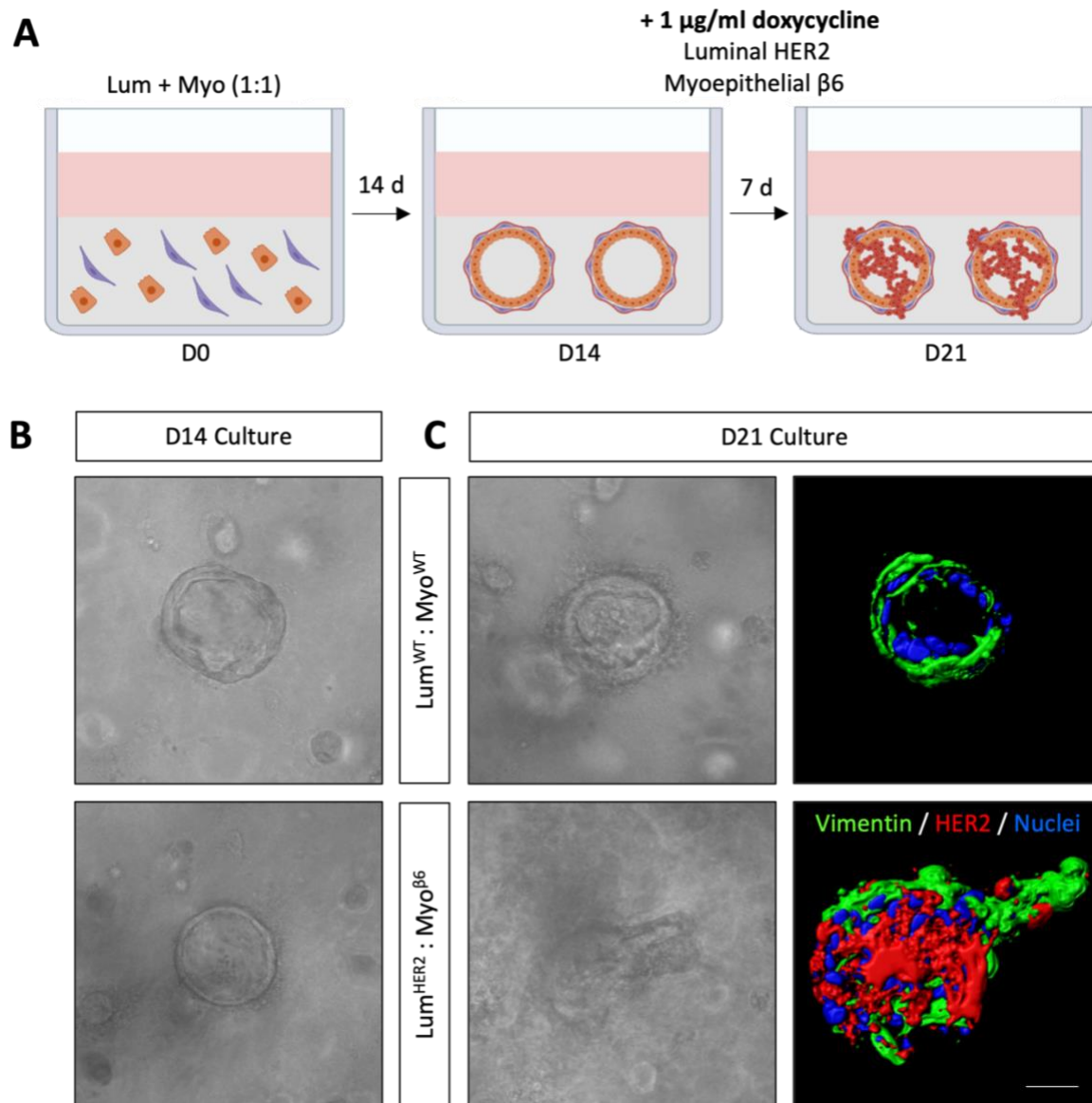


**Figure 3.1 Generation of integrin  $\beta 6$ -inducible cells.**

**(A)** Linear schematic of ITGB6 pINDUCER21 construct. The main components include a tetracycline response element (TRE), the ITGB6 sequence flanked by attB sites, an EF-1 $\alpha$  promoter, a reverse tetracycline controlled transactivator (rtTA), an internal ribosome entry site (IRES) and a sequence encoding green fluorescent protein (GFP). The ITGB6 pINDUCER21 plasmid was constructed by Dr. E. Carter. **(B,C)** Cells were transduced with the ITGB6 pINDUCER21 construct to generate integrin  $\beta 6$ -inducible 1089 (1089 <sup>$\beta 6$</sup> ) and primary cells (Myo <sup>$\beta 6$</sup> ). Cells were treated with 1  $\mu\text{g/ml}$  doxycycline for 48 h and harvested for RNA and protein extraction. Integrin  $\beta 6$  mRNA expression and protein expression in (B) 1089 <sup>$\beta 6$</sup>  and (C) Myo <sup>$\beta 6$</sup>  models. qPCR data are presented as mean  $\pm$  SEM with biological replicates indicated by different colours. For the Myo <sup>$\beta 6$</sup>  model, data presented is from three different donors. \* $p < 0.05$  (Two-tailed t-test).

### 3.2.2. Integrin $\beta 6$ expression drives invasion in primary ductal model of DCIS

Previous work by Dr. E. Carter has demonstrated that primary luminal and myoepithelial cells can be recombined in collagen-I gel cultures to form native ductal structures with a correctly orientated bilayer (245). For this model, primary luminal and myoepithelial cells were combined in a 1:1 ratio and embedded in 2 mg/ml collagen-I gels (Fig 3.2A). Ductal structures formed at approximately 14 days post-embedding (Fig 3.2B) after which cultures were treated with 1  $\mu$ g/ml doxycycline for 7 days to induce expression of virally transduced constructs encoding HER2 and integrin  $\beta 6$  in the luminal and myoepithelial cells, respectively. Induction of HER2 in the luminal compartment of this model, once ductal structures had formed, led to filling of the ducts, reminiscent of DCIS. Strikingly, concomitant expression of integrin  $\beta 6$  in the myoepithelial compartment promoted the formation of invasive projections, with vimentin-positive myoepithelial cells leading and HER2-positive luminal cells following (Fig 3.2C).

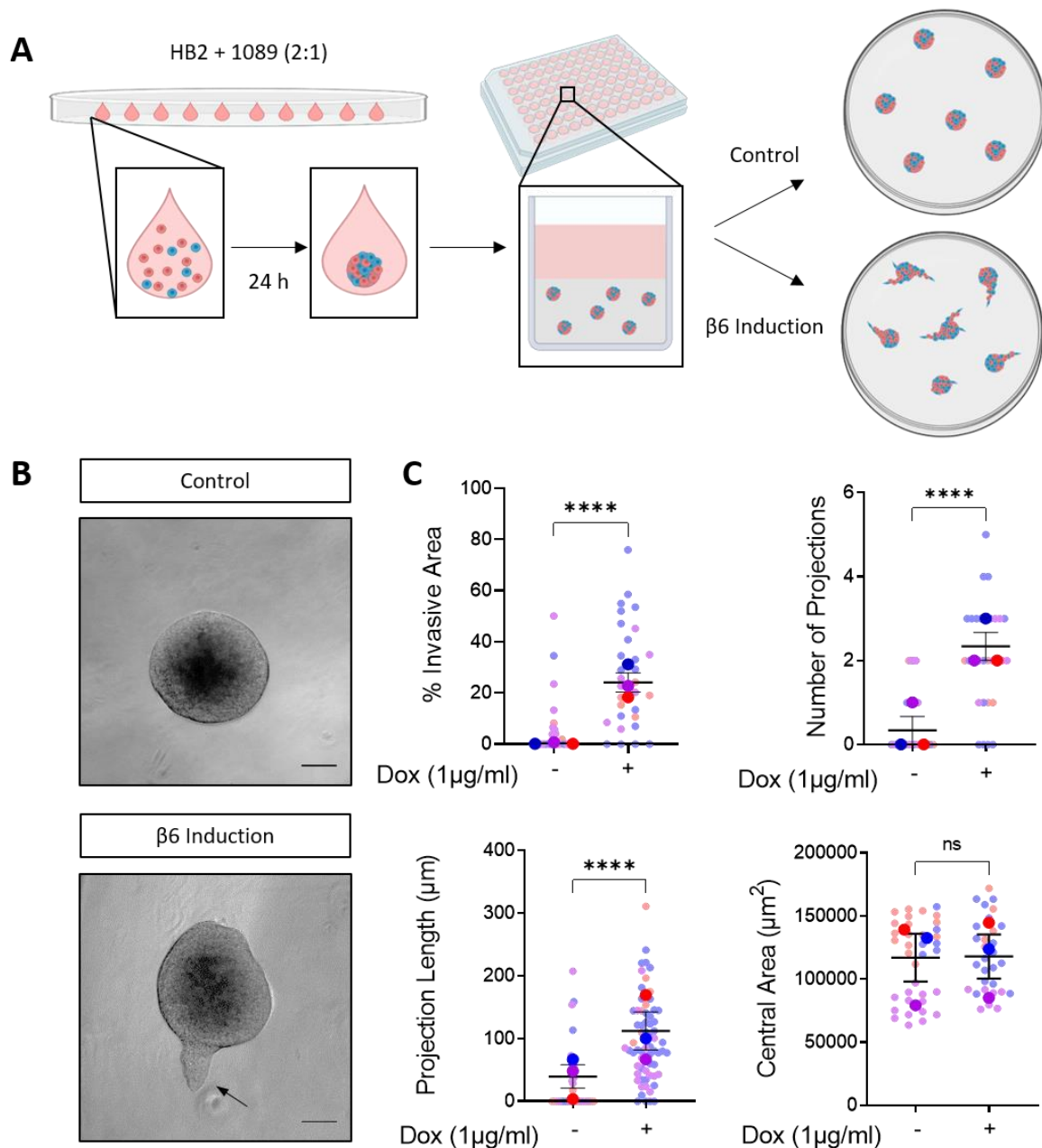


**Figure 3.2 Integrin  $\beta 6$  expression drives invasion in primary ductal model of DCIS.**

**(A)** Schematic of primary ductal model. Primary luminal and myoepithelial cells were combined in a 1:1 ratio and resuspended in a 2 mg/ml collagen-I gel. Upon formation of ductal structures 14 d post embedding, cultures were treated with 1  $\mu\text{g/ml}$  doxycycline for 7 d to induce expression of luminal HER2 and myoepithelial integrin  $\beta 6$ . **(B)** Representative light micrographs of cultures at 14 d. **(C)** Representative light micrographs and Imaris projections of ductal structures after 21 d of culture with vimentin; green, HER2; red, DAPI; blue. Scale bar = 20  $\mu\text{m}$ . The model developed was developed by Dr. E. Carter. For this experiment, S. Gibson set up the cultures, while Dr. E. Carter performed the immunofluorescent imaging and Imaris reconstruction.

### 3.2.3. Myoepithelial integrin $\beta 6$ expression drives invasion in cell line sphere model

While the ductal model is a faithful reconstruction of the ductal architecture in an *in vitro* environment, the model is difficult to quantify and requires lengthy culture time that limits mechanistic studies. As an alternative, we developed two short term 3D co-culture systems, combining either 1089 <sup>$\beta 6$</sup>  or Myo <sup>$\beta 6$</sup>  cells with the non-tumorigenic HB2 cell line (244) into collagen-I gels. For the first model, an HB2/1089 <sup>$\beta 6$</sup>  cell line model was constructed, where 1089 <sup>$\beta 6$</sup>  cells were combined with HB2 cells in hanging drops to form heterocellular spheroids, which were then embedded into collagen-I gels (Fig 3.3A). To induce myoepithelial integrin  $\beta 6$  expression, spheroids were then treated with 1  $\mu\text{g/ml}$  doxycycline. Upon overexpression in myoepithelial cells, we observed striking formation of invasive projections within 4 days post integrin  $\beta 6$ -induction (Fig 3.3B). To quantify the extent of invasion, four measurements were taken: 1) the percentage of invasive area, 2) the number of projections, 3) the projection length, and 4) the central spheroid area. Sphere quantification and morphometric analysis revealed a significant increase in invasion, with integrin  $\beta 6$ -induced spheroids displaying an invasive area of approximately 20%. The number of projections and projection length were also increased to an average of approximately 2 versus 0 ( $\beta 6$  vs. Control) and 100 versus 50  $\mu\text{m}$  ( $\beta 6$  vs Control), respectively (Fig 3.3C). For the central area, no significant change was observed between conditions, indicating that  $\beta 6$  expression had no effect on cell growth and proliferation. Taken together, these data implicate myoepithelial integrin  $\beta 6$  expression as a driver of invasion (Fig 3.3C).

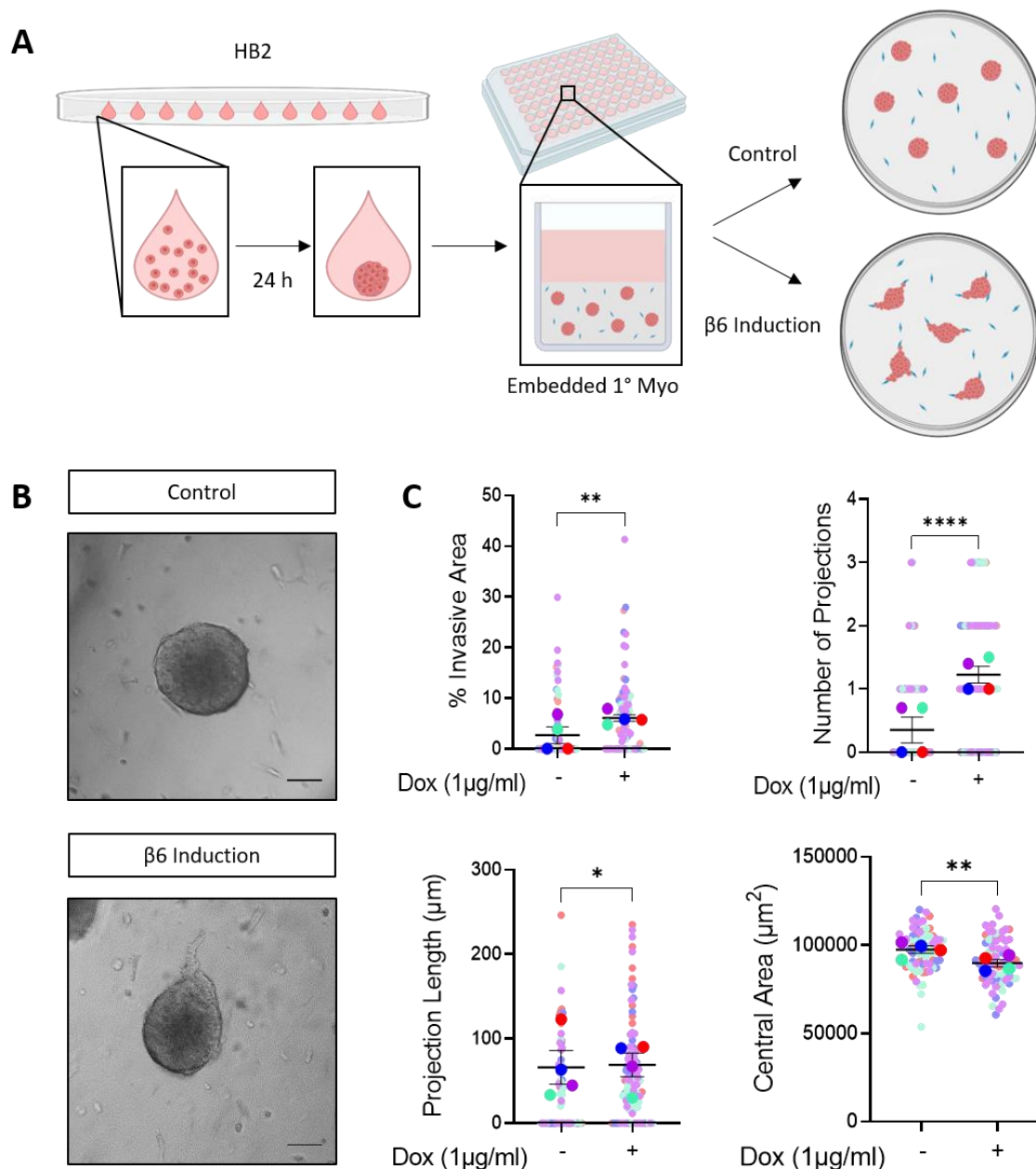


**Figure 3.3 Myoepithelial integrin  $\beta 6$  expression drives invasion in HB2/1089<sup>iB6</sup> model.**

**(A)** Schematic summarising HB2/1089 spheroid cell line model. HB2 luminal and 1089<sup>iB6</sup> cells were combined in methylcellulose hanging drops to form heterocellular spheroids overnight. The follow day, spheroids were collected and embedded in 4 mg/ml collagen gels. Spheres were treated with 1 µg/ml of doxycycline to induce myoepithelial integrin  $\beta 6$  expression. **(B)** Representative brightfield images at 4 d post treatment. **(C)** Summary graphs showing % invasive area, number of projections, projection length (µm) and central area (µm<sup>2</sup>). Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. ns = not significant, \*\*\*\*p<0.0001 (Mann Whitney U Test). Scale bar = 100 µm

### 3.2.4. Myoepithelial integrin $\beta 6$ expression drives invasion in primary sphere model

To validate the invasive role of myoepithelial integrin  $\beta 6$  expression in the context of primary cells, a second co-culture model was developed. In this model, Myo<sup>i $\beta 6$</sup>  cells were also combined with the HB2 luminal cells. However, as primary Myo<sup>i $\beta 6$</sup>  cells collected in the centre of the sphere when combined with HB2 cells in methylcellulose hanging drops, and were then trapped, this model comprised of HB2 monoculture spheroids embedded in collagen gels containing primary Myo<sup>i $\beta 6$</sup>  cells instead (Fig 3.4A). Upon induction of myoepithelial integrin  $\beta 6$  expression in the Myo<sup>i $\beta 6$</sup>  cells for 3 days, we observed a similar increase in invasion, characterised by the presence of protrusions out of the core spheroid (Fig 3.4B). Congruent with the cell line model, the percentage of invasive area, the number of projections and the projection length showed a significant increase with integrin  $\beta 6$ -induction (Fig 3.4C), demonstrating the role of myoepithelial integrin  $\beta 6$  in promoting the invasive phenotype across both cell line and primary models.



**(A)** Schematic summarising HB2/1°Myo spheroid cell line model. HB2 luminal cells were resuspended in methylcellulose hanging drops to form monoculture spheroids overnight. The following day, spheroids were collected and embedded into 4 mg/ml Myo<sup>iβ6</sup>-containing collagen gels. Spheroids were treated with 1 μg/ml of doxycycline to induce myoepithelial integrin β6 expression. **(B)** Representative brightfield images at 3 d post doxycycline treatment. **(C)** Summary graphs showing % invasive area, number of projections, projection length (μm) and central area (μm<sup>2</sup>). Data are presented as mean ± SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. ns = not significant, \*p<0.05, \*\*p<0.01, \*\*\*\*p<0.0001 (Mann Whitney U Test). Scale bar = 100 μm. For this experiment, Dr. E. Carter and S. Gibson performed two biological replicates each, while S. Gibson performed quantification analysis across all biological replicates.

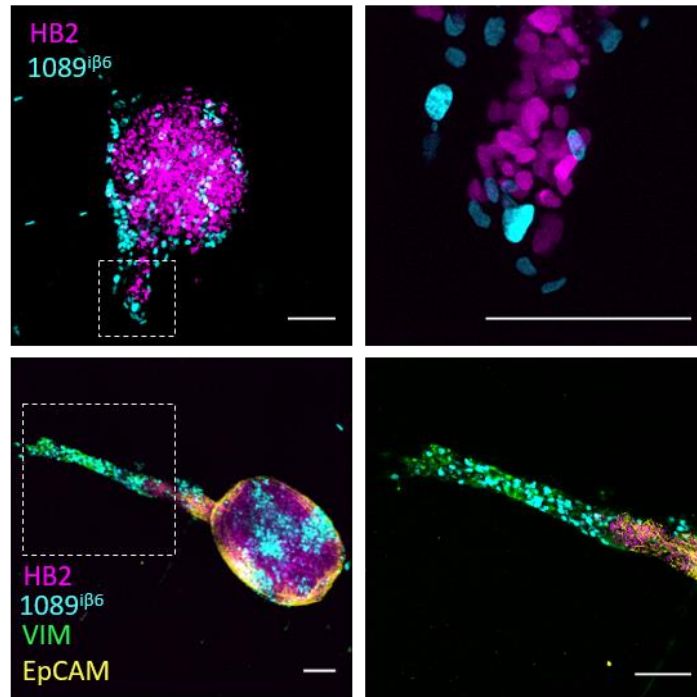
### 3.2.5. HB2 invasion is led by myoepithelial cells

Collective invasion is a process whereby cells invade as a multicellular unit (265-267). For cancer cells to invade while maintaining their epithelial phenotype, they can rely upon the motile cells in the microenvironment to facilitate their invasion. This is typically observed in solid malignancies, where cancer cells rely on migratory stromal cells such as CAFs to create tracks and physically pull cancer cells into their surrounding matrix (265, 266).

To distinguish between luminal and myoepithelial cell types within the invading projections of integrin  $\beta 6$ -expressing spheroids, cells were transduced with constructs encoding fluorescently tagged histone proteins, prior to the recombination into spheres. Confocal analysis of fluorescently labelled cells demonstrated that, while both myoepithelial and luminal compartments were present in the invasive projections, the invasion was myoepithelial-led, with HB2 luminal cells following (Fig 3.5, 3.6). These findings suggest that, in this context, myoepithelial cells act as the 'leader' cells, whereas luminal cells act as 'follower' cells.

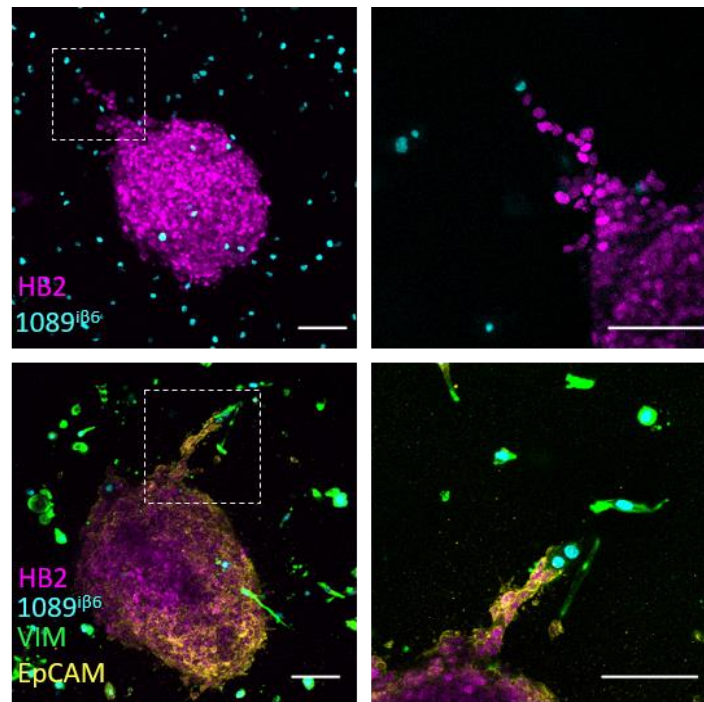
Characterisation of invading cells by immunofluorescent analysis revealed that invading myoepithelial cells were vimentin positive, consistent with existing literature (268). Interestingly, luminal cells maintained expression of the epithelial specific marker EpCAM (Fig 3.5, 3.6), demonstrating their dependence on the myoepithelial compartment, and maintenance of their epithelial phenotype observed in collective tumour invasion (105, 265, 267, 269, 270).





**Figure 3.5** HB2 invasion is led by myoepithelial cells in HB2/1089<sup>iβ6</sup> model.

Cells were transduced with lentivirus encoding fluorescently-tagged histone proteins. Spheres comprised of fluorescently labelled HB2/1089<sup>iβ6</sup> spheroids were embedded in collagen gels and treated with 1 µg/ml doxycycline to induce myoepithelial integrin β6 expression. Representative fluorescence images of invading spheroids at 4 d with HB2 (magenta) 1089<sup>iβ6</sup> (cyan), vimentin (green) and EpCAM (yellow). Scale bar = 100 µm.



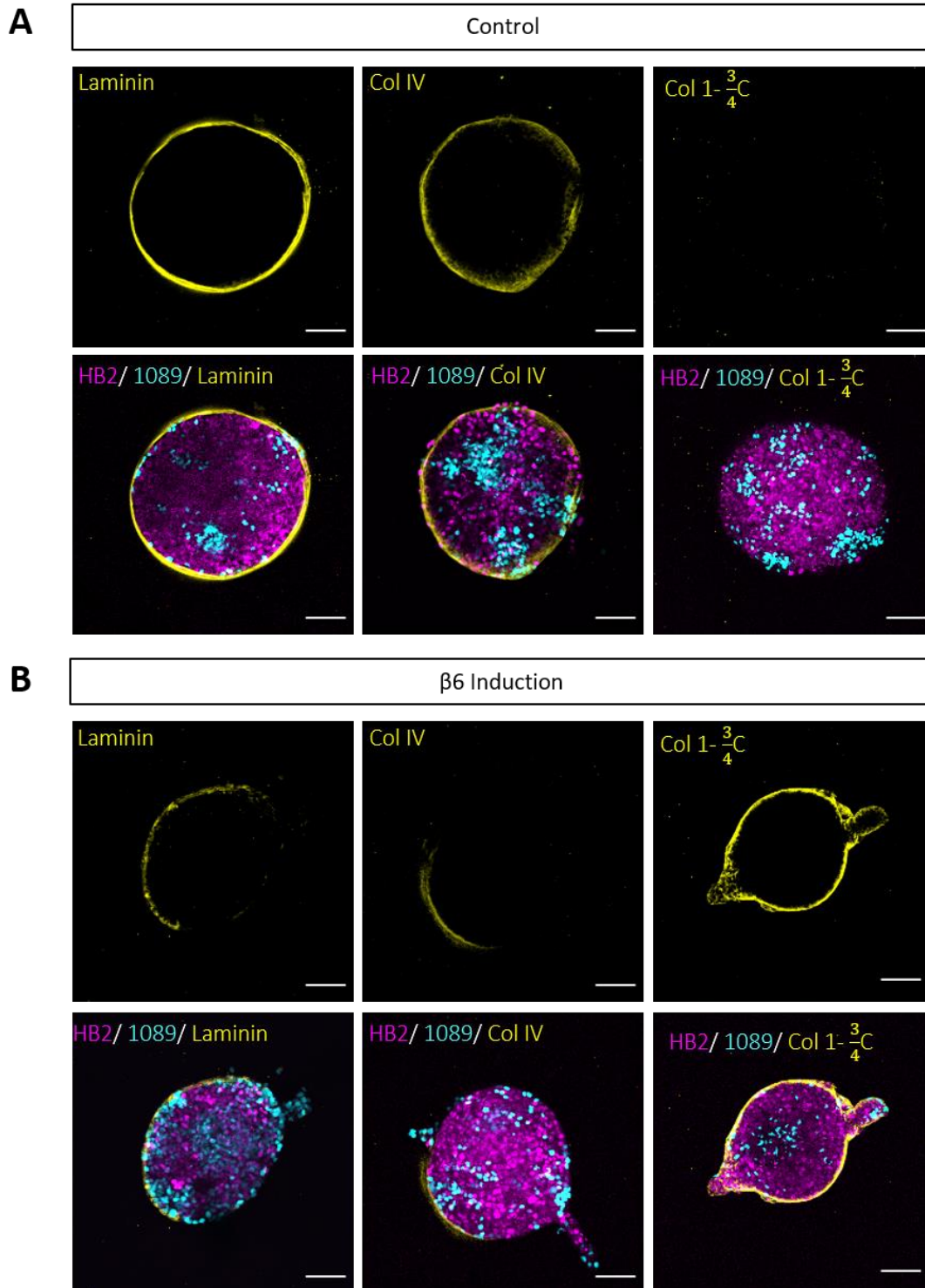
**Figure 3.6** HB2 invasion is led by myoepithelial cells in HB2/Myo<sup>iβ6</sup> model.

Cells were transduced with lentivirus encoding fluorescently-tagged histone proteins. HB2 monoculture spheroids were embedded in collagen gels containing Myo<sup>iβ6</sup> cells and treated with 1 µg/ml doxycycline to induce myoepithelial integrin β6 expression. Representative fluorescence images of invading spheroids at 4 d with HB2 (magenta), Myo<sup>iβ6</sup> (cyan), vimentin (green) and EpCAM (yellow). Scale bar = 100 µm.

### **3.3. Integrin $\beta 6$ expression drives invasion through a protease-dependent mechanism**

#### **3.3.1. Invading spheroids remodel their basement membrane upon integrin $\beta 6$ expression**

To determine whether spheroids deposited their own BM and if this was perturbed upon myoepithelial integrin  $\beta 6$  expression, confocal imaging was used to visualise the major BM proteins, laminin and collagen-IV. Due to the nature of the HB2/Myo<sup>i $\beta 6$</sup>  model, with primary myoepithelial cells embedded in the surrounding collagen gel, BM analysis was performed only on the HB2/1089<sup>i $\beta 6$</sup>  model, where matrix deposition from the heterocellular sphere could be visualised. Reflective of an intact and deposited BM, continuous expression of both laminin and collagen-IV was observed in equatorial sections of non-invading control spheroids (Fig 3.7A). Despite minimal changes to matrisome expression upon integrin  $\beta 6$ -induction (Appendix 1.1), invading integrin  $\beta 6$ -induced spheroids displayed a loss in their BM, resembling the loss typically seen in the transition from DCIS to IDC (Fig 3.7B). Moreover, a striking increase in cleaved collagen-I was detected in the gels of integrin  $\beta 6$ -induced spheroids compared to non-invasive controls, suggesting that while control spheres undergo minimal matrix proteolysis, this is significantly enhanced upon expression of integrin  $\beta 6$  in the myoepithelial compartment (Fig 3.7B).



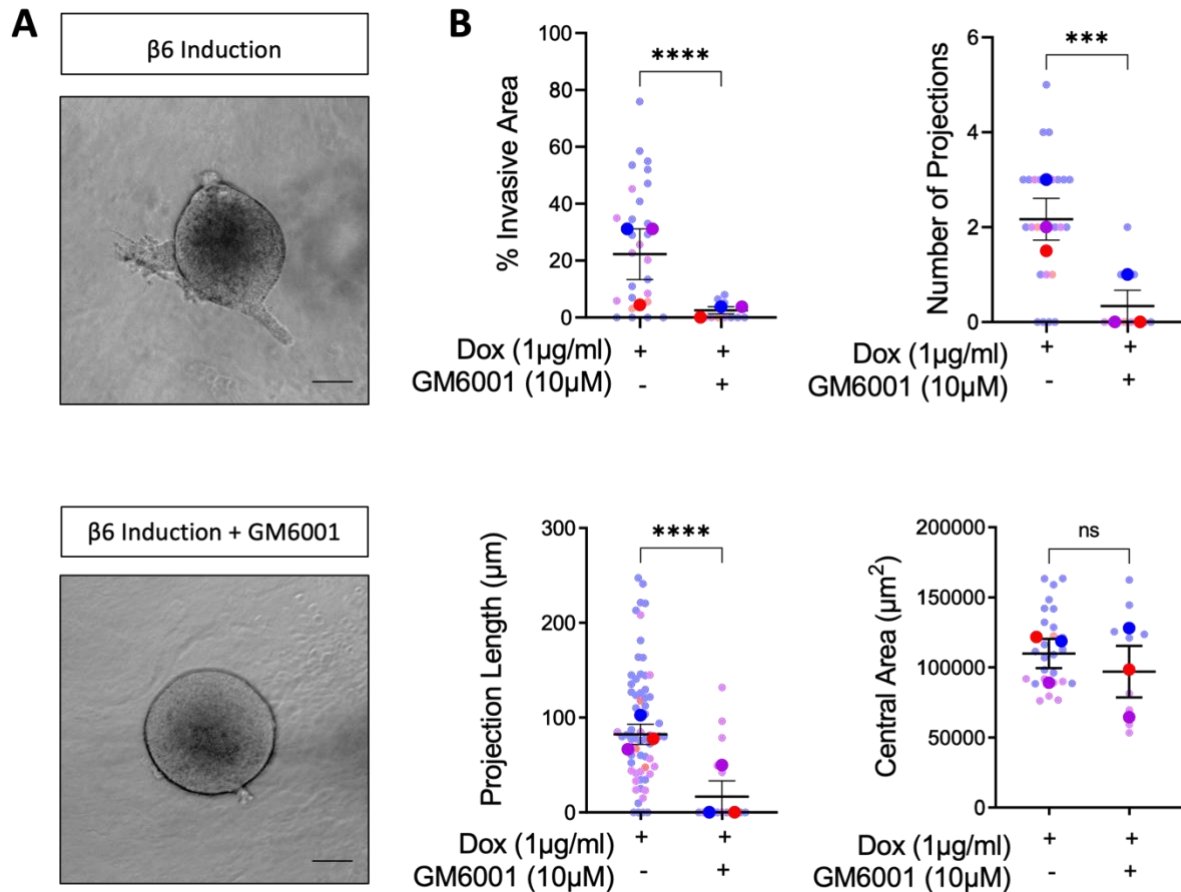
**Figure 3.7** BM remodelling is associated with invasion in the HB2/1089<sup>β6</sup> model.

Spheres comprised of fluorescently labelled HB2/1089<sup>β6</sup> spheroids were embedded in collagen gels and treated with 1 μg/ml doxycycline to induce myoepithelial integrin β6 expression. Representative fluorescence images of **(A)** control and **(B)** β6-induced spheroids at 4 d post treatment with HB2 (magenta), 1089<sup>β6</sup> (cyan) and laminin, collagen-IV or cleaved collagen-I (yellow). Scale bar = 100 μm.

### 3.3.2. Integrin $\beta 6$ -dependent invasion requires metzincin activity

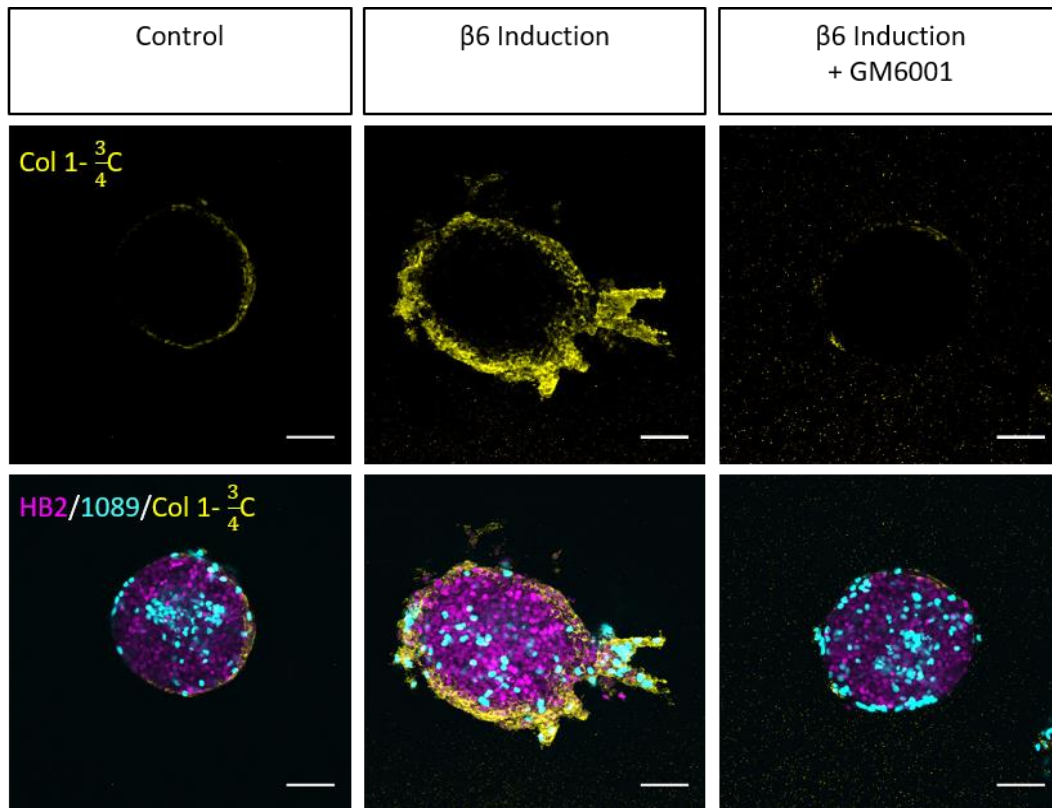
Stromal-led invasion can be protease dependent or independent, depending on cell behaviour and matrix composition (265, 271, 272). We next hypothesised that, based on the expected pore size of the rat-tail derived collagen-I, as well as the increase in cleaved collagen-I staining, the observed integrin  $\beta 6$ -induced myoepithelial-led invasion was protease dependent (273). To confirm this, spheroids were treated with a broad-spectrum metzincin inhibitor, GM6001 (10  $\mu$ M), which targets several MMPs, ADAMs and ADAMTS zinc endopeptidases. Upon GM6001 treatment of  $\beta 6$ -induced spheres, the percentage of invasive area, number of projections and projection length were significantly reduced, demonstrating that inhibition of metzincins is sufficient to block integrin  $\beta 6$ -dependent invasion in this model (Fig 3.8A,B). Indeed, analysis of collagen-I cleavage using an antibody which recognises collagen-I in its cleaved form, showed a loss of integrin  $\beta 6$ -induced proteolysis following GM6001 treatment (Fig 3.9). Changes to the central area of GM6001-treated HB2/1089<sup>i $\beta 6$</sup>  spheres were non-significant, indicating that reduced invasion was not due to drug toxicity and reduced cell growth (Fig 3.8B).

Reduced invasion following GM6001 treatment was also recapitulated in the Myo<sup>i $\beta 6$</sup>  model, with a significant decrease in the percentage of invasive area, number of projections and projection length, highlighting the requirement for a protease-mediated mechanism in facilitating invasion across both cell line and primary myoepithelial models (Fig 3.10A,B). Interestingly, a significant reduction in the central area was observed in this model, suggesting that while protease activity may not be required for spheroid growth in the cell line spheroid model, it may be required in the primary spheroid model (Fig 3.10B).



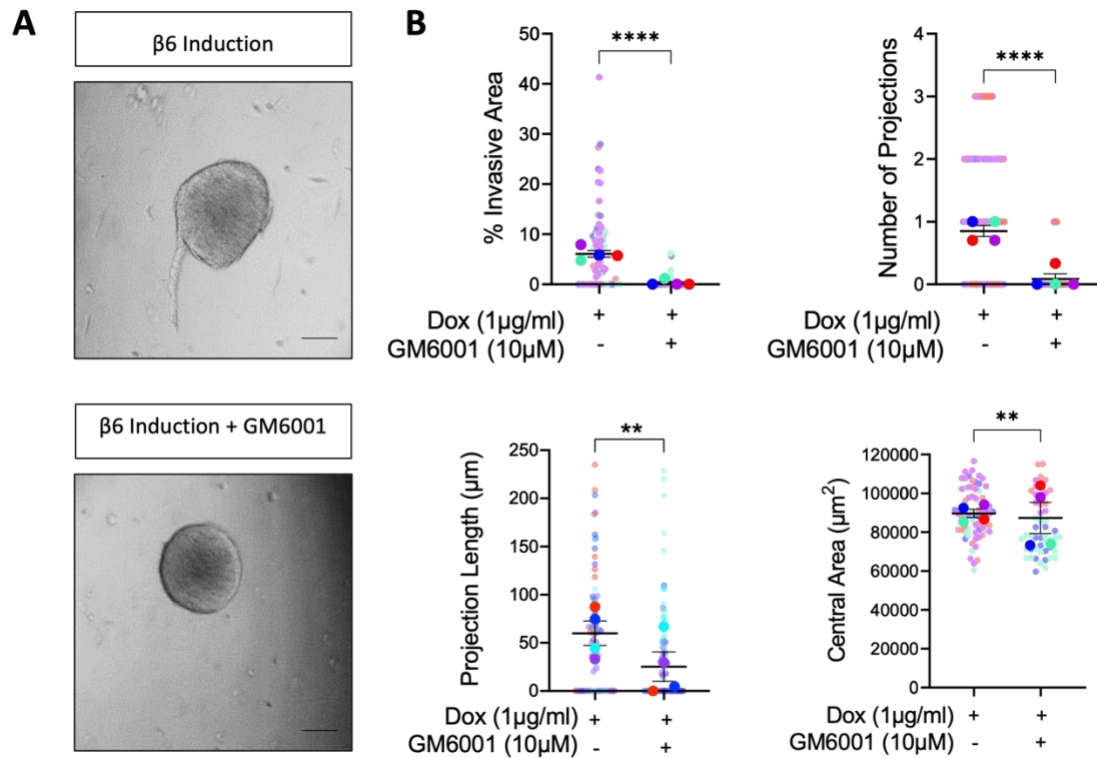
**Figure 3.8 Integrin  $\beta 6$ -driven invasion is protease-dependent in the HB2/1089 <sup>$\beta 6$</sup>  model.**

HB2 luminal and 1089<sup>ib6</sup> cells were combined in a methylcellulose cell suspension to form hanging drops. Overnight, cells formed spheres, which were embedded into a 4 mg/ml collagen gels the following day. **(A)** Representative light micrographs of spheres at 4 d, treated with 1  $\mu$ g/ml of doxycycline to induce myoepithelial integrin  $\beta 6$  expression and 10  $\mu$ M GM6001 to inhibit metzincin activity. **(B)** Summary graphs showing % invasive area, number of projections, projection length ( $\mu$ m) and central area ( $\mu$ m<sup>2</sup>). Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. ns = not significant, \*\*\*p<0.001, \*\*\*\*p<0.0001 (Mann Whitney U Test). Scale bar = 100  $\mu$ m.



**Figure 3.9** Protease activity is required for collagen-I cleavage in HB2/1089<sup>iB6</sup> model.

HB2 luminal and 1089<sup>iB6</sup> cells were combined in a methylcellulose cell suspension to form hanging drops. Overnight, cells formed spheres, which were embedded into 4 mg/ml collagen gels the following day. Spheres were treated with 1  $\mu$ g/ml of doxycycline to induce myoepithelial integrin  $\beta 6$  expression and 10  $\mu$ M GM6001 to inhibit metzincin activity. Representative fluorescence images of spheroids at 4 d post treatment with HB2 (magenta), 1089<sup>iB6</sup> (cyan) and cleaved collagen-I (yellow). Scale bar = 100  $\mu$ m.



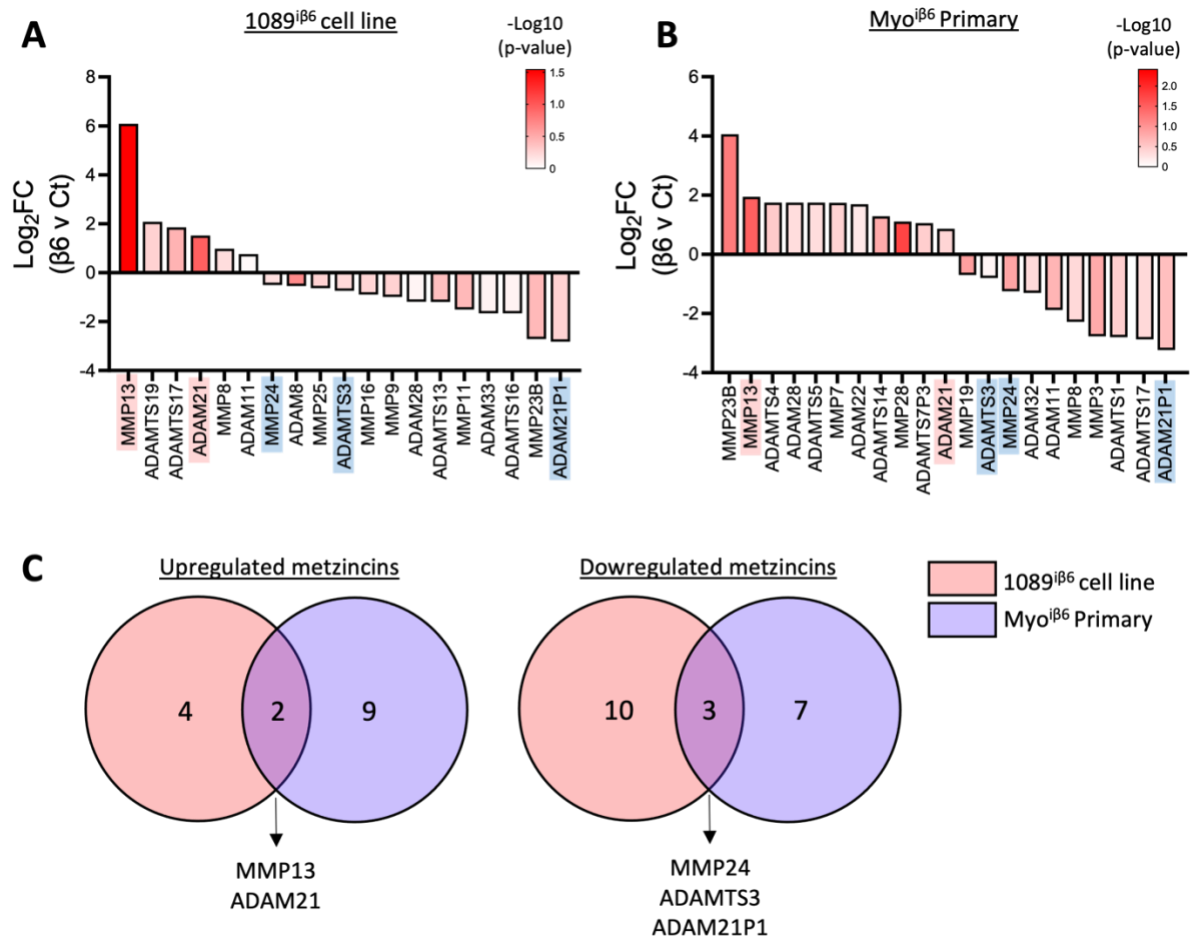
**Figure 3.10 Integrin  $\beta 6$ -driven invasion is protease-dependent in the HB2/Myo <sup>$\beta 6$</sup>  model.**

HB2 monoculture spheroids were embedded in 4 mg/ml Myo <sup>$\beta 6$</sup> -containing collagen gels. **(A)** Representative light micrographs of spheres at 3 d, treated with 1  $\mu$ g/ml of doxycycline to induce myoepithelial integrin  $\beta 6$  expression and 10  $\mu$ M GM6001 to inhibit metzincin activity. **(B)** Summary graphs showing % invasive area, number of projections, projection length ( $\mu$ m) and central area ( $\mu$ m<sup>2</sup>). Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$  (Mann Whitney U Test). Scale bar = 100  $\mu$ m. For this experiment, Dr. E. Carter and S. Gibson performed two biological replicates each, while S. Gibson performed quantification analysis across all biological replicates.



### 3.3.3. Integrin $\beta 6$ alters metzincin expression

In order to understand the effects of integrin  $\beta 6$  induction on metzincin expression, we examined changes in the transcriptional levels of metzincins in both 1089<sup>i $\beta 6$</sup>  and Myo<sup>i $\beta 6$</sup>  cells grown in 2D. To achieve this, RNA was extracted from myoepithelial cells 48 h after treatment with doxycycline treatment and subjected to RNA sequencing (RNAseq). The differential expression analysis of RNAseq data revealed notable changes in several metzincins upon integrin  $\beta 6$  induction (Fig 3.11A,B). Interestingly, differences were observed between the two models, most likely due to the genotypic and phenotypic variation in cells lines upon serial passage and genetic drift. Differences were also observed between independent patient donors, highlighting the heterogeneity across patient samples. Despite this, consistent upregulation in MMP13 and ADAM21, and a downregulation in MMP24, ADAMTS3 and ADAM21P1 were observed across both models (Fig 3.11C). These data indicate that myoepithelial cell upregulation of integrin  $\beta 6$ , leads to alterations in metzincin expression, which may facilitate invasive progression.



**Figure 3.11** Overexpression of integrin  $\beta 6$  in myoepithelial cells alters metzincin expression.

Myoepithelial cells grown in 2D were treated with 1  $\mu\text{g}/\text{ml}$  doxycycline for 48 h to induce myoepithelial integrin  $\beta 6$  expression prior to extraction of RNA. RNAseq differential metzincin expression analysis from of  $\beta 6$ - versus control (Ct) in **(A)** 1089 <sup>$\beta 6$</sup>  and **(B)** Myo <sup>$\beta 6$</sup>  cells from two primary donors. **(C)** Venn diagrams showing overlapping similarities in metzincin expression changes between 1089 <sup>$\beta 6$</sup>  (red) and Myo <sup>$\beta 6$</sup>  (purple) models.

### 3.4. Discussion

#### Results summary

- Upregulation of myoepithelial integrin  $\beta 6$ , a marker of high-risk DCIS, drives invasion of luminal cells into surrounding matrix.
- Heterocellular spheroids lay down their own physiologically relevant basement membrane, which is remodelled during invasion.
- Myoepithelial-led invasion in 3D is dependent on protease activity.
- Integrin  $\beta 6$  upregulation alters metzincin expression in myoepithelial cells.

The expression of epithelial integrin  $\beta 6$  has previously been implicated as a driver of progression in cancer, where its expression is associated with an aggressive tumour phenotype and reduced survival (194, 208, 274-276). However, most studies have focused on the upregulation of integrin  $\beta 6$  in cancer cells rather than its expression in other epithelial-derived cells of the tumour microenvironment, such as myoepithelial cells.

In this study, we present three models of DCIS progression that incorporate integrin  $\beta 6$ -inducible myoepithelial cells, along with a luminal cell compartment. These models enable the dissection of myoepithelial mechanisms underlying early-stage breast cancer progression. By utilising 3D models of the tumour microenvironment, we can investigate cellular crosstalk and more closely mimic physiological behaviours, offering a valuable tool for researchers (277). Our heterocellular model systems provides advantage over others, such as the MCF10A cell line model, which tends to exhibit a mixture of luminal and basal characteristics in during spheroid formation (262, 263). Importantly,, our models allow for the manipulation of distinct myoepithelial and luminal compartments, while

maintaining their cellular phenotypes, thus recapitulating the behaviours seen in human breast tumours. Furthermore, we demonstrate that spheroids deposit their own physiologically relevant BM in our models, which they subsequently remodel and degrade during cellular invasion.

We show that upregulation of integrin  $\beta 6$ , a high-risk marker for DCIS, drives myoepithelial-led invasion of luminal cells in both cell line and primary spheroid models. Notably, luminal cells are unable to invade without the presence of a myoepithelial conduit (Appendix 1.2), suggesting that myoepithelial cells expressing integrin  $\beta 6$  may acquire characteristics similar to CAFs. This phenomenon is commonly observed during stromal-led invasion, where cancer cells maintain their epithelial phenotype and exploit non-malignant cells in the microenvironment to facilitate invasion (107, 267). The role of myoepithelial cells as tumour-promoters is supported by spatial analysis of DCIS, which identifies the myoepithelium as one of the most influential features predictive of invasive recurrence (58). Interestingly, in a murine model of breast cancer invasion, where myoepithelial cells are less abundant, leading cells have been shown to adopt the basal CK14+ characteristics of myoepithelial cells (270). However, in our model, which incorporates a myoepithelial cell compartment, this provides evidence that cancer cells can rely on resident myoepithelial cells and maintain their luminal state during invasion.

Finally, we implicate integrin  $\beta 6$ -driven invasion to be protease-dependent, with invasion abrogated upon pharmacological inhibition of metzincins. Validation of key proteases, as well as novel substrates, which may mediate myoepithelial-led invasion, will be key in furthering our understanding of DCIS progression.

## **Chapter 4. MMP13 is required for integrin $\beta$ 6-dependent myoepithelial-led invasion**

## 4.1. Introduction

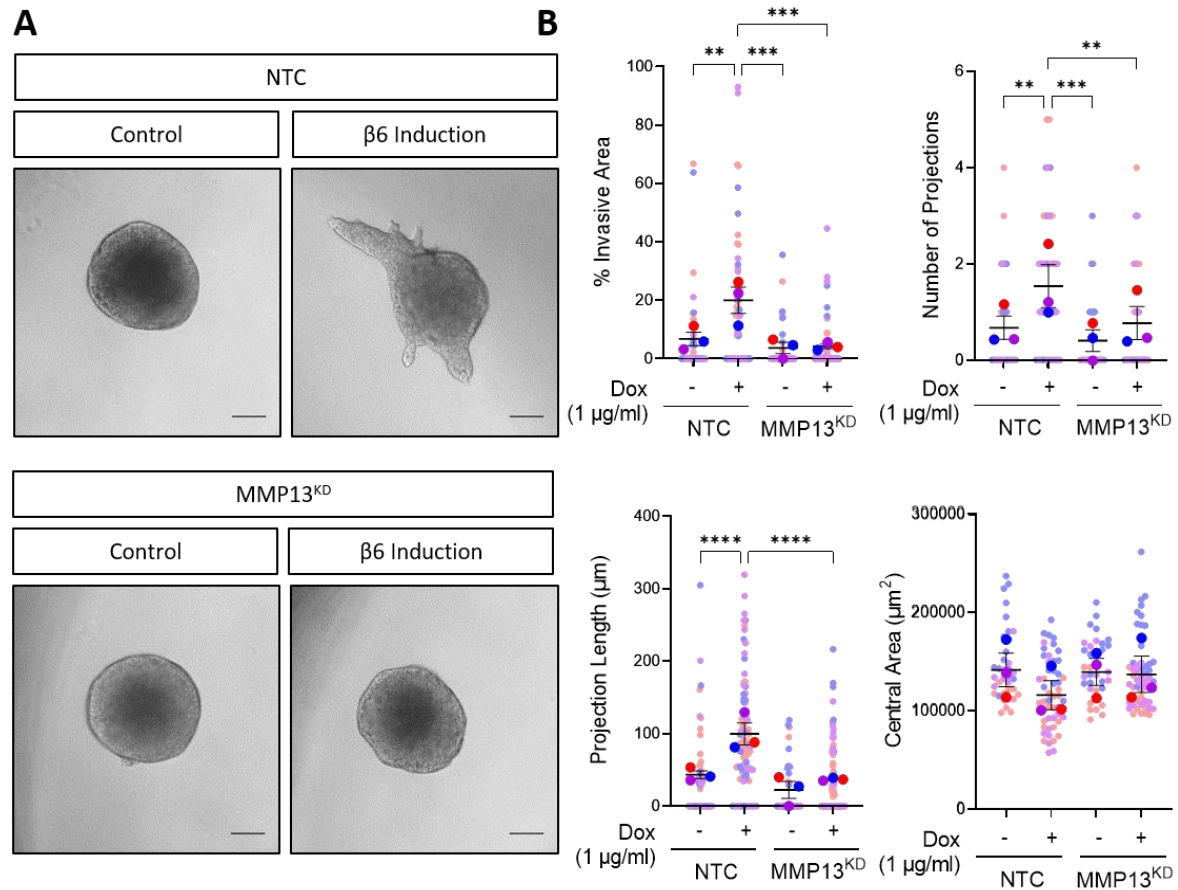
During the progression of DCIS to IDC, degradation of the BM is required to enable the invasion of cancer cells into the surrounding stroma. This invasion is typically associated with increased protease expression and subsequent cleavage of dense matrices to create paths for migration. Integrins are key players in this process and can regulate protease expression, as well as activate pro-MMPs through physical interaction. For example, in cancer cells, integrin  $\alpha\beta6$  has been shown to promote MMP2 and MMP9 expression (195, 196, 270), while  $\alpha\beta3$  activates MMP2 (197, 198, 278).

Previously, we showed that upregulation of integrin  $\beta6$  drives myoepithelial-led invasion of breast luminal cells into a collagen matrix. As integrin  $\beta6$  expression was coupled with alterations in metzincin expression, we hypothesise that these changes may present as a fundamental step in promoting DCIS progression to IDC. In this chapter, we investigate the role of MMP13 in facilitating  $\beta6$ -induced myoepithelial-led invasion using our previously established models of DCIS progression. We dissect the mechanism whereby MMP13 is transcriptionally regulated and validate MMP13 as a potential driver of invasion *in vivo*.

## 4.2. Integrin $\beta 6$ -driven invasion is MMP13 dependent

### 4.2.1. Knock-down of myoepithelial MMP13 reduces invasion

As the most significant change in protease expression upon integrin  $\beta 6$  induction was an upregulation in MMP13 across both 1089<sup>i $\beta 6$</sup>  and Myo<sup>i $\beta 6$</sup>  models, we hypothesised that MMP13 could be responsible for driving the invasive phenotype. To address this, we next examined whether loss of myoepithelial MMP13 was sufficient to reduce invasion. This was achieved by transfecting 1089<sup>i $\beta 6$</sup>  cells with non-targeting control or MMP13-targeting siRNA (MMP13<sup>KD</sup>) prior to their combination with HB2 luminal cells in heterocellular spheroids (Fig 4.1A). Effective knockdown of myoepithelial MMP13 was confirmed by qPCR prior to embedding spheres into collagen gels (Appendix 2.1). Strikingly, knockdown of myoepithelial MMP13 in integrin  $\beta 6$ -induced spheroids significantly diminished the percentage of invasive area to non-integrin  $\beta 6$ -induced levels. In addition, the number of projections and the projection length were reduced significantly compared to non-targeting control conditions (Fig 4.1B). Taken together, these data indicate that  $\beta 6$ -induced invasion occurs through an MMP13-dependent mechanism.



**Figure 4.1 Integrin β6-driven invasion is MMP13 dependent in the HB2/1089<sup>iβ6</sup> model.**

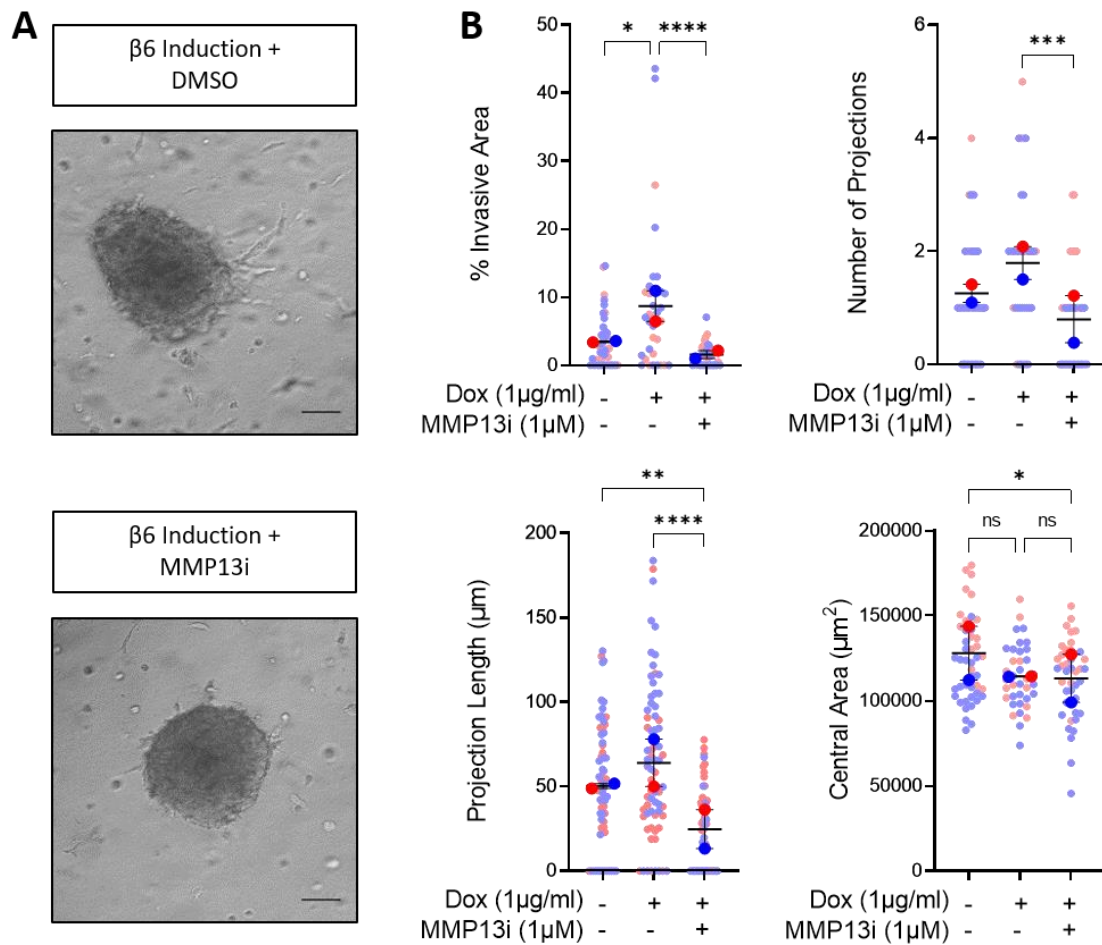
Prior to co-culture, 1089<sup>iβ6</sup> cells were transiently transfected with non-targeting control (NTC) or MMP13 siRNA (MMP13<sup>KD</sup>). Transfected 1089<sup>iβ6</sup> cells were then combined with HB2 cells to form spheres, which were embedded in collagen gels and treated with 1 μg/ml doxycycline to induce integrin β6 expression. **(A)** Representative light micrographs of spheres at 4 d, with either myoepithelial-NTC or -MMP13<sup>KD</sup> and myoepithelial integrin β6 expression. **(B)** Summary graphs showing % invasive area, number of projections, projection length (μm) and central area (μm<sup>2</sup>). Data are presented as mean ± SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \*\*p<0.01, \*\*\*p<0.001, \*\*\*\*p<0.0001 (Kruskal-Wallis test with multiple comparisons). Scale bar = 100 μm.



#### 4.2.2. Treatment with MMP13i reduces $\beta$ 6-driven invasion

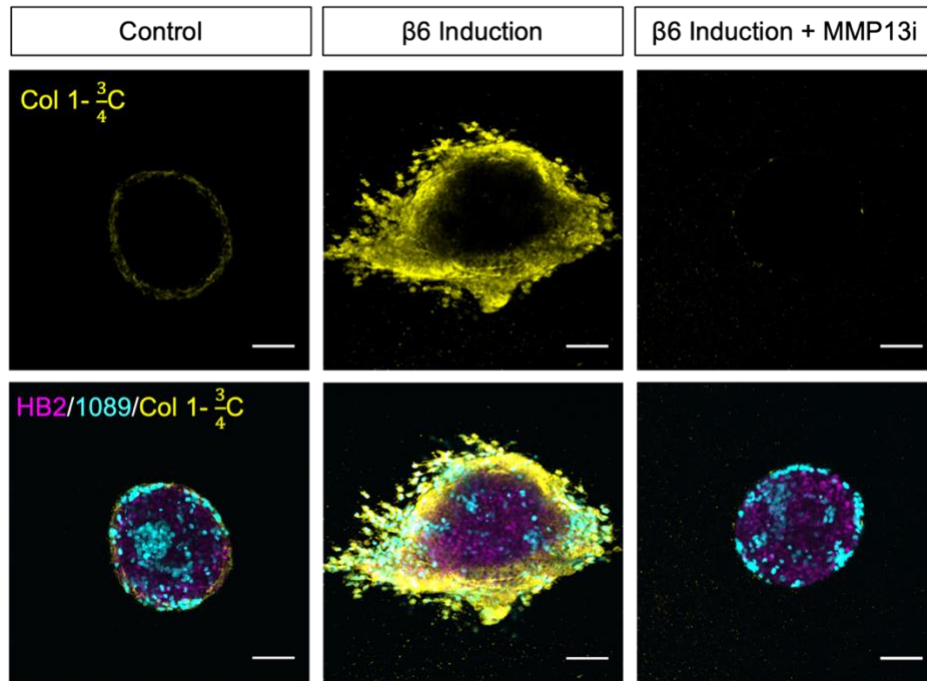
To confirm that this invasive phenotype was dependent on the proteolytic activity of MMP13 and not just its expression, integrin  $\beta$ 6-induced spheroids were treated with an MMP13 specific inhibitor (DB04760; 1  $\mu$ M) (279). The percentage of invasive area, number of projections and the projection length were reduced significantly across both the HB2/1089 <sup>$\beta$ 6</sup> and HB2/Myo <sup>$\beta$ 6</sup> models, demonstrating that both MMP13 expression and its activity are required to facilitate invasion (Fig 4.2, 4.3). To observe changes in collagen-I cleavage following MMP13 inhibition, HB2/1089 <sup>$\beta$ 6</sup> spheroids were subjected to immunofluorescent analysis of collagen-I proteolysis. Levels of cleaved collagen-I were reduced in MMP13i treated conditions (Fig 4.4), consistent with the observed reduction in invasion.





**Figure 4.3 Integrin  $\beta 6$ -driven invasion is dependent on MMP13 activity in the HB2/Myo <sup>$\beta 6$</sup>  model.**

HB2 spheres were embedded in Myo <sup>$\beta 6$</sup> -containing collagen gels and treated with 1  $\mu\text{g}/\text{ml}$  doxycycline to induce integrin  $\beta 6$  expression, and 1  $\mu\text{M}$  MMP13i (DB04760) to inhibit MMP13 activity. **(A)** Representative light micrographs of spheres at 3 d post treatment. **(B)** Summary graphs showing % invasive area, number of projections, projection length ( $\mu\text{m}$ ) and central area ( $\mu\text{m}^2$ ). Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. ns = not significant, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$  (Kruskal-Wallis test with multiple comparisons). Scale bar = 100  $\mu\text{m}$ .



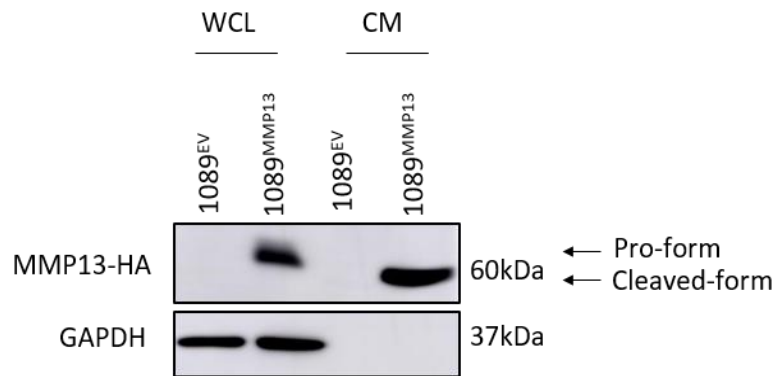
**Figure 4.4 Invasion requires MMP13-dependent collagen-I cleavage in HB2/1089 <sup>$\beta 6$</sup>  model.**

Spheres comprised of fluorescently labelled HB2/1089 <sup>$\beta 6$</sup>  spheroids were embedded in collagen gels and treated with 1  $\mu$ g/ml doxycycline to induce myoepithelial integrin  $\beta 6$  expression, and 1  $\mu$ M MMP13i (DB04760), to inhibit MMP13 activity. Representative fluorescence images of invading spheroids at 4 d post treatment with HB2 (magenta), Myo <sup>$\beta 6$</sup>  (cyan) and cleaved collagen-I (yellow). Scale bar = 100  $\mu$ m.

#### 4.2.3. Integrin $\beta 6$ -independent overexpression of MMP13 drives invasion

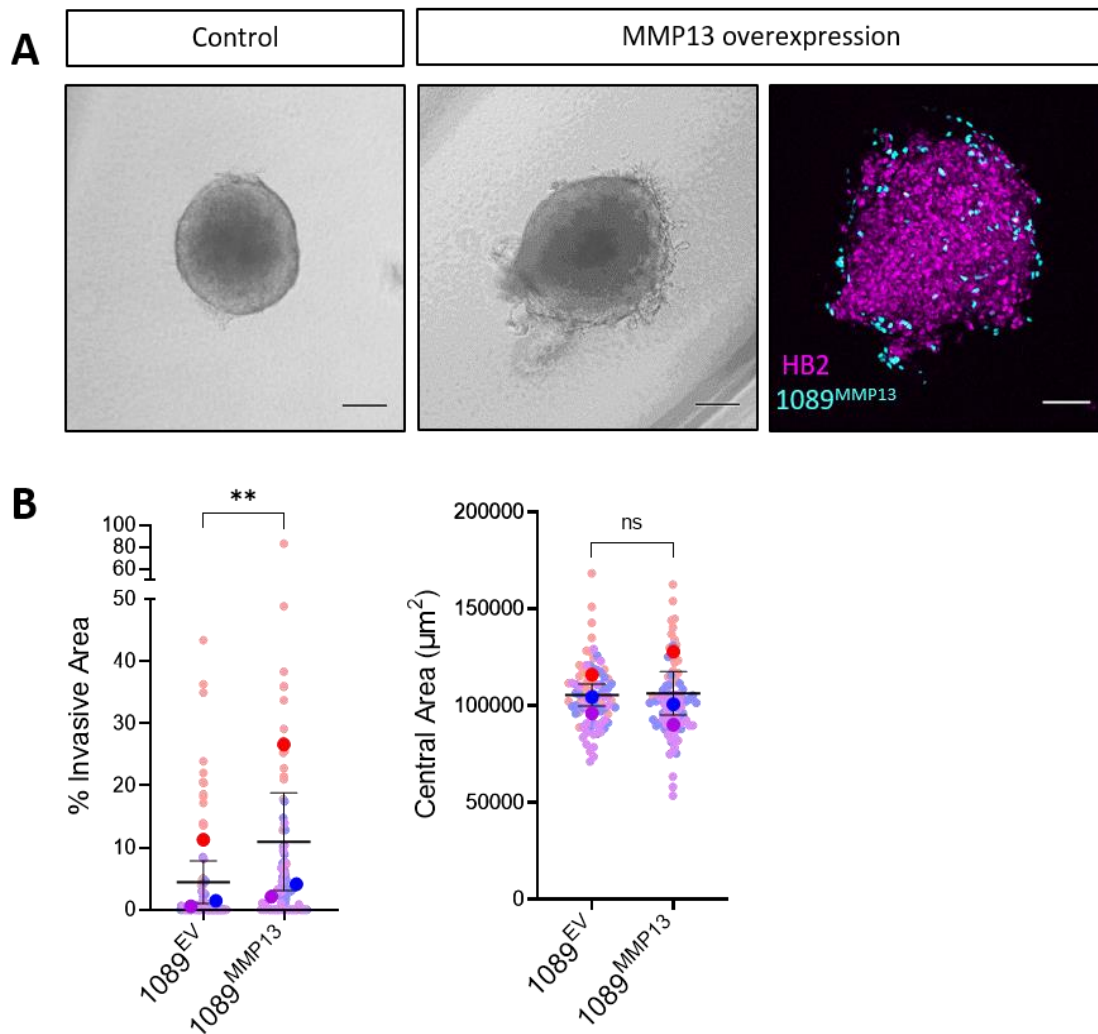
We next wanted to determine whether overexpression of MMP13 could drive invasion independently of integrin  $\beta 6$ . To achieve this, we engineered a constitutive MMP13-HA tag-expressing lentiviral construct using the pLV-CMV-Puro destination vector (Fig 2.1) and transfected non integrin  $\beta 6$ -inducible parental 1089 cells (1089<sup>MMP13</sup>). Efficacy of MMP13 overexpression was confirmed by western blot, with a significant increase in MMP13-HA detected in both the whole cell lysate and conditioned medium, compared to the 1089 empty vector control (1089<sup>EV</sup>) (Fig 4.5).

To assess the ability of MMP13 to drive invasion, 1089<sup>MMP13</sup> cells were then combined with HB2 cells to form spheroids prior to embedding in collagen gels. Strikingly, integrin  $\beta 6$  independent expression of MMP13 in 1089<sup>MMP13</sup> cells was able to drive invasion compared to the empty vector (EV) 1089<sup>EV</sup> control, with a significant increase in the percentage of invasive area (Fig 4.6A,B). Noticeably, 1089<sup>MMP13</sup> cells exhibited more dispersed invasion out from the sphere (Fig 4.6A), compared to integrin  $\beta 6$ -induced spheroids, likely due to higher overall expression of MMP13. Taken together, this supports the role of myoepithelial-derived MMP13 as an active driver of integrin  $\beta 6$ -dependent myoepithelial-led invasion.



**Figure 4.5 Validation of pLV CMV MMP13 HA-Tag construct.**

1089 cells were transfected with constructs constitutively expressing MMP13 HA-Tag and media and cells were harvested after 72 h. Western blots using anti-HA antibodies confirmed sufficient overexpression of MMP13 in the whole cell lysate (WCL) and conditioned media (CM).



**Figure 4.6 MMP13 drives invasion independently of integrin  $\beta 6$  in the HB2/1089 model.**

HB2/1089 spheres, consisting of either 1089<sup>EV</sup> or 1089<sup>MMP13</sup> were embedded in collagen gels. **(A)** Representative light micrographs of spheres 4 d post embedding in collagen. **(B)** Summary graphs showing % invasive area and central area ( $\mu\text{m}^2$ ). Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. ns = not significant, \*\* $p < 0.01$  (Mann Whitney U Test). Scale bar = 100  $\mu\text{m}$ .

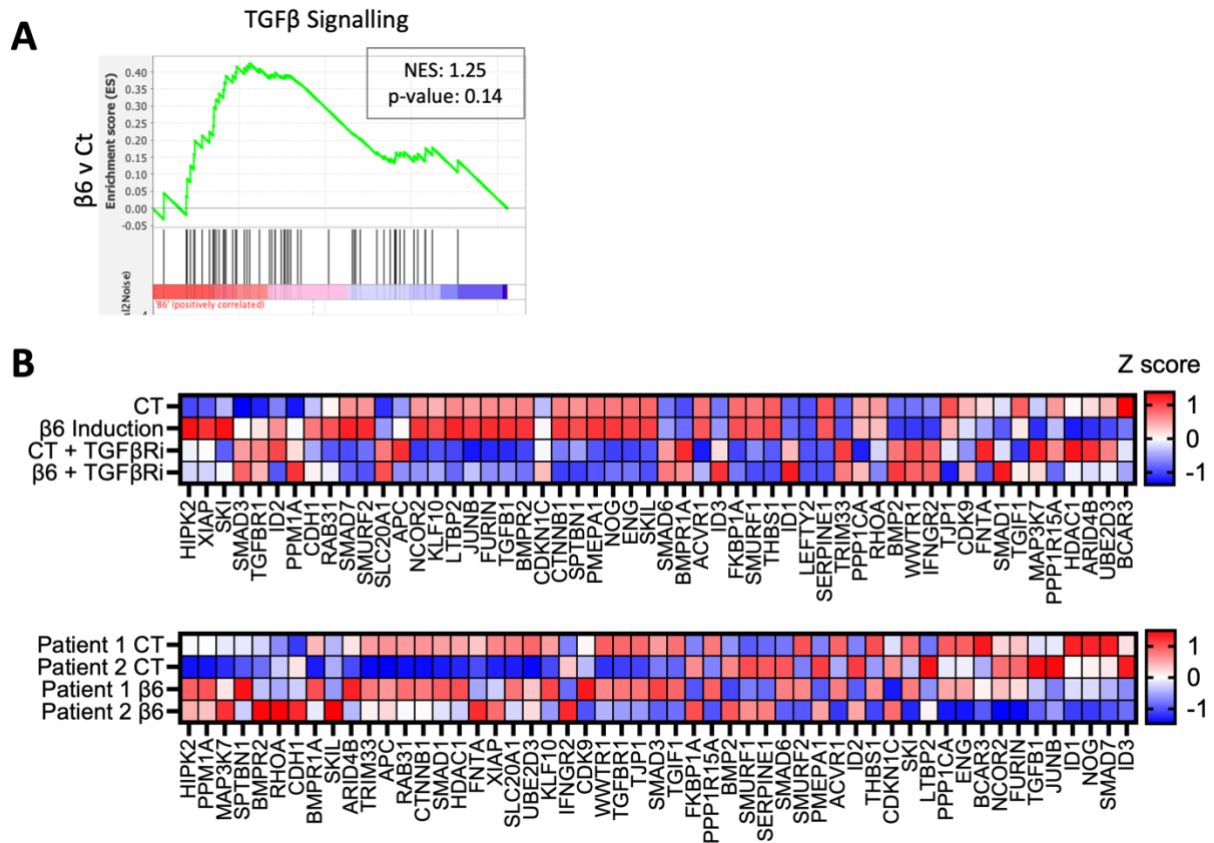
### 4.3. Integrin $\beta 6$ -driven MMP13 upregulation is TGF $\beta$ -dependent

#### 4.3.1. Integrin $\beta 6$ expression drives canonical TGF $\beta$ signalling

Having determined integrin  $\beta 6$ -induced MMP13 to be a driver of invasion, we next sought to investigate the mechanism by which integrin  $\beta 6$  expression was driving increased MMP13 expression. As well-characterised ligands of integrin  $\alpha v\beta 6$  include the TGF $\beta$  pro-peptide (280-282), we hypothesised that integrin  $\beta 6$  expression would increase TGF $\beta$  signalling. Indeed, gene set enrichment analysis (GSEA) of RNAseq data suggested an enrichment of the TGF $\beta$  signalling signature upon integrin  $\beta 6$  induction in 1089 <sup>$\beta 6$</sup>  cells (Fig 4.7A,B), with TGF $\beta$  signalling components also upregulated in Myo <sup>$\beta 6$</sup>  cells (Fig 4.7B).

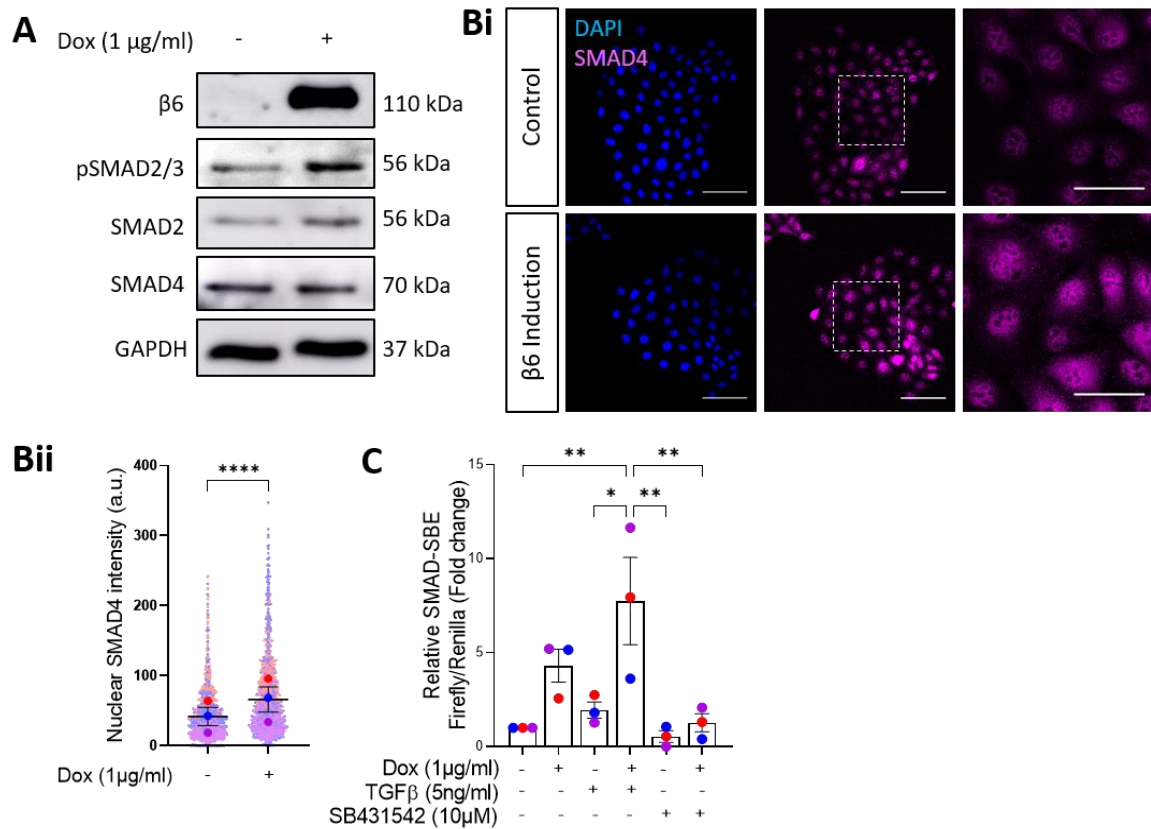
Increased basal canonical TGF $\beta$  signalling was also observed by western blot analysis, revealing enhanced phosphorylation of SMAD2/3 following integrin  $\beta 6$  induction in 1089 <sup>$\beta 6$</sup>  cells (Fig 4.8A). Although total SMAD4 expression remained unchanged following integrin  $\beta 6$  induction, an increase in its nuclear localisation was observed by immunofluorescence, suggesting an increase in SMAD4-mediated transcription (Fig 4.8B). To confirm this, cells were transfected with a SMAD-SMAD binding element (SBE) luciferase reporter construct, prior to integrin  $\beta 6$  induction alongside TGF $\beta$ R inhibition and/or TGF $\beta$  stimulation. Indicative of increased TGF $\beta$  signalling, luminescence analysis revealed an increase in SMAD driven transcription upon integrin  $\beta 6$  induction, both with and without TGF $\beta$  stimulation (Fig 4.8C). These data confirm that integrin  $\beta 6$  expression drives increased TGF $\beta$  signalling in breast myoepithelial cells.





**Figure 4.7 TGFβ signalling is enriched following myoepithelial integrin β6 induction.**

1089<sup>iβ6</sup> or primary Myo<sup>iβ6</sup> cells grown in 2D were treated with 1 μg/ml doxycycline for 48 h to induce integrin β6 expression. 1089<sup>iβ6</sup> cells were also treated with 10 μM of TGFβR inhibitor (SB435142). Differential expression analysis based on RNAseq data was used to perform gene set enrichment analysis. **(A)** Enrichment plot of TGFβ signalling gene set in integrin β6-induced versus control 1089<sup>iβ6</sup> cells. NES = normalised enrichment score. **(B)** Heat maps showing expression of genes from the TGFβ signalling gene set in 1089<sup>iβ6</sup> or Myo<sup>iβ6</sup> cells upon integrin β6 induction. For this experiment, C. Mein performed the RNA sequencing and provided the normalised counts. E. Tomas Bort performed the differential expression analysis which S. Gibson used to perform the GSEA. GSEA plots were extracted from the GSEA software, while heatmaps were generated by S. Gibson.

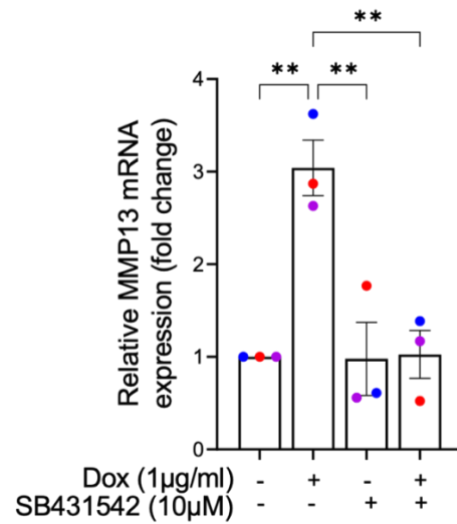


**Figure 4.8 Integrin  $\beta 6$  induction drives canonical TGF $\beta$  signalling.**

1089<sup>i $\beta 6$</sup>  cells grown in 2D were treated with 1  $\mu\text{g/ml}$  doxycycline for 48 h to induce integrin  $\beta 6$  expression. **(A)** Western blots using integrin  $\beta 6$ , pSMAD2/3, SMAD2, SMAD4 and GAPDH-directed antibodies. **(Bi)** Representative immunofluorescence staining of DAPI (blue) and SMAD4 (magenta) and **(Bii)** quantification of nuclear SMAD4 intensity in 1089<sup>i $\beta 6$</sup>  cells. Data are presented as mean  $\pm$  SEM where each dot represents one cell with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. Scale bar = 100  $\mu\text{m}$  (or 50  $\mu\text{m}$  for inserts). \*\*\*\* $p < 0.0001$  (Two-tailed t-test). **(C)** Relative luminescence of 1089<sup>i $\beta 6$</sup>  cells grown in 2D and transfected with SMAD-SBE luciferase constructs prior to treatment with 1  $\mu\text{g/ml}$  doxycycline for 48 h to induce integrin  $\beta 6$  expression, 10  $\mu\text{M}$  of TGF $\beta$ RI (SB435142), and/or 5 ng/ml TGF $\beta$  stimulation. Data are presented as  $\pm$  SEM, with biological replicates indicated by different colours. \* $p < 0.05$ , \*\* $p < 0.01$  (One-way ANOVA with multiple comparisons). The SMAD-SBE Luciferase reporter construct was provided by Dr. A. Cameron.

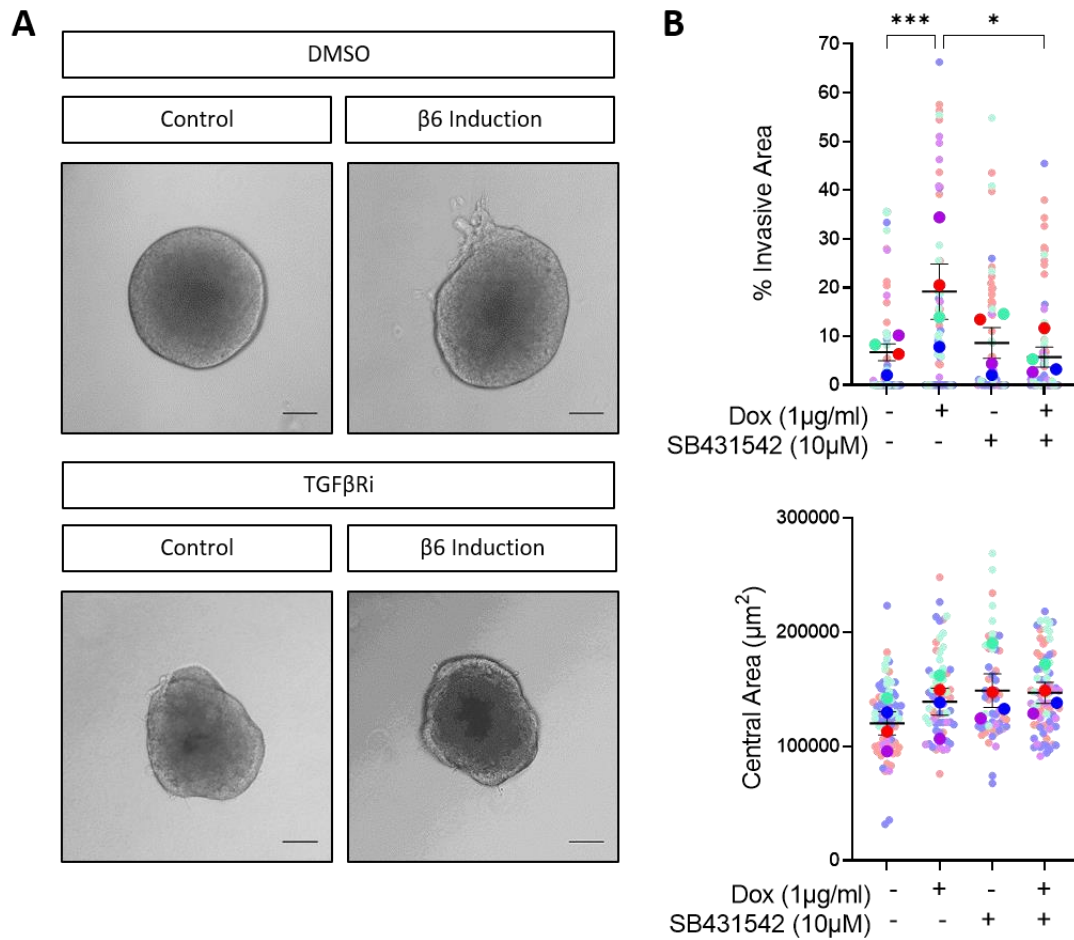
#### 4.3.2. Inhibition of TGF $\beta$ signalling blocks integrin $\beta$ 6-driven MMP13 upregulation

To determine whether integrin  $\beta$ 6-induced MMP13 upregulation was dependent on TGF $\beta$  signalling, integrin  $\beta$ 6-induced 1089<sup>i $\beta$ 6</sup> cells were treated with the TGF $\beta$ R inhibitor SB431542 for 48 h prior to qPCR analysis. Transcriptional upregulation of MMP13 was blocked upon integrin  $\beta$ 6 induction, indicating that integrin  $\beta$ 6-driven MMP13 upregulation occurs through a TGF $\beta$ -dependent mechanism (Fig 4.9). To confirm the role of TGF $\beta$  signalling in facilitating integrin  $\beta$ 6-induced invasion in a 3D context, HB2/1089<sup>i $\beta$ 6</sup> spheres were treated with SB431542 for 4 days and the effect on invasion was assessed. Accordingly, integrin  $\beta$ 6-induced spheres showed reduced invasion following TGF $\beta$ R inhibition (Fig 4.10).



**Figure 4.9 Integrin  $\beta 6$ -driven upregulation of MMP13 is dependent on TGF $\beta$  signalling.**

1089 <sup>$\beta 6$</sup>  cells grown in 2D were treated with 1 µg/ml doxycycline for 48 h to induce myoepithelial integrin  $\beta 6$  expression and 10 µM TGF $\beta$ Ri (SB435142) to inhibit TGF $\beta$  signalling. qPCR analysis was performed to determine relative MMP13 mRNA expression compared to control. Data are presented as mean  $\pm$  SEM. \*\* $p < 0.01$  (One-way ANOVA with multiple comparisons).



**Figure 4.10  $\beta 6$ -driven invasion is dependent on TGF $\beta$  signalling in the HB2/1089<sup>i $\beta 6$</sup>  model.**

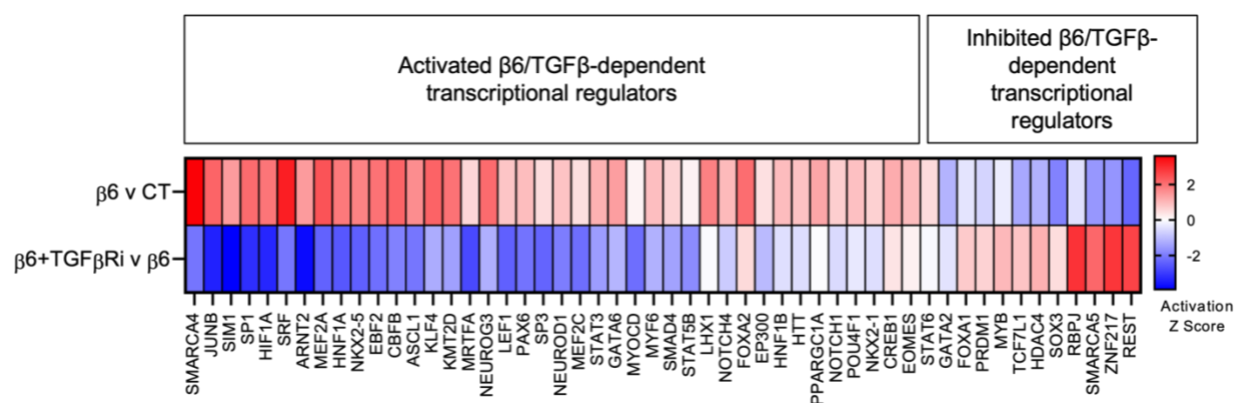
HB2/1089<sup>i $\beta 6$</sup>  spheres were embedded in collagen gels and treated with 1  $\mu\text{g}/\text{ml}$  doxycycline to induce integrin  $\beta 6$  expression and 10  $\mu\text{M}$  TGF $\beta$ Ri (SB435142) to inhibit TGF $\beta$  signalling. **(A)** Representative light micrographs of spheres at 4 d post treatment. **(B)** Summary graphs showing % invasive area and central area ( $\mu\text{m}^2$ ). Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \* $p < 0.05$ , \*\*\* $p < 0.001$  (Kruskal-Wallis test with multiple comparisons). Scale bar = 100  $\mu\text{m}$ .

## 4.4. TGF $\beta$ -dependent expression of MMP13 requires activation of EP300

### 4.4.1. Identification of integrin $\beta$ 6/TGF $\beta$ -dependent transcriptional regulators

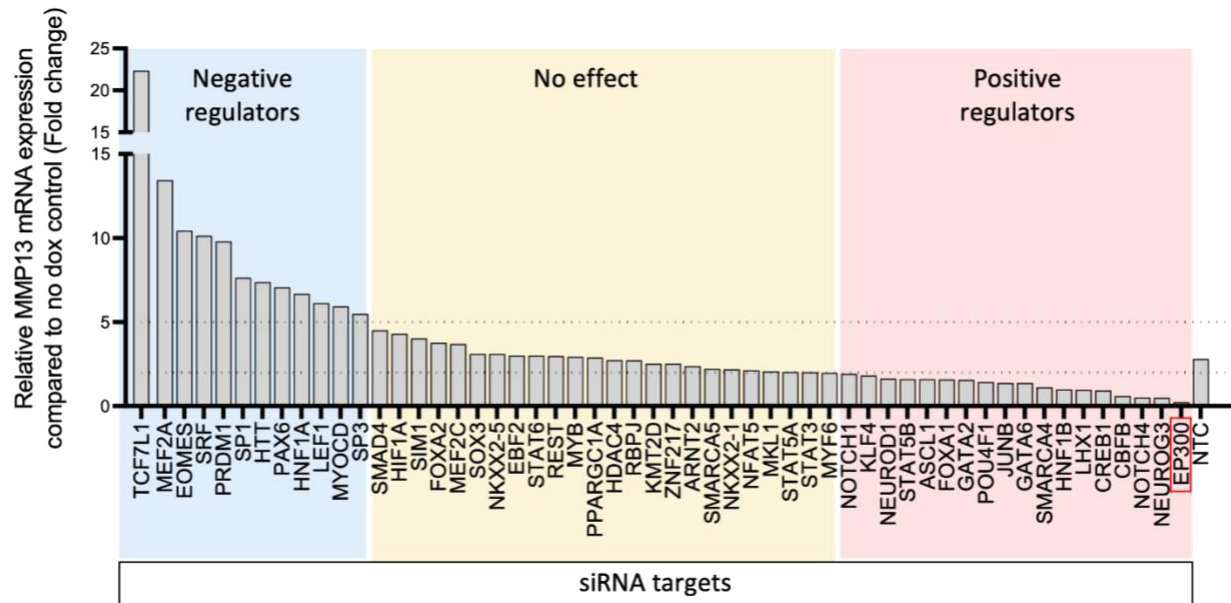
After identifying integrin  $\beta$ 6-driven MMP13 upregulation to be TGF $\beta$ -dependent, we next investigated direct transcriptional regulators downstream of TGF $\beta$  signalling that could regulate MMP13 expression. Given our finding that MMP13 expression was regulated independently of SMAD4 (Appendix 2.2), we used Ingenuity Pathway Analysis to predict integrin  $\beta$ 6-driven, TGF $\beta$ -dependent transcriptional regulators, determined from our 1089<sup>i $\beta$ 6</sup> RNAseq data. This approach identified 54 integrin  $\beta$ 6/TGF $\beta$ -dependent hits. While 43 were activated upon integrin  $\beta$ 6-induction and subsequently inhibited following TGF $\beta$ Ri treatment, 11 were inhibited upon integrin  $\beta$ 6-induction and subsequently activated following TGF $\beta$ Ri treatment (Fig 4.11).

To identify regulators specific to MMP13, each of the 54 hits were individually knocked down by siRNA, and the impact on integrin  $\beta$ 6-driven MMP13 transcription was assessed by qPCR analysis. As integrin  $\beta$ 6-induced MMP13 mRNA upregulation typically fell between 2-5-fold in NTC conditions, knockdown of regulators that resulted in below 2-, or above 5- fold, were classed as positive and negative regulators, respectively (Fig 4.12). This approach identified the histone acetyl transferase EP300 as a potential MMP13 regulator, as its knockdown had the greatest effect on blocking integrin  $\beta$ 6-driven MMP13 expression (Fig 4.12, Appendix 2.3).



**Figure 4.11 Activation Z-scores of integrin  $\beta 6$ /TGF $\beta$ -dependent transcriptional regulators.**

1089<sup>i</sup> $\beta 6$  cells grown in 2D were treated with 1  $\mu$ g/ml doxycycline to induce myoepithelial integrin  $\beta 6$  expression and 10  $\mu$ M TGF $\beta$ Ri (SB435142) to inhibit TGF $\beta$  signalling, for 48 h. Differential expression analysis based on RNAseq data was used to predict the activation score of upstream transcriptional regulators using Ingenuity Pathway Analysis Software. Heat map showing activation Z scores of integrin  $\beta 6$ /TGF $\beta$ -dependent transcriptional regulators. Activated transcriptional regulators are in red, while inhibited transcriptional regulators are in blue. For this experiment, E. Tomas Bort performed the differential expression analysis from RNAseq normalised counts, which S. Gibson used to perform the upstream regulator analysis using IPA Qiagen software. Heatmap was produced by S. Gibson.



**Figure 4.12** SiRNA screen reveals EP300 as a potential transcriptional regulator of MMP13.

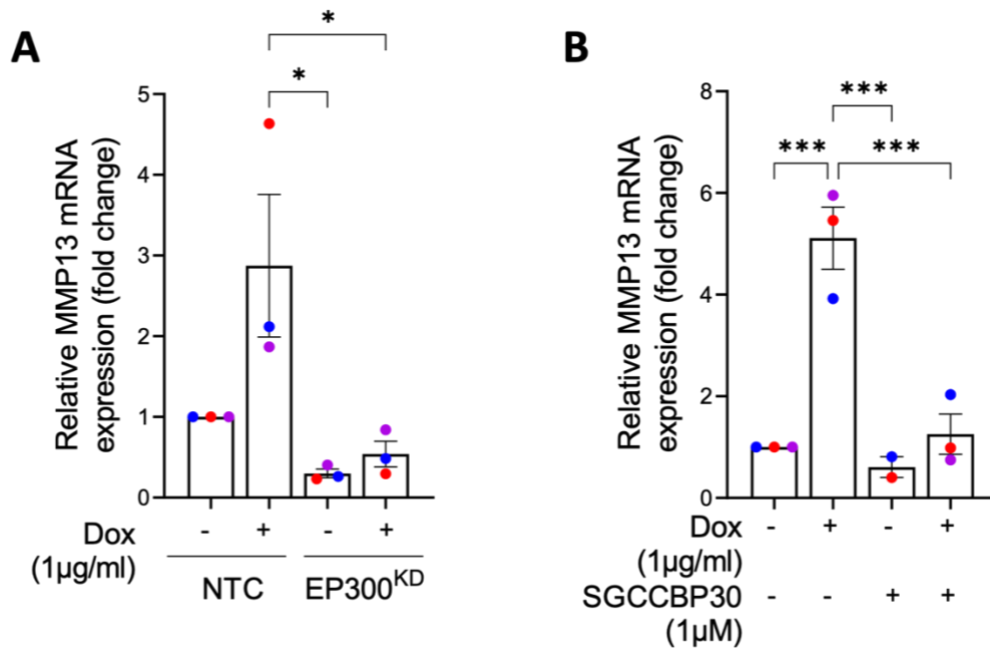
1089<sup>iB6</sup> cells grown in 2D were subjected to siRNA transfection against candidate integrin  $\beta$ 6/TGF $\beta$ -dependent transcriptional regulators prior to treatment with 1  $\mu$ g/ml doxycycline for 48 h to induce myoepithelial integrin  $\beta$ 6 expression and MMP13 upregulation. Relative MMP13 mRNA expression of siRNA-treated 1089<sup>iB6</sup> cells compared to non-targeting control (NTC), 72 h post transfection. Fold change < 2 indicates positive regulators of MMP13 (red) and fold change > 5 indicates negative regulators of MMP13 (blue).



#### 4.4.2. Validation of EP300 as a driver of MMP13 expression

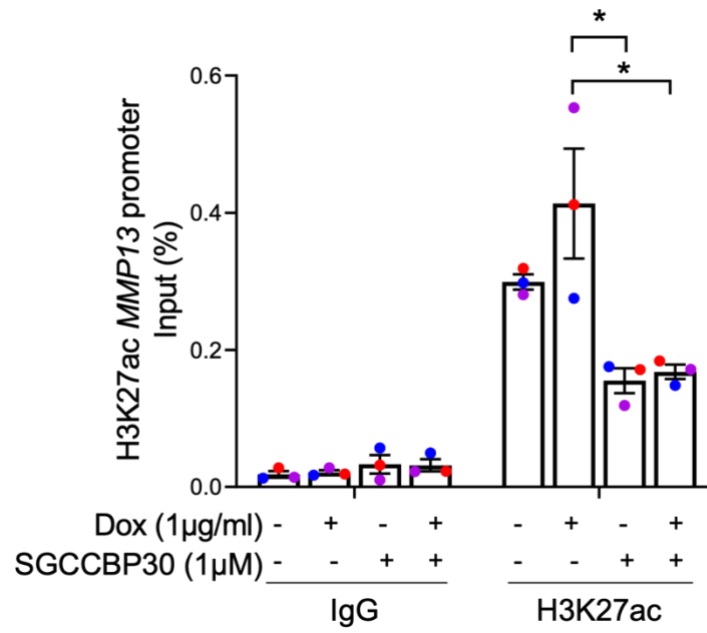
As siRNA-mediated knockdown of EP300 was able to block integrin  $\beta$ 6-mediated upregulation of MMP13 (Fig 4.13A), we next tested whether inhibition of EP300 activity could yield similar results. Using the dual CREBBP/EP300 inhibitor, SGCCBP30, which works by binding to bromodomains required for chromatin binding (283), we treated integrin  $\beta$ 6-induced 1089<sup>i $\beta$ 6</sup> cells and assessed MMP13 mRNA expression by qPCR analysis. Treatment with SGCCBP30 yielded a loss in integrin  $\beta$ 6-dependent MMP13 upregulation, further implicating EP300 activity as a requirement for integrin  $\beta$ 6-induced MMP13 expression (Fig 4.13B).

Next, to confirm the regulation of MMP13 by EP300, ChIP was performed to analyse for enrichment of H3K27ac, an epigenetic histone acetylation mark associated with transcriptional activation, at the *MMP13* promoter region following EP300 inhibition. Integrin  $\beta$ 6-induced 1089<sup>i $\beta$ 6</sup> cells were treated with SGCCBP30 prior to chromatin extraction and immunoprecipitation with an H3K27ac targeted antibody. ChIP-qPCR analyses using genomic primer pairs specific for the *MMP13* promoter suggested an enrichment of H3K27ac following integrin  $\beta$ 6 induction, although this was not significant. Strikingly, treatment with SGCCBP30 was able to diminish levels of H3K27ac at the *MMP13* promoter region upon integrin  $\beta$ 6 induction (Fig 4.14), suggestive of a mechanism where EP300 activation drives transcriptomic expression of MMP13.



**Figure 4.13 MMP13 upregulation is dependent on expression and activity of EP300.**

**(A)** 1089<sup>β6</sup> cells grown in 2D were subjected to EP300-targeted siRNA (EP300<sup>KD</sup>) transfection and treated with 1 µg/ml doxycycline for 48 h to induce integrin β6 expression. RNA was extracted 72 h post transfection and subjected to qPCR analysis. Relative mRNA expression of MMP13 compared to the non-targeting control (NTC), non-integrin β6-induced condition. **(B)** 1089<sup>β6</sup> cells grown in 2D and treated with 1 µg/ml doxycycline and/or 1 µM SGCCBP30 for 48 h to block EP300 activity. Harvested RNA was subjected to qPCR analysis. Relative mRNA expression of MMP13 compared to the DMSO-treated, non-integrin β6-induced condition. Data are presented as mean ± SEM. \*p<0.05, \*\*\*p<0.001 (One-way ANOVA with multiple comparisons). The SGCCBP30 inhibitor was provided by Dr. B. Peck.



**Figure 4.14 EP300 regulates MMP13 expression through H3K27ac enrichment.**

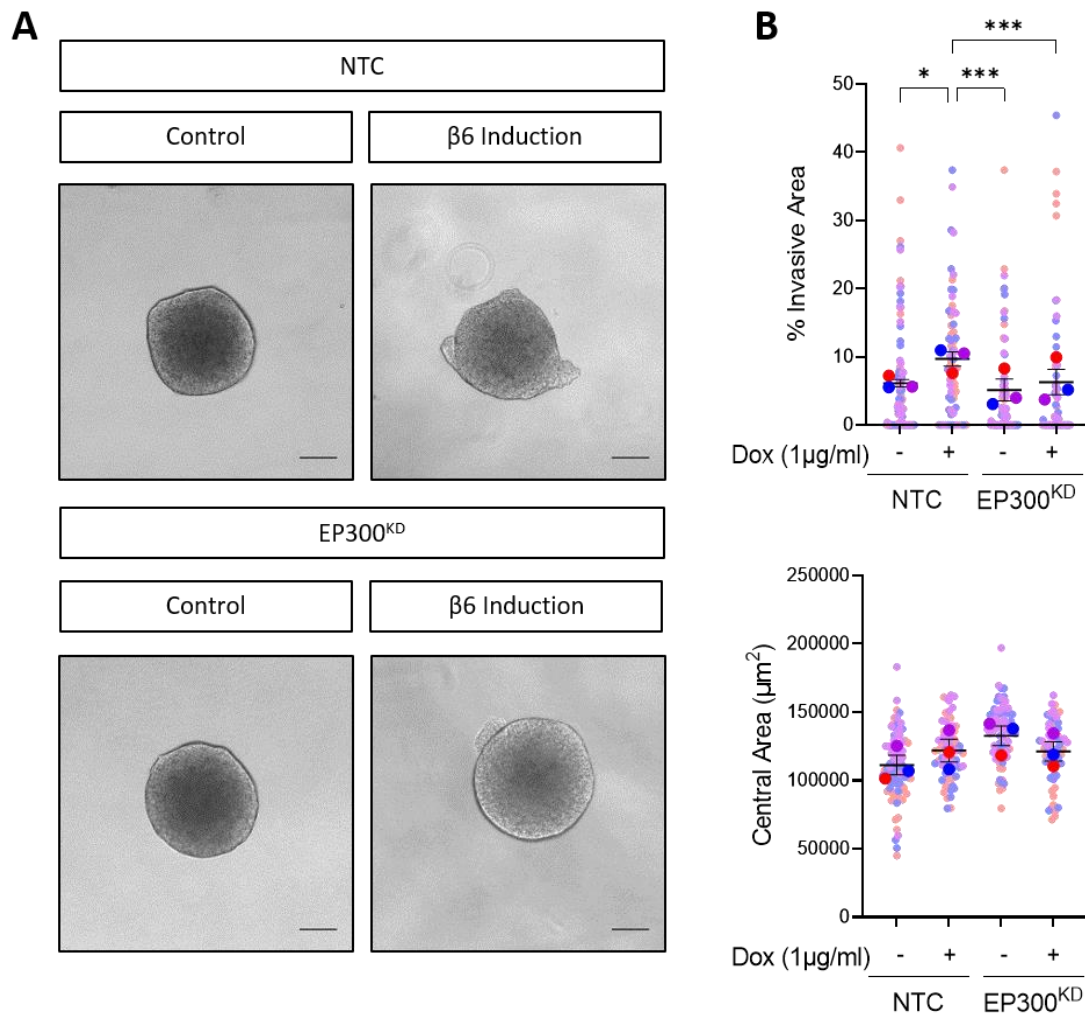
1089<sup>iβ6</sup> cells grown in 2D were treated with 1 µg/ml doxycycline and/or 1 µM SGCCBP30 for 48 h. Chromatin was harvested and subjected to immunoprecipitation using an anti-H3K27ac antibody. ChIP-qPCR analysis for enrichment of H3K27ac at the *MMP13* promoter region using genomic *MMP13* primers. Data are presented as mean ± SEM. \*p<0.05 (One-way ANOVA with multiple comparisons). The SGCCBP30 inhibitor was provided by Dr. B. Peck.

#### 4.4.3. Knock-down of EP300 reduces integrin $\beta$ 6-driven invasion

To confirm the role of EP300 as a driver of invasion in a 3D context, we next examined whether loss of myoepithelial EP300 in the HB2/1089 <sup>$\beta$ 6</sup> sphere model was sufficient to reduce invasion. Effective knockdown of EP300 by siRNA was confirmed by western blot analysis prior to embedding spheres into collagen gels (Fig 4.15) (Appendix 2.3). Consistent with the previous findings implicating EP300 as a regulator of integrin  $\beta$ 6 driven MMP13 expression, EP300 siRNA-treated spheres (EP300<sup>KD</sup>) displayed reduced invasion compared to NTC counterparts, characterised by a significant decrease in the percentage of invasive area despite integrin  $\beta$ 6 induction (Fig 4.15). These data confirm that EP300 expression is required to drive integrin  $\beta$ 6-induced invasion.

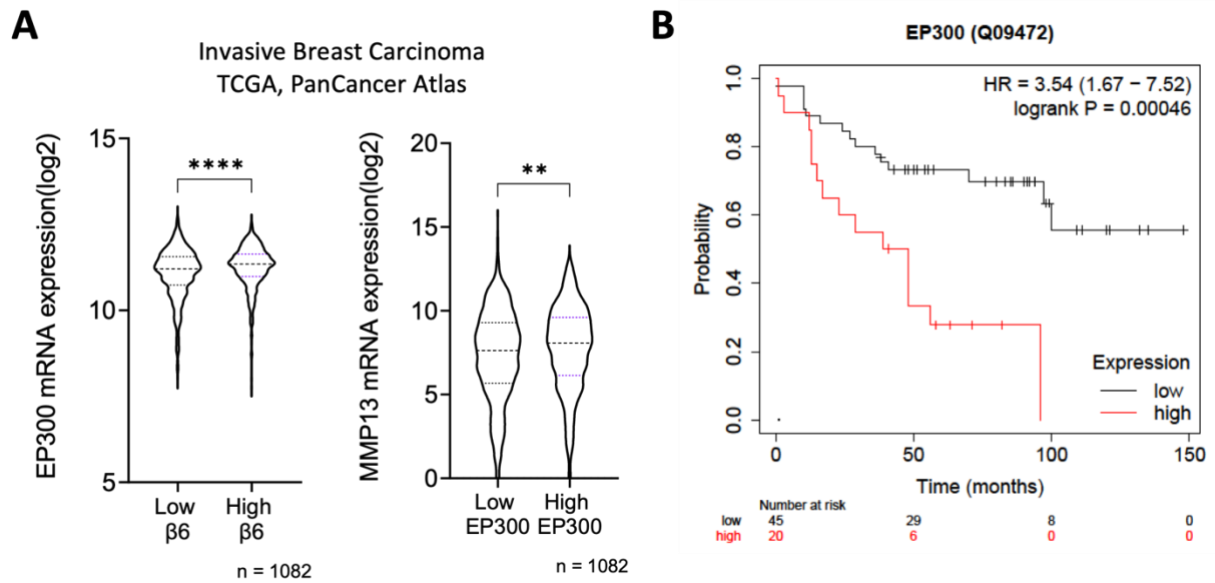
#### 4.4.4. EP300 expression is associated with poor prognosis in breast cancer

In line with its potential role in DCIS, data from publicly available invasive breast carcinoma repositories showed that high RNA expression of integrin  $\beta$ 6 in the bulk tumour was associated with higher levels of EP300. Furthermore, MMP13 expression was also significantly increased in high EP300 expressing patients (Fig 4.16A), supportive of our *in vitro* data. Survival analysis revealed that patients with tumours expressing higher levels of EP300 protein displayed poorer overall survival, with a hazard ratio of 3.54, further supporting that EP300 may also play a role in facilitating the later stages of breast cancer progression (Fig 4.16B).



**Figure 4.15 EP300 expression is required for  $\beta 6$ -driven invasion.**

Prior to co-culture, 1089 <sup>$\beta 6$</sup>  cells were transiently transfected with non-targeting control (NTC) or EP300 siRNA (EP300<sup>KD</sup>). Transfected 1089 <sup>$\beta 6$</sup>  cells were then combined with HB2 cells to form spheres, which were embedded in collagen gels and treated with 1  $\mu\text{g}/\text{ml}$  doxycycline to induce integrin  $\beta 6$  expression. **(A)** Representative light micrographs of spheres at 4 d, with either myoepithelial-NTC or -EP300<sup>KD</sup> and myoepithelial integrin  $\beta 6$  expression. **(B)** Summary graphs showing % invasive area and central area ( $\mu\text{m}^2$ ). Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \* $p < 0.05$ , \*\*\* $p < 0.001$  (Kruskal-Wallis test with multiple comparisons). Scale bar = 100  $\mu\text{m}$ .

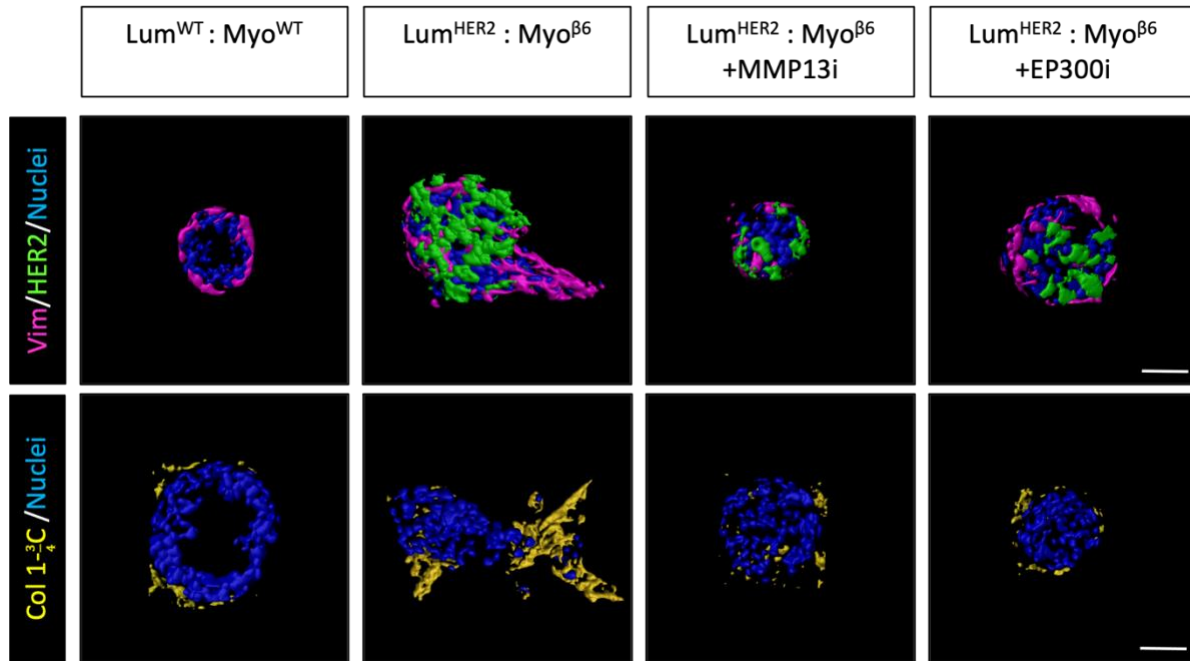


**Figure 4.16 EP300 expression is associated with poor prognosis in breast cancer.**

**(A)** EP300 and MMP13 mRNA expression in patient samples taken from the breast invasive carcinoma, TCGA, PanCancer Atlas data set. Low or high integrin  $\beta 6$  or EP300 classification was determined based on expression above or below integrin  $\beta 6$  or EP300 median expression. \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$  (Two-tailed t-test). **(B)** Kaplan Meier plots showing association between EP300 protein expression and overall survival from the Tang *et al.* breast cancer data set (284).

#### 4.4.5. Integrin $\beta$ 6-driven invasion is dependent on EP300/MMP13 in a duct model

Having shown that blocking the myoepithelial EP300/MMP13 axis was sufficient to prevent invasion of integrin  $\beta$ 6-expressing spheroid models, we next aimed to recapitulate this in our ductal model of DCIS, where primary luminal and myoepithelial cells from reduction mammoplasty patients are recombined in collagen-I to form ductal cultures (245). Upon the formation of these structures at 14 days post embedding, cultures were treated with doxycycline to induce luminal HER2 and myoepithelial integrin  $\beta$ 6 expression, which results in luminal filling and myoepithelial-led invasion (Fig 3.2). To assess the role of the EP300/MMP13 axis, cultures were also treated with EP300i (SGCCBP30), or MMP13i (DB04760) to block EP300 and MMP13, respectively. Under integrin  $\beta$ 6-induced conditions, cleaved collagen-I analysis was consistent with findings observed in the HB2/1089 <sup>$\beta$ 6</sup> sphere model, displaying increased collagen proteolysis surrounding myoepithelial-led protrusions (Fig 4.17). Strikingly, as with our spheroid model, inhibition of MMP13 reduced collagen proteolysis as well as invasion into the surrounding matrix, confirming MMP13 as an active driver of integrin  $\beta$ 6-dependent invasion (Fig 4.17). Targeting of EP300 was also sufficient to reduce both invasion and collagen-I cleavage, highlighting EP300 as a critical regulator of MMP13 transcription in this model (Fig 4.17).



**Figure 4.17 Integrin  $\beta 6$ -driven invasion is EP300/MMP13-dependent in ductal model.**

Primary luminal and myoepithelial ductal structures formed 14 d post collagen embedding were subsequently treated with 1  $\mu$ g/ml doxycycline for a further 7 d to induce transgene expression. To inhibit activity of MMP13 and EP300, ducts were treated with 1  $\mu$ M MMP13i (DB04760) and 1  $\mu$ M EP300i (SGCCBP30), respectively. Imaris reconstructions of luminal/myoepithelial ductal structures after 21 d of culture with HER2 luminal expression, myoepithelial integrin  $\beta 6$  expression and inhibition of MMP13 or EP300 where Nuclei (blue), vimentin (magenta), HER2 (green) and cleaved collagen (yellow). Scale bar = 20  $\mu$ m. This experiment was performed by Dr. E. Carter.

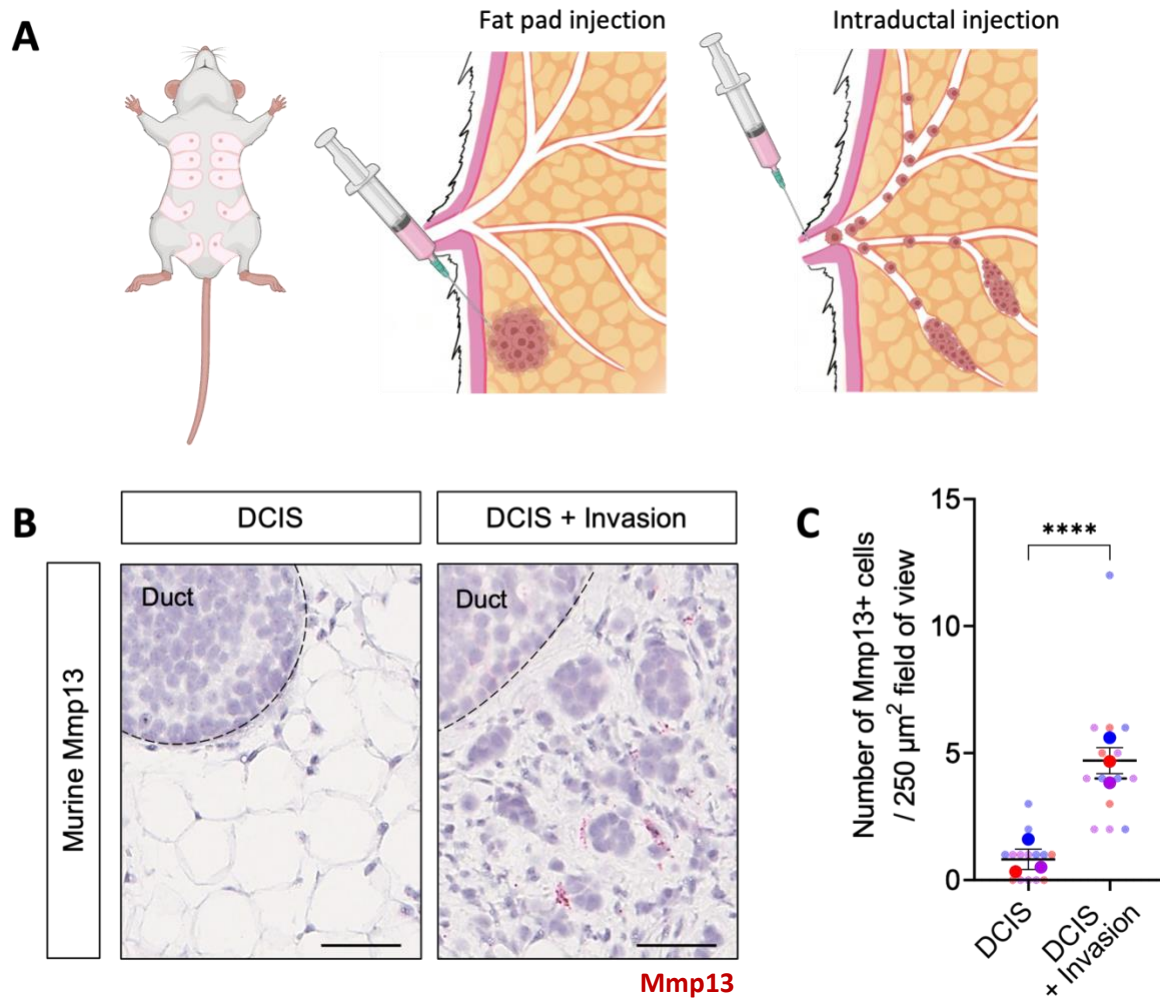


## 4.5. MMP13 expression *in vivo* and in clinical cases

### 4.5.1. Stromal Mmp13 expression is associated with invasion *in vivo*

To investigate our findings *in vivo*, we next used the mouse mammary intraductal (MIND) model to examine the presence of Mmp13 in the tumour microenvironment of invasive margins (259, 285). In this model, MCF7 cancer cells are xenografted into the mouse mammary ductal tree and can be used to study DCIS progression, with the dispersal of tumour cells into the stroma occurring at approximately 12 weeks post intraductal injection (Fig 4.18A) (259, 285). The model presents advantages when studying DCIS, as tumours develop with a greater involvement of the microenvironment and demonstrate features more consistent with clinical breast cancer than traditional fat pad xenografts (259).

To detect mouse Mmp13 in the microenvironment, while avoiding detection of cancer cell derived MMP13, RNAscope analysis was performed, using mouse-specific Mmp13 probes. Interestingly, stromal Mmp13 expression was localised to the periphery of collectively invading cancer cells, compared to non-invasive areas of DCIS where Mmp13 expression was significantly reduced (Fig 4.18B,C). These data suggest that while MMP13 expression is not required for facilitating DCIS growth, it could be associated with invasive progression.

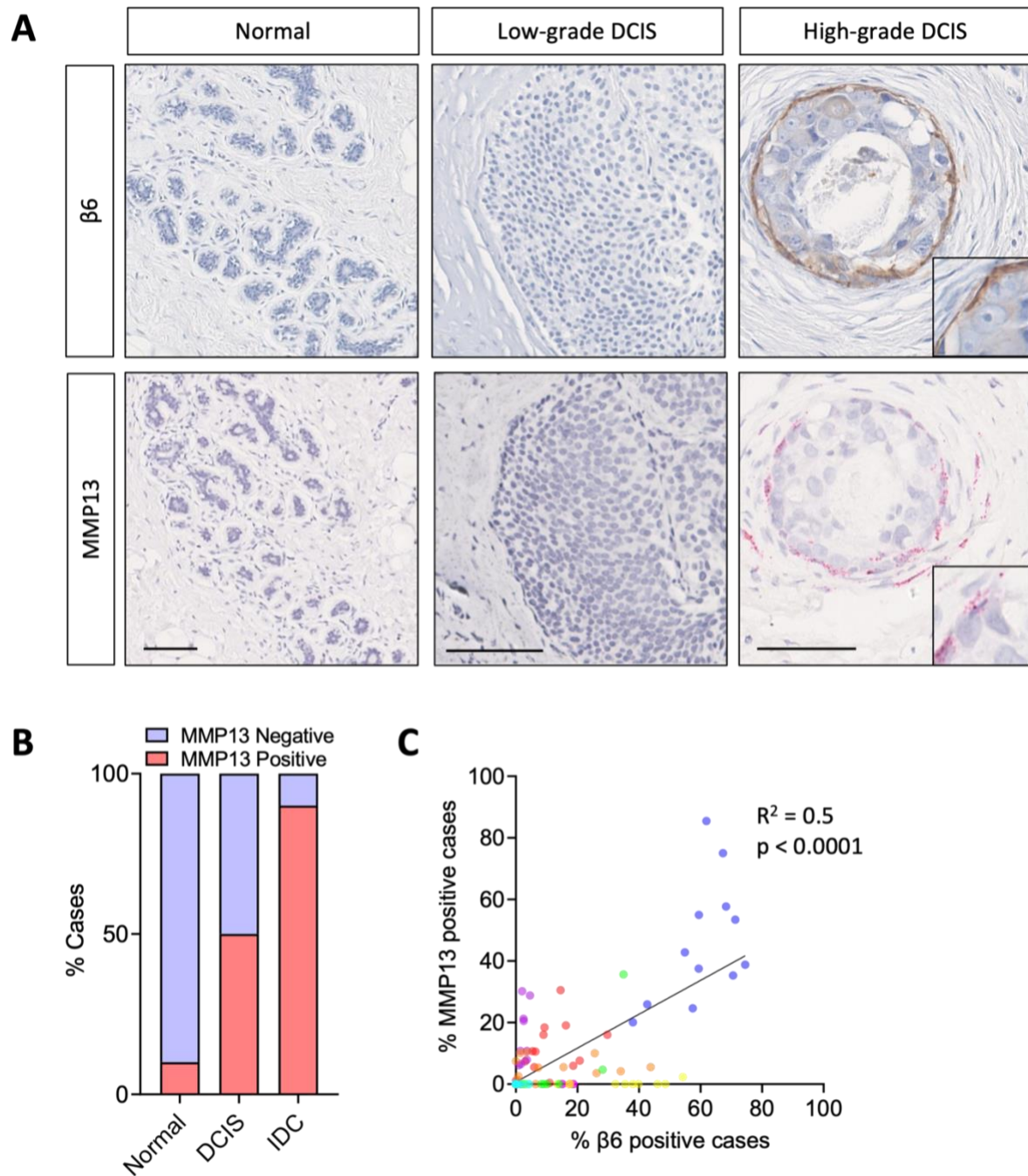


**Figure 4.18 Stromal Mmp13 is associated with invasion in intraductal mouse model.**

**(A)** Summary schematic of mammary fat pad versus intraductal mouse models of breast cancer progression. **(B)** Representative Mmp13 RNAscope images of non-invasive and invasive fronts from MCF7-MIND mice at 12 weeks post intraductal injection. Mmp13 expression is indicated by presence of red dots. Scale bar = 100 μm. **(C)** Graph showing number of Mmp13+ cells per 250 μm<sup>2</sup> field of view. Data are presented as mean ± SEM where each dot represents one field of view with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \*\*\*\*p<0.0001 (Two-tailed t-test) n = 3 mice. The intraductal model developed was developed by Prof. C. Briskens group. For this experiment, Dr. A. Agnoletto provided the tissue sections, while S. Gibson performed the RNAscope and quantification analyses.

#### 4.5.2. Myoepithelial cells express MMP13 in high-grade DCIS

To investigate myoepithelial MMP13 expression in the context of integrin  $\beta 6$  positive DCIS, 30 human breast samples, comprising 10 normal breast tissue, 10 DCIS and 10 IDC were analysed for expression of MMP13 and integrin  $\beta 6$  by RNAscope and immunohistochemistry, respectively. Strikingly, whereas 90% of normal samples were MMP13 negative (9 out of 10), MMP13 was detected in the myoepithelial and fibroblast cells of 50% (5 out of 10) of DCIS cases (Fig 4.19A,B). Notably, MMP13 positive cases were also high-grade DCIS, while MMP13 negative were either intermediate or low-grade DCIS. In MMP13 positive cases, MMP13 expression was absent from cancer cells and primarily localised to integrin  $\beta 6$  positive myoepithelial cells, as well as fibroblast cells in the tumour periphery, with a positive correlation observed between the percentage of MMP13 and integrin  $\beta 6$  positive cells per duct (Fig 4.19A,C).

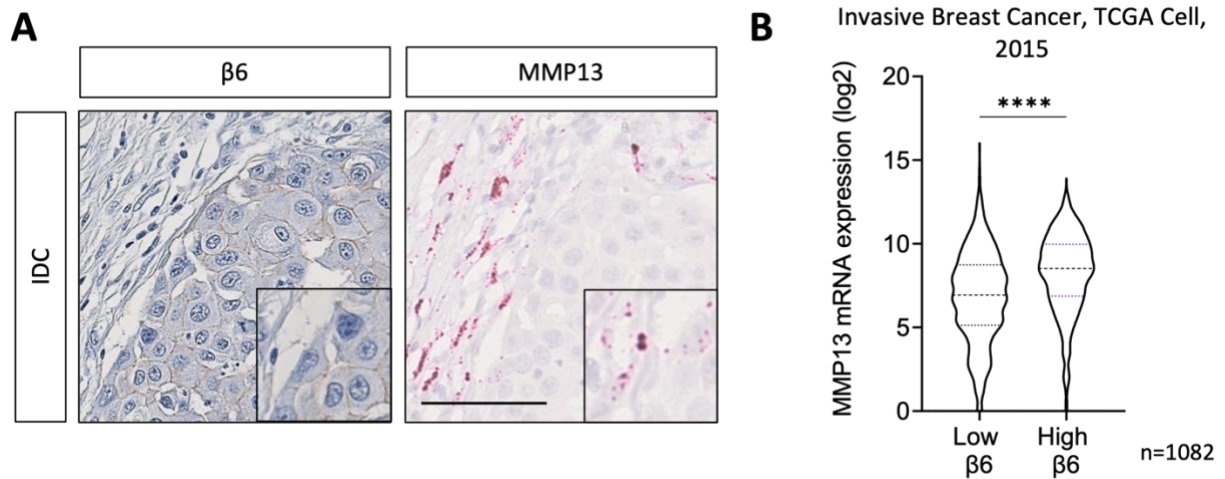


**Figure 4.19** MMP13 is localised to integrin  $\beta 6$ -expressing myoepithelial cells in DCIS.

Detection of MMP13 expression by RNAscope was performed in normal ( $n = 10$ ), DCIS ( $n = 10$ ) and IDC ( $n = 10$ ) cases. **(A)** Representative images of integrin  $\beta 6$  IHC (brown) staining and MMP13 RNAscope (red) detection in normal and DCIS patients. Scale bar = 100  $\mu\text{m}$ . **(B)** Summary graph showing percentage of MMP13 positive cases across disease states. **(C)** Graph showing percentage of MMP13 and integrin  $\beta 6$  positive cells/per duct across DCIS patients. Data are presented as individual points where each dot represents one duct, with different patients represented by different colours. (Simple linear regression). For this experiment, Prof. L. Jones provided the tissue sections, while N. Rahman performed the  $\beta 6$  IHC and S. Gibson performed the MMP13 RNAscope and quantification analyses of ducts.

### 4.5.3. MMP13 is expressed in IDC cases

In IDC, expression of MMP13 was observed in 90% (9 out of 10) of cases (Fig 4.19B), with expression localised to fibroblasts surrounding areas of invasion (Fig 4.20A). While integrin  $\beta 6$  expression was mostly absent in the cancer cell compartment of DCIS, cancer cells in IDC displayed increased integrin  $\beta 6$  expression, congruent with findings in a study of integrin  $\beta 6$  in breast cancer (207). These data suggest that once the initial stages of invasion have occurred, cancer cells and fibroblasts may upregulate integrin  $\beta 6$  and MMP13, respectively, to support invasion at later stages of disease. Supporting this, patient data from the breast invasive carcinoma TCGA PanCancer atlas data set, revealed a significant association between high integrin  $\beta 6$  and high MMP13 RNA expression in breast tumours (Fig 4.20B). Taken together, this implicates a progressive increase in MMP13 activity from normal breast tissue to DCIS and IDC, implicating a role for integrin  $\beta 6$  driven MMP13 in facilitating invasion.



**Figure 4.20** Expression of MMP13 is increased in IDC.

**(A)** Representative images of integrin  $\beta 6$  IHC (brown) staining and MMP13 RNAscope (red) detection in IDC patients. Scale bar = 100  $\mu\text{m}$ . **(B)** MMP13 mRNA expression in patient samples taken from the breast invasive carcinoma, TCGA Cell 2015 data set. Low or high integrin  $\beta 6$  classification was determined based on expression above or below median integrin  $\beta 6$  expression. \*\*\*\* $p < 0.0001$  (Two-tailed t-test). For this experiment, Prof. L. Jones provided the tissue sections, while N. Rahman performed the  $\beta 6$  IHC and S. Gibson performed the MMP13 RNAscope.

## 4.6. Discussion

### Results summary

- Integrin  $\beta 6$ -dependent myoepithelial-led invasion requires MMP13 expression and activity.
- Integrin  $\beta 6$ -driven upregulation of MMP13 is dependent on activation of the histone acetyltransferase EP300 upon enhanced TGF $\beta$  signalling.
- Expression of Mmp13 is associated with areas of invasion in intraductal mouse model of DCIS.
- MMP13 is elevated in myoepithelial cells of clinical high-grade DCIS cases.

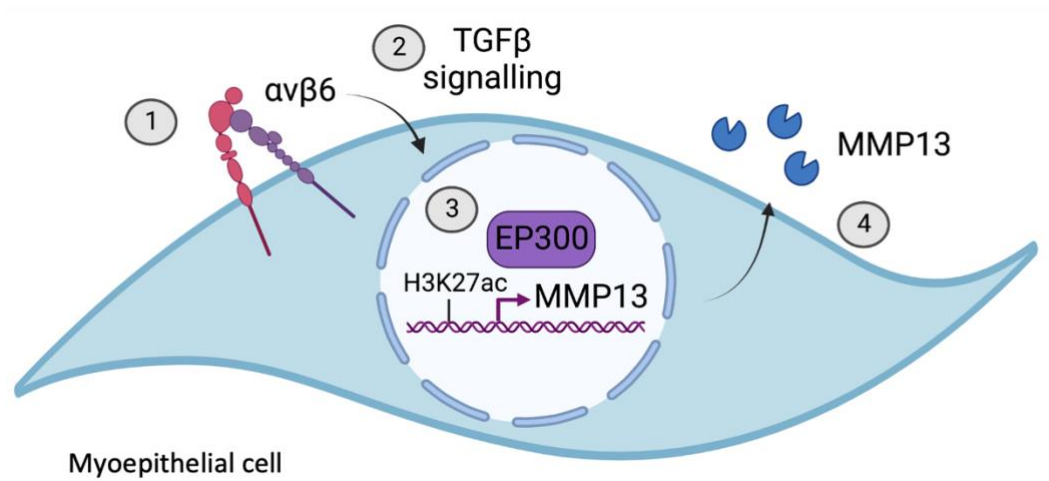
MMP13 has previously been implicated as a marker in the transition of DCIS to IDC, where its expression has been localised to stromal-myofibroblasts in areas of microinvasion (286). However, the functional significance of myoepithelial MMP13 as an active driver of disease progression has yet to be established. Here, we demonstrate that inhibition or knockdown of myoepithelial-derived MMP13 can significantly reduce invasion, implicating MMP13 as a crucial protease in facilitating the progression of early-stage breast cancer. These findings were supported by analyses of clinical DCIS samples, where MMP13 expression was observed in myoepithelial cells positive for integrin  $\beta 6$ . In IDC cases, MMP13 expression was increased further and detected in both myofibroblasts and cancer cells. This suggests that after the loss of the myoepithelial cell layer in the initial stages of invasion, integrin  $\beta 6$ -expressing cancer cells may upregulate MMP13 expression, alongside myofibroblasts, to enhance their invasive capabilities. This is reinforced by data from the breast invasive carcinoma TCGA cohort where we identified a positive correlation between expression of integrin  $\beta 6$  and MMP13 (287). Despite the robust and dramatic effect of MMP13 knockdown on invasion in our models, several other myoepithelial metzincins may also be involved in promoting the transition of DCIS to IDC (72, 160, 180).

We further show that MMP13 upregulation, mediated by integrin  $\beta 6$ , occurs via a TGF $\beta$ -dependent mechanism and can be blocked through inhibition of TGF $\beta$ R. These findings align with existing literature where integrin  $\beta 6$  drives the upregulation of several MMPs through TGF $\beta$  signalling (35). While TGF $\beta$  is often associated with promoting tumour progression in the later stages of disease, typically due to mutation or allelic loss of pathway components (211), we demonstrate that TGF $\beta$  signalling in the microenvironment may influence progression during the early stages of disease. This is further supported by immunohistochemical analysis, where overexpression of TGF $\beta$ Rs is associated with a worse prognosis in breast cancer (222, 288).

Our data suggest that while integrin  $\beta 6$ -driven MMP13 expression in myoepithelial cells is TGF $\beta$ -dependent, its transcriptional regulation occurs independently of canonical SMAD pathways, as SMAD4 knockdown had minimal effects on MMP13 expression. The most striking result from our integrin  $\beta 6$ /TGF $\beta$ -dependent transcriptional regulator screen, showed that knockdown of the histone acetyltransferase, EP300, resulted in significantly diminished MMP13 expression. These results are aligned with previous studies on MMP13 regulation in osteoblasts, where EP300 and RUNX2 are required to mediate activation of MMP13 during bone remodelling (289, 290). Indeed, our experiments show that in the context of myoepithelial cells, EP300 activity is critical for MMP13 expression, as either knockdown or inhibition were sufficient to block integrin  $\beta 6$ -driven MMP13 upregulation. Further analysis of publicly available datasets revealed that high expression of EP300 in breast cancer patients was associated with reduced overall survival, further supporting the notion that EP300 expression and activity contribute to disease progression (284). Overall, we provide evidence that myoepithelial cells play an active role in facilitating the invasion of luminal cells into the surrounding matrix. This is mediated by increased TGF $\beta$  signalling, which facilitates epigenetic



regulation of *MMP13* through EP300 (Fig 4.21). This mechanism may facilitate DCIS progression and provides a rationale for further investigation into myoepithelial *MMP13* as a high-risk marker.



**Figure 4.21 Proposed mechanism of the  $\beta 6$ /MMP13 axis in myoepithelial-led invasion.**

Upon upregulation of myoepithelial integrin  $\beta 6$  expression, TGF $\beta$  signalling is increased to drive activation of EP300 and subsequent expression of MMP13. Secretion of MMP13 into the surrounding stroma supports myoepithelial-led invasion via proteolytic remodelling of the basement membrane.

**Chapter 5. Loss of ADAMTS3 in DCIS-myoepithelial cells contributes to  
invasive progression through increased fibronectin/integrin  $\alpha 5\beta 1$   
signalling**

## 5.1. Introduction

While proteases are classically known for their pro-tumoural roles in facilitating invasion and disease progression, it is become increasingly clear that some proteases can also possess tumour repressive functions (135, 138). For example, loss of the tumour-suppressive protease, MMP8, has been shown to contribute to DCIS progression by reducing the ability of cells to adhere to the matrix (160).

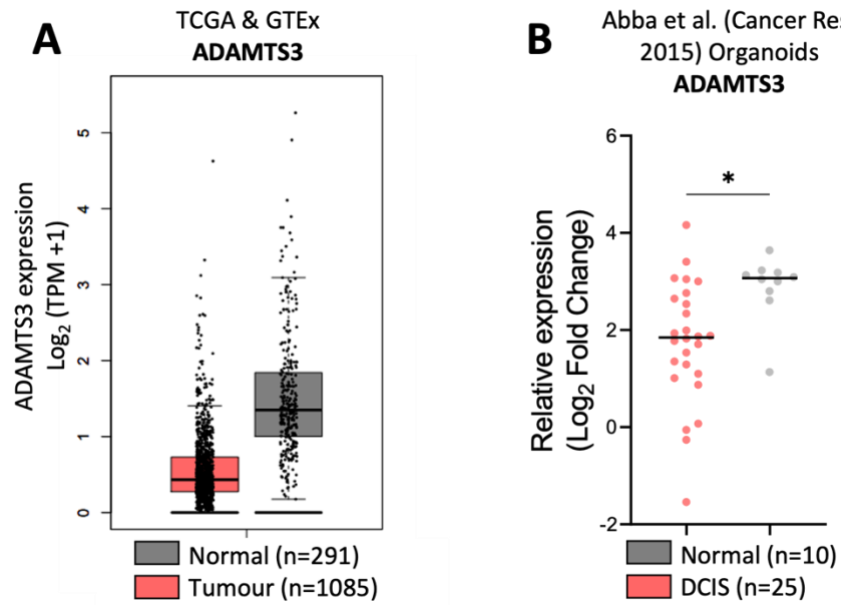
While we have previously shown that integrin  $\beta 6$  expression drives upregulation of MMP13, we also show that metzincins MMP24, ADAMTS3 and ADAM21P1 are instead downregulated (Fig 3.11). ADAMTS3 is a member of the procollagen N-proteinase subfamily, which functions to cleave the N-terminal propeptide of collagen trimers during collagen processing to reduce solubility and facilitate fibril assembly (291). More recently, it has been shown that ADAMTS3 may have broader functions, including the activation of pro-VEGF-C through proteolytic cleavage during lymphangiogenesis (291-293). These findings are supported by other studies where examination of substrates for this family of enzymes identified additional ECM components, with an emphasis on matrix-associated proteins related to TGF $\beta$  signalling (294). This suggests that the substrate repertoire of ADAMTS3 is greater than initially thought, implicating the existence of other functions besides collagen processing.

In this chapter, we assess the role of myoepithelial ADAMTS3 as a potential suppressor of invasion and utilise a degradomic workflow to assess novel targets of ADAMTS3.

## 5.2. Myoepithelial ADAMTS3 loss drives invasion

### 5.2.1. ADAMTS3 expression is lost in DCIS and IDC

Previous expression profiling of ADAMTS members in a breast carcinoma cohort has revealed expression changes in ten ADAMTS proteases, one of which is a downregulation in the procollagen N-proteinase, ADAMTS3 (180). Mining of publicly available TCGA and GTEx cohorts corroborated this finding, demonstrating a reduction of ADAMTS3 in bulk breast tumours compared to healthy breast tissue (Fig 5.1A). To provide a more comprehensive understanding of ADAMTS3-expressing cell types in the healthy breast, we utilised single cell RNAseq data obtained from the Human Protein Atlas, which identified fibroblasts and myoepithelial cells as the main source of ADAMTS3 (Appendix 3.1). As the ADAMTS3-loss observed in breast cancer could be due to the absence of an ADAMTS3-expressing myoepithelial population, we next interrogated publicly available data from a study that performed transcriptomic profiling of DCIS organoids (295). Interestingly, ADAMTS3 loss was observed at the DCIS stage, compared to healthy control organoids, highlighting that ADAMTS3 is already lost during the early stages of disease progression (Fig 5.1B).

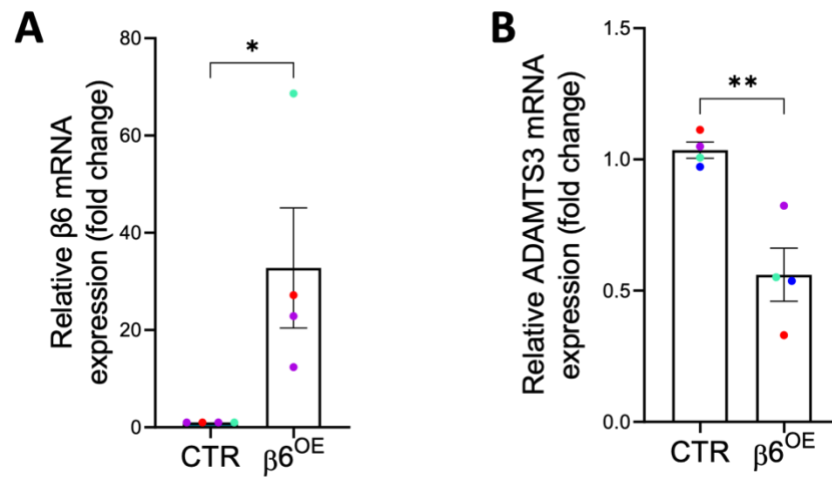


**Figure 5.1 ADAMTS3 loss is observed in DCIS and IDC.**

**(A)** ADAMTS3 mRNA expression ( $\text{Log}_2(\text{TPM} + 1)$ ) in normal (n=291) versus tumour (n=1085) breast tissue. Data extracted from breast cancer TCGA and GTEx data sets. **(B)** ADAMTS3 mRNA expression ( $\text{Log}_2$  fold change) in normal (n=10) versus DCIS (n=25) organoids. Data taken from Abba *et al.* study (GSE69994) (295). \* $p < 0.05$  (Two-tailed t-test).

### 5.2.2. Integrin $\beta 6$ -induction drives ADAMTS3 loss in myoepithelial cells

To investigate the relevance of ADAMTS3 loss in the DCIS-associated myoepithelium, we utilised the previously generated integrin  $\beta 6$ -inducible 1089 myoepithelial cell line to assess changes to ADAMTS3 expression. Upon overexpression of integrin  $\beta 6$  in 1089<sup>i $\beta 6$</sup>  cells (Fig 5.2A), we observed a significant downregulation of ADAMTS3 compared to non  $\beta 6$ -induced control cells (Fig 5.2B), congruent with previous RNAseq data (Fig 3.11). This finding is consistent with other studies that report a downregulation of proteases upon integrin  $\beta 6$  expression, such as loss of myoepithelial MMP8 expression in DCIS (160).



**Figure 5.2 Integrin  $\beta 6$  expression drives ADAMTS3 loss.**

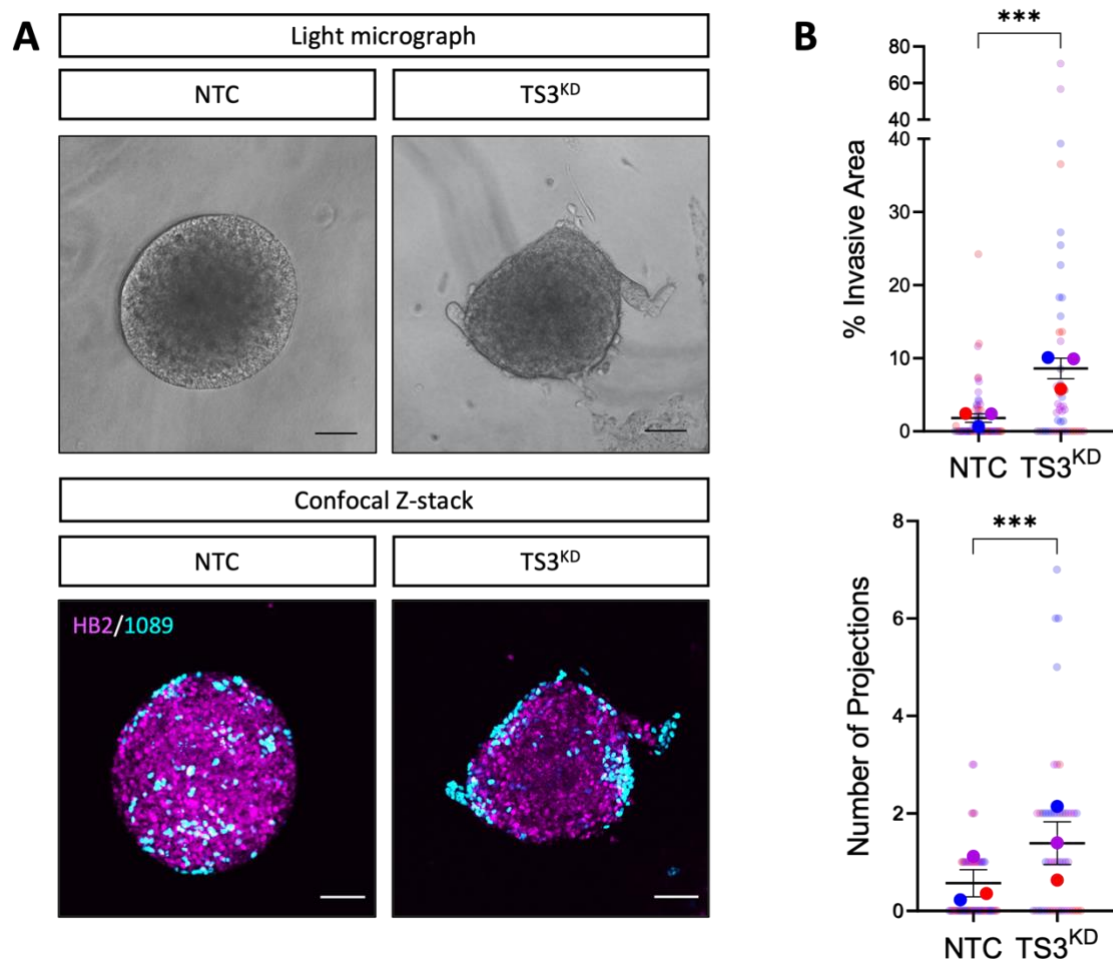
1089 <sup>$\beta 6$</sup>  cells grown in 2D were treated with 1  $\mu\text{g}/\text{ml}$  doxycycline for 48 h to induce  $\beta 6$  induction. Relative mRNA expression of **(A)** integrin  $\beta 6$  and **(B)** ADAMTS3 upon integrin  $\beta 6$  overexpression ( $\beta 6^{OE}$ ) compared to control (CTR) in the 1089 <sup>$\beta 6$</sup>  myoepithelial cell line. \* $p < 0.05$ , \*\* $p < 0.01$  (Two-tailed t-test).



### 5.2.3. ADAMTS3 loss drives myoepithelial-led invasion

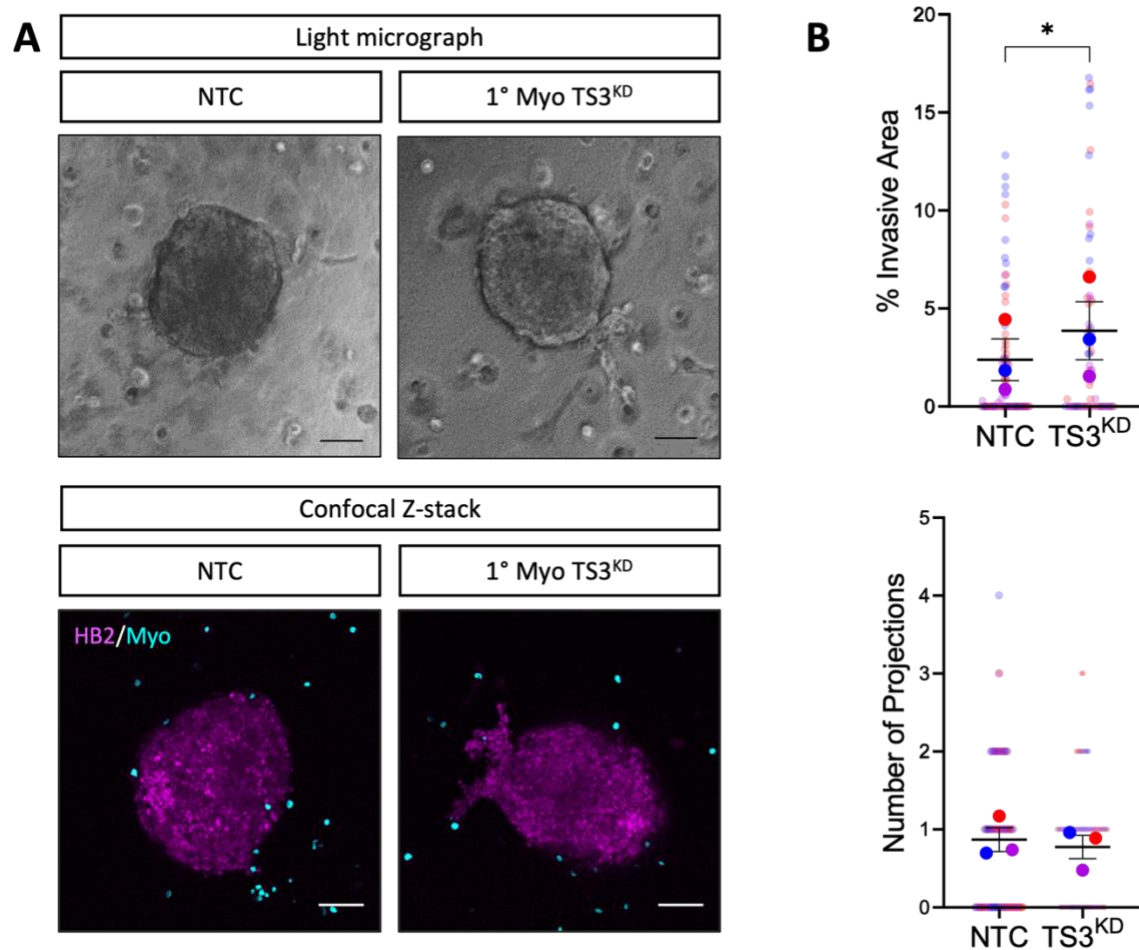
To understand the biological consequence of myoepithelial ADAMTS3 loss on the invasive progression of DCIS, we utilised our previously established HB2/1089 spheroid model (Fig 3.3A). Prior to combination with HB2 cells in heterocellular spheroids, 1089 cells were transfected with non-targeting control or ADAMTS3-targeted siRNA (Appendix 3.2). Strikingly, upon siRNA-mediated knockdown of myoepithelial ADAMTS3 (TS3<sup>KD</sup>) in HB2/1089 spheroids, we observed an increase in invasion, compared to non-targeting control spheroids (Fig 5.3A). Quantification of invading spheroids confirmed a significant increase in the invasive area and the number of invading projections, implicating myoepithelial ADAMTS3 as a suppressor of invasion (Fig 5.3B). Subsequent confocal analysis of invasive protrusions, utilising fluorescently tagged histone constructs, revealed that invading luminal cells were led by myoepithelial cells (Fig 5.3A), in line with previous findings observed upon integrin  $\beta$ 6 induction in this model (Fig 3.5).

These data were recapitulated, although to a lesser extent, in our primary HB2/Myo model consisting of HB2 monoculture spheroids embedded in collagen gels containing primary myoepithelial cells, with knockdown of ADAMTS3 significantly increasing the invasive area (Fig 5.4).



**Figure 5.3 Loss of myoepithelial ADAMTS3 drives invasion in HB2/1089 model.**

Prior to co-culture, 1089 cells were transiently transfected with non-targeting control (NTC) or ADAMTS3 siRNA (TS3<sup>KD</sup>). Transfected 1089 cells were then combined with HB2 cells to form spheres which were embedded in collagen gels. **(A)** Representative light micrographs and fluorescent images (HB2; magenta, 1089; cyan) of spheres at 4 d, with either myoepithelial NTC or TS3<sup>KD</sup>. **(B)** Summary graphs showing % invasive area and number of projections. Data are presented as mean ± SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \*\*\*p < 0.001 (Mann Whitney U test) Scale bar = 100 µm. For this experiment, A. Tan performed two biological replicates, while S. Gibson performed the third biological replicate. Confocal images were captured by S. Gibson.

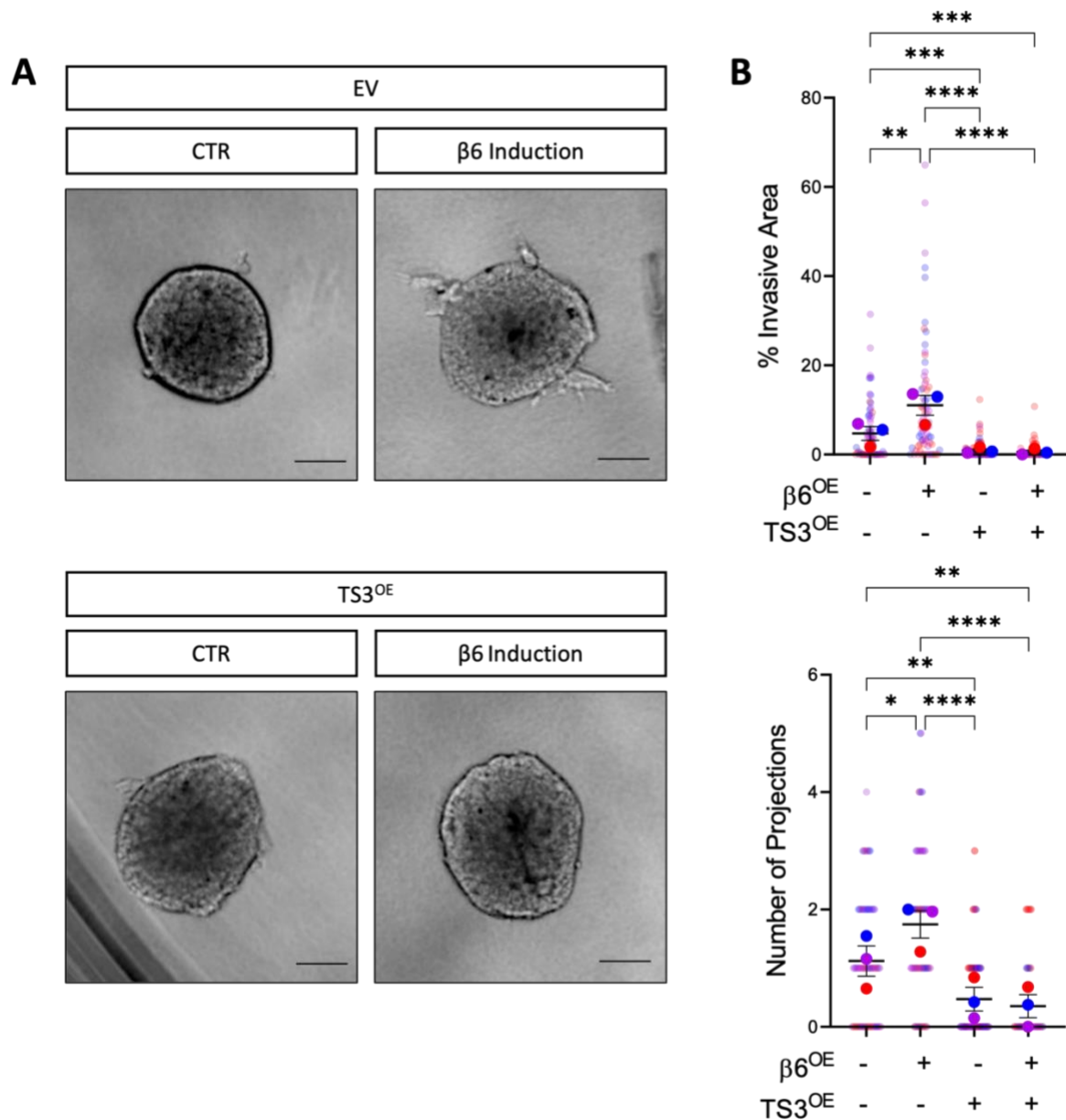


**Figure 5.4 Loss of myoepithelial ADAMTS3 drives invasion in HB2/Myo model.**

Prior to co-culture, primary Myo cells were transiently transfected with non-targeting control (NTC) or ADAMTS3 siRNA (TS3<sup>KD</sup>). HB2 spheres were embedded in Myo-containing collagen gels. **(A)** Representative light micrographs of spheres at 3 d, with either myoepithelial NTC or TS3<sup>KD</sup>. **(B)** Summary graphs showing % invasive area and number of projections. Data are presented as mean ± SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \* $p < 0.05$ , \*\*\* $p < 0.001$  (Mann Whitney U test). Scale bar = 100  $\mu\text{m}$ .

#### 5.2.4. Myoepithelial ADAMTS3 expression inhibits invasion

As loss of myoepithelial ADAMTS3 promoted invasion across both spheroid models, we hypothesised that ADAMTS3 expression may suppress invasion and restrict disease progression. To confirm the suppressive effect of ADAMTS3 on invasion, we designed an ADAMTS3 overexpressing construct and generated ADAMTS3 expressing 1089<sup>iβ6</sup> cells on an integrin β6-inducible background (Appendix 3.3). While integrin β6 overexpression (β6<sup>OE</sup>) drove invasion of non-ADAMTS3 overexpressing spheroids, consistent with our previous findings, simultaneous overexpression of ADAMTS3 (TS3<sup>OE</sup>) abrogated this effect (Fig 5.5A). For example, spheroid analysis revealed a significant decrease in the invasive area and number of projections between dual overexpression of both integrin β6 and ADAMTS3, compared to spheroids with integrin β6<sup>OE</sup> expression alone (Fig 5.5B). Taken together, these data suggest that in the healthy breast, expression of ADAMTS3 may protect ductal integrity, in line with the natural tumour-suppressive function of the myoepithelium (59). However, loss of myoepithelial-ADAMTS3 in DCIS may facilitate a mechanism whereby myoepithelial cells can become altered and drive the invasion of luminal cells into the surrounding stroma.



**Figure 5.5 Myoepithelial expression of ADAMTS3 inhibits integrin  $\beta 6$ -induced invasion**

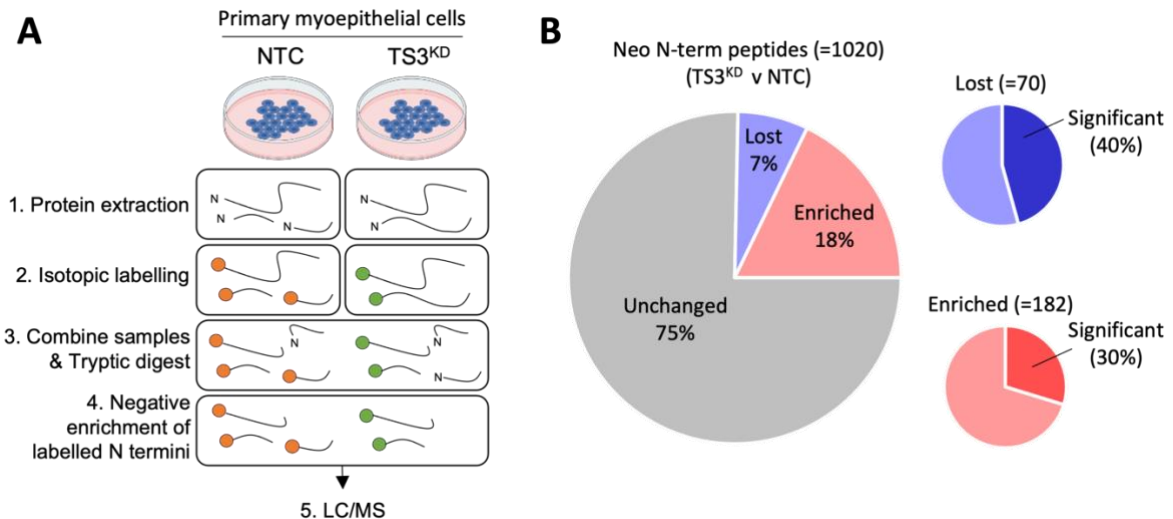
Prior to co-culture, 1089 <sup>$\beta 6$</sup>  cells were transduced with empty vector (EV) or ADAMTS3 overexpressing (TS3<sup>OE</sup>) constructs. Transduced 1089 <sup>$\beta 6$</sup>  cells were then combined with HB2 cells to form spheres, which were embedded in collagen gels and treated with 1  $\mu$ g/ml doxycycline to induce integrin  $\beta 6$  expression. **(A)** Representative light micrographs of spheres at 4 d post treatment, with integrin  $\beta 6$  induction and/or expression of EV or TS3<sup>OE</sup>. **(B)** Summary graphs showing % invasive area and number of projections. Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$  (Kruskal-Wallis test with multiple comparisons). Scale bar = 100  $\mu$ m.

### 5.3. Degradomic analysis reveals ADAMTS3 substrate candidates

#### 5.3.1. Analysis of ADAMTS3 degradome

We next sought to investigate how ADAMTS3 may restrict invasion by interrogating ADAMTS3 substrates. To identify candidate ADAMTS3 substrates, we utilised an N-terminus amine isotopic labelling of substrates (N-TAILS) workflow to analyse changes in the abundance of cleavage events following ADAMTS3 knockdown (Fig 2.3, 5.6A) (256). In this method, N-termini are isotopically labelled prior to the depletion of internal tryptic peptides, allowing for the simultaneous detection of both natural and protease generated neo-N-termini (256). Thus, by identifying neo-N-termini peptides with a reduced abundance upon ADAMTS3 loss, candidate ADAMTS3 substrates can be predicted.

While analysis of 3D cultures would have provided a more physiologically relevant understanding of the degradome, high levels of protein input are required for the workflow and were unattainable from this model. Instead, a 2D approach was utilised and the secretomes of ADAMTS3 siRNA treated-primary myoepithelial cells, were subjected to N-TAILS. Out of the 1020 neo-N-term peptides that were detected, 7% were lost (defined by a reduction in 25% or more) and 18% were enriched (defined by an increase in 25% or more) following ADAMTS3 knockdown compared to the non-targeting control samples (Fig 5.6B).



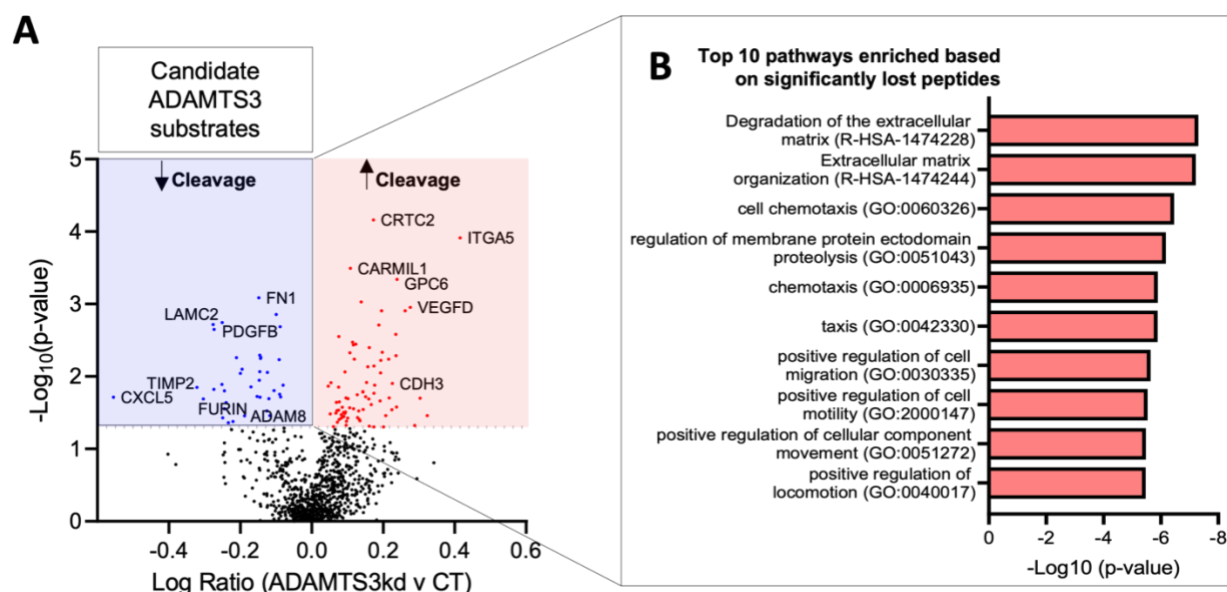
**Figure 5.6 Summary of detected neo-N-term peptides.**

**(A)** Primary myoepithelial cells were transiently transfected with non-targeting control (NTC) or ADAMTS3 siRNA (TS3<sup>KD</sup>). Secretomes were collected 72 h post transfection for N-TAILS. **(B)** Large pie chart shows summary of detected neo-N-term peptides, while smaller pie charts show significantly enriched (red; ratio>1.25,  $p<0.05$ ) or lost (blue; ratio<0.75,  $p<0.05$ ) peptides. For this experiment, protein was extracted by S. Gibson, while E. Madzharova performed the TAILS workflow. Pie charts were generated by S. Gibson.

### 5.3.2. ADAMTS3 substrates are associated with matrix organisation and migration

Out of the peptides that were lost upon ADAMTS3 knockdown, significant hits ( $p < 0.05$ ) included those belonging to parental proteins fibronectin (FN1), laminin-C2 (LAMC2), TIMP metalloproteinase inhibitor 2 (TIMP2) as well as growth factors and additional proteases (Fig 5.7A). Interestingly, increased cleavage of the myoepithelial P-cadherin (CDH3) was also observed following ADAMTS3 loss, congruent with the reported loss of myoepithelial differentiation markers associated with DCIS progression (296, 297). Subsequent analysis of candidate ADAMTS3 substrates revealed an enrichment for pathways associated with matrix organisation, as well as those associated with cell migration and motility (Fig 5.7B). These data concur with the phenotypic observations in our 3D models, suggesting that loss of ADAMTS3 substrate cleavage may positively regulate pathways associated with invasion.





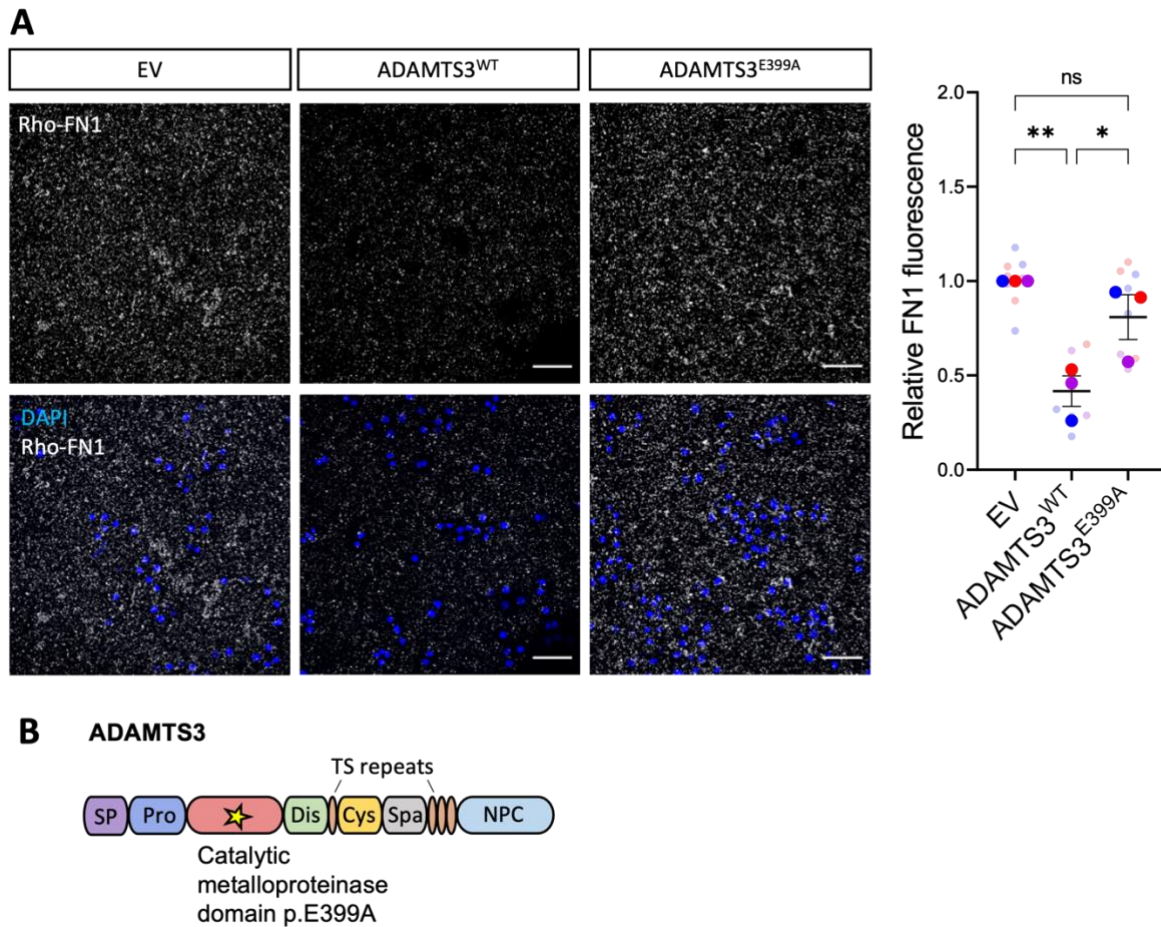
**Figure 5.7** Pathway enrichment analysis of candidate ADAMTS3 substrates.

**(A)** Volcano plot showing parent proteins of neo-N-term peptides enriched (red) or lost (blue) upon ADAMTS3 knockdown. Candidate ADAMTS3 substrates are highlighted in the blue box. **(B)** Candidate ADAMTS3 substrates were subjected to pathway analysis using Metascape software (298). Summary of top 10 enriched pathways based on significantly lost peptides. For this experiment, peptide mapping to parental proteins was performed by Prof. U. auf dem Keller, while S. Gibson generated volcano plots and performed the pathway enrichment analysis.

## 5.4. Proteolytic activity of ADAMTS3 is required for cleavage of fibronectin

### 5.4.1. ADAMTS3 mediates fibronectin cleavage

As the most significant change of the candidate ADAMTS3 substrates was a loss in fibronectin cleavage, we next aimed to validate fibronectin as an ADAMTS3 substrate. To achieve this, 1089 cells with or without ADAMTS3 overexpression were seeded on coverslips coated with rhodamine-labelled fibronectin (299, 300). Fluorescence of fibronectin was then measured as a readout for intact fibronectin, with any fluorescent loss indicative of fibronectin cleavage and degradation (299). Interestingly, 1089 cells overexpressing wild-type ADAMTS3 (ADAMTS3<sup>WT</sup>) displayed approximately 50% less fluorescence compared to 1089 cells transduced with the empty vector (EV) control, indicative of reduced fibronectin levels. These data suggest that at high ADAMTS3 levels, increased cleavage of fibronectin may occur (Fig 5.8A). To then ascertain whether fibronectin cleavage was dependent on the proteolytic activity of ADAMTS3, we designed a construct expressing the enzymatically inactive form of ADAMTS3. To achieve this, we introduced a point mutation in the active site of its catalytic domain, converting glutamic acid at amino acid position 399, which is critical for proteolytic function, to alanine (ADAMTS3<sup>E399A</sup>) (Fig 5.8B) (Appendix 3.3). Fluorescent fibronectin levels remained unchanged in ADAMTS3<sup>E399A</sup> cells compared to the EV condition, confirming that ADAMTS3-mediated fibronectin cleavage is dependent on its proteolytic activity rather than its expression alone (Fig 5.8A).

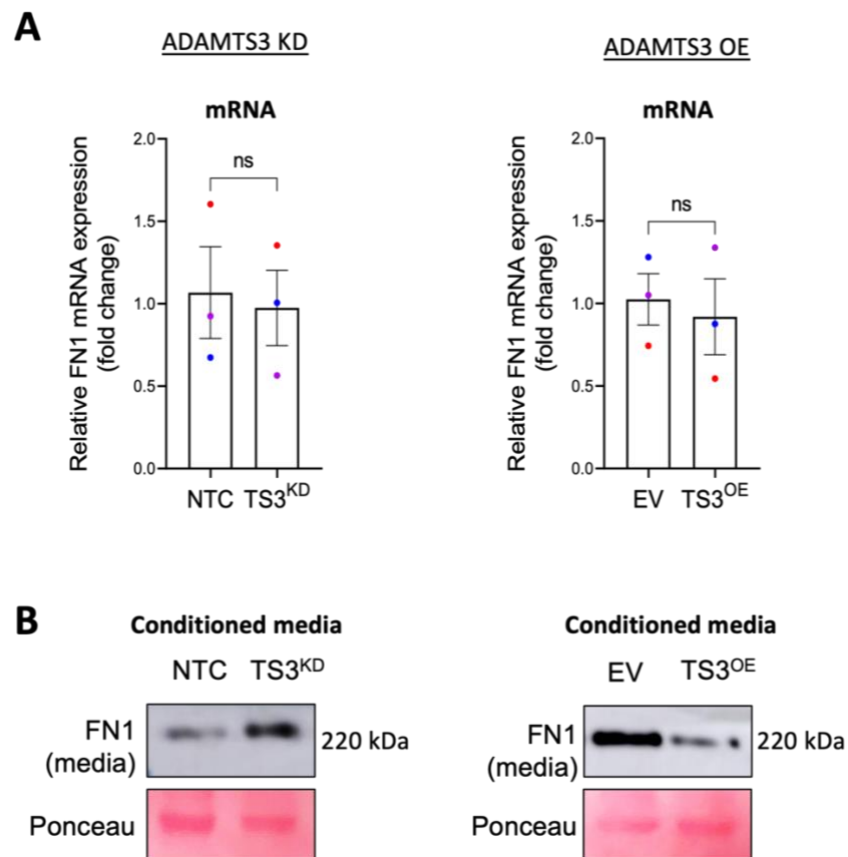


**Figure 5.8** Proteolytic activity of ADAMTS3 is required for cleavage of fibronectin.

1089 cells (EV, ADAMTS3<sup>WT</sup> or ADAMTS3<sup>E399A</sup>) were seeded in 2D on rhodamine-labelled fibronectin (Rho-FN1) coated coverslips for 24 h. **(A)** Representative confocal Z-stack images with DAPI (blue) and Rho-FN1 (white). Quantified relative fluorescence intensity of Rho-FN1. Data are presented as mean  $\pm$  SEM where each dot represents one area of interest with biological replicates indicated by different colours. Average of experimental replicates indicated as larger-sized points. (One-way ANOVA with multiple comparisons). Scale bar = 100  $\mu$ m. **(B)** Schematic of ADAMTS3 structure. Catalytically inactive ADAMTS3 was generated by introducing a point mutation in its active site (p.E399A). SP; signal peptide, Pro; pro-peptide domain, Dis; disintegrin domain, TS; thrombospondin, Cys; cysteine-rich domain, Spa; spacer, NPC; N-terminal pro collagen domain.

#### **5.4.2. ADAMTS3 expression regulates levels of fibronectin in conditioned media**

As high ADAMTS3 levels resulted in increased cleavage and degradation of exogenously introduced fibronectin, we next sought to observe any changes to the levels of endogenous fibronectin in 1089 cultures upon ADAMTS3 knockdown or overexpression. Interestingly, while knockdown or overexpression of ADAMTS3 had no change on fibronectin mRNA expression (Fig 5.9A), differences in the level of fibronectin protein were observed in the conditioned media (Fig 5.9B). Upon siRNA mediated ADAMTS3 loss, levels of fibronectin in cell culture supernatant were increased. However, when ADAMTS3 was instead overexpressed, levels of fibronectin were decreased in the cell culture supernatant (Fig 5.9B). Taken together, these data imply a mechanism whereby ADAMTS3 may restrict levels of fibronectin through increased substrate cleavage and degradation.



**Figure 5.9 ADAMTS3 regulates levels of fibronectin in culture medium.**

For knockdown experiments, 1089 cells were transfected with non-targeting control (NTC) or ADAMTS3 (TS3<sup>KD</sup>) siRNA, and RNA/ medium was harvested at 72 h. For 1089 overexpression experiments, transduced empty vector (EV) or ADAMTS3 overexpressing (TS3<sup>OE</sup>) cells were seeded in 2D, and RNA/medium was harvested at 48 h. **(A)** Relative FN1 mRNA expression upon ADAMTS3 siRNA knockdown (TS3<sup>KD</sup>) or overexpression (TS3<sup>OE</sup>) in 1089 cells. ns = not significant (Two-tailed t-test). **(B)** Western blot showing levels of full-length FN1 present in medium from TS3<sup>KD</sup> or TS3<sup>OE</sup> 1089 cells. Ponceau used as loading control.

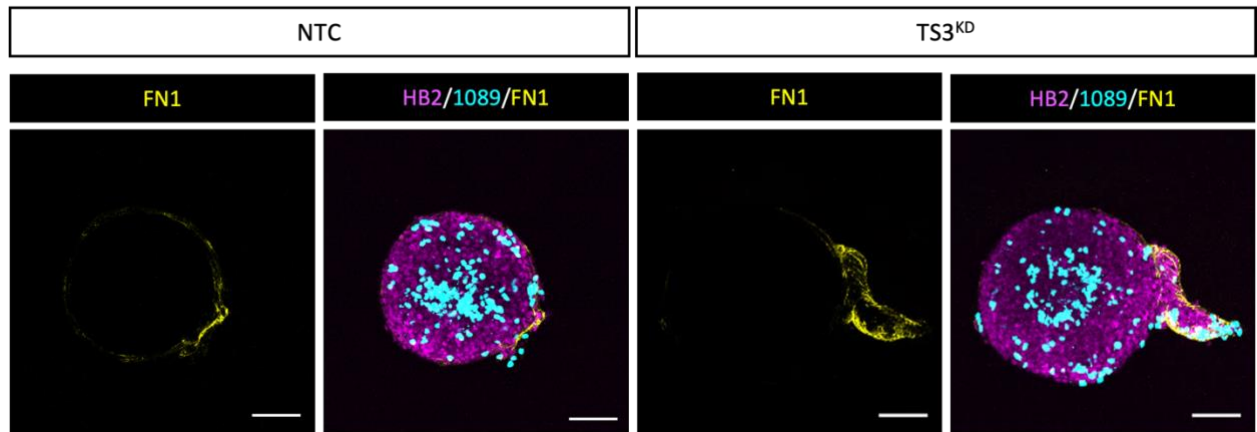
## **5.5. Fibronectin-dependent RGD integrin signalling drives invasion upon ADAMTS3 loss**

### **5.5.1. Increased fibronectin is localised to protrusions of invading spheroids**

Increased fibronectin expression has been shown to favour disease progression, correlating with increased cancer cell invasion and metastasis (128, 202, 301). We therefore sought to investigate if ADAMTS3 dependent regulation of fibronectin was controlling invasion in this context. We first examined fibronectin deposition in our HB2/1089 spheroid assay. Interestingly, immunofluorescent analysis of invading spheroids revealed that fibronectin was localised to their invasive protrusions, suggesting that fibronectin levels may be associated with driving invasion in this model (Fig 5.10).

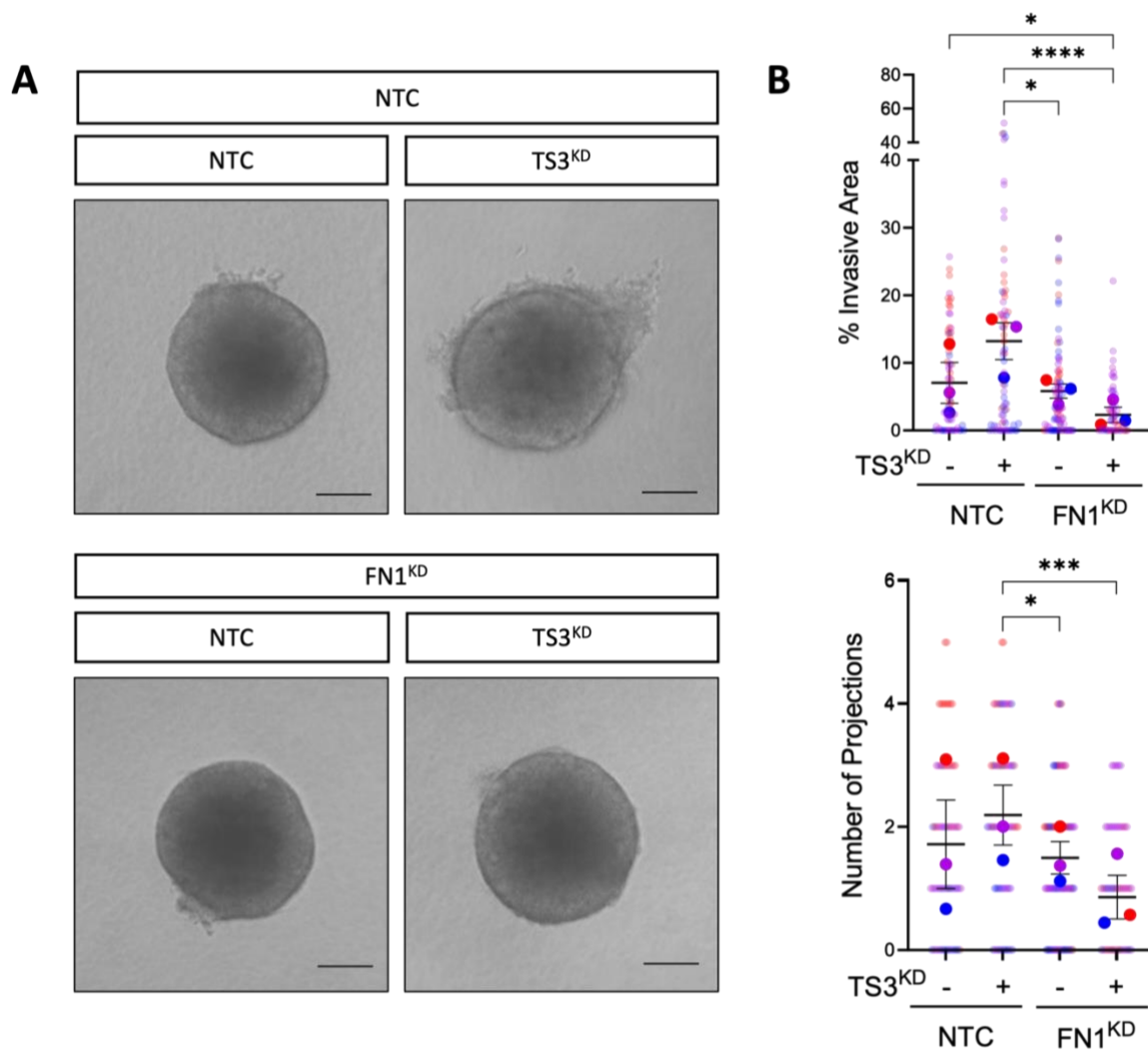
### **5.5.2. Knockdown of myoepithelial fibronectin blocks invasion upon ADAMTS3 loss**

To confirm the invasion-promoting role of fibronectin in our spheroid model, we performed siRNA-mediated knockdown of both fibronectin and ADAMTS3 in 1089 cells prior to their combination with HB2 cells for spheroid embedding (Appendix 3.4). In line with our hypothesis, dual knockdown of both myoepithelial fibronectin and ADAMTS3 mitigated the enhanced invasion observed with ADAMTS3 knockdown alone (Fig 5.11). Equally, supplementing fibronectin into spheroid-containing collagen gels promoted invasion when compared to gels without exogenously added fibronectin, further supporting the invasive role of fibronectin in this model (Fig 5.12).



**Figure 5.10 Fibronectin is localised to invading protrusions.**

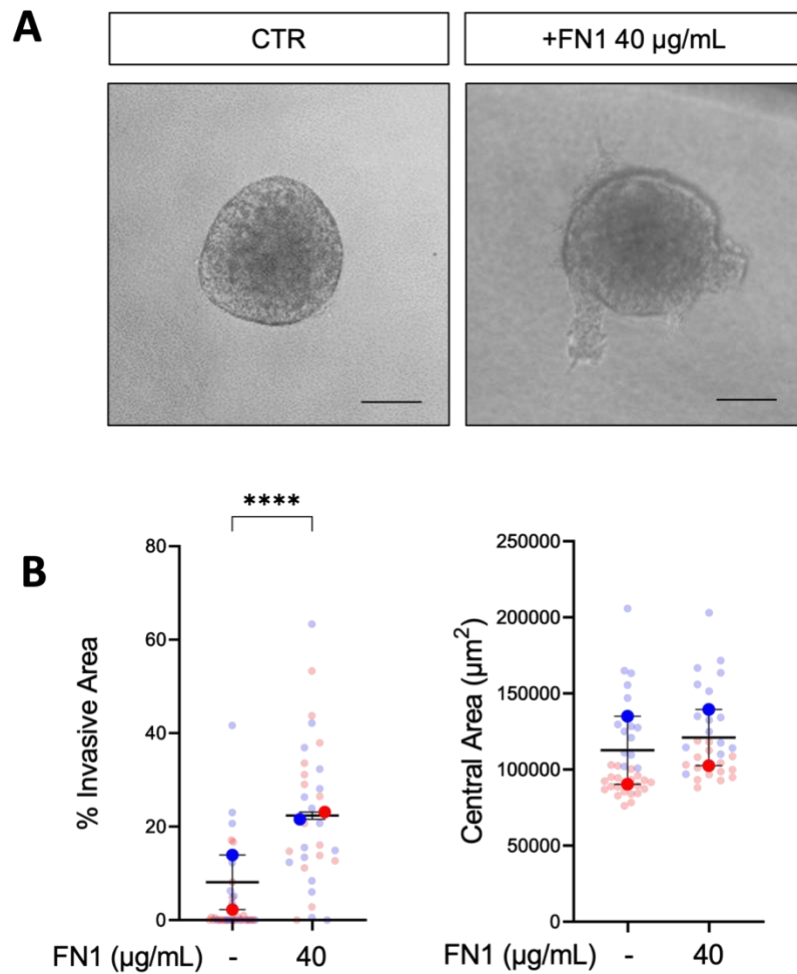
Prior to co-culture, 1089 cells were transiently transfected with non-targeting control (NTC) or ADAMTS3 siRNA (TS3<sup>KD</sup>). Transfected 1089 cells were combined with HB2 cells to form spheres, which were embedded in collagen gels. Representative fluorescence images of fibronectin (FN1) (yellow) in HB2 (magenta)/1089 (cyan) spheroids 4 d post collagen embedding. Scale bar = 100  $\mu$ m.



**Figure 5.11 Loss of myoepithelial fibronectin reduces invasion in HB2/1089 model.**

Prior to co-culture, 1089 cells were transiently transfected with non-targeting control (NTC), ADAMTS3 (TS3<sup>KD</sup>) and/or FN1 (FN1<sup>KD</sup>) siRNA. Transfected 1089 cells were then combined with HB2 cells to form spheres, which were embedded in collagen gels. **(A)** Representative light micrographs of spheres at 4 d, with either myoepithelial NTC, TS3<sup>KD</sup> and/or FN1<sup>KD</sup>. **(B)** Summary graphs showing % invasive area and number of projections. Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \* $p < 0.05$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$  (Kruskal-Wallis test with multiple comparisons) Scale bar = 100  $\mu$ m.





**Figure 5.12 Fibronectin gels drive invasion in HB2/1089 model.**

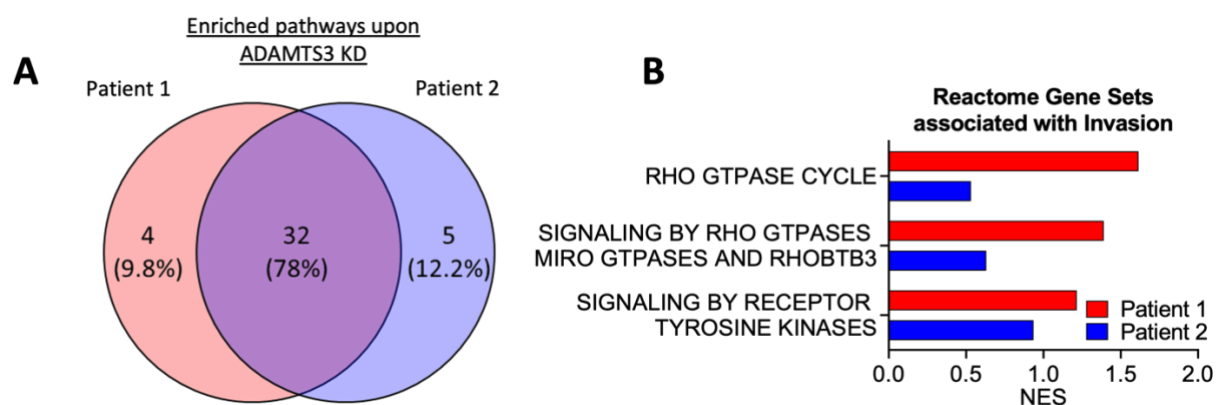
HB2/1089 spheroids were embedded in collagen gels supplemented with 40  $\mu\text{g/mL}$  of exogenous FN1. **(A)** Representative light micrographs of HB2/1089 spheroid cultures at 4 d post embedding. **(B)** Summary graphs showing % invasive area and central area. Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \*\*\*\* $p < 0.0001$  (Mann Whitney U Test). Scale bar = 100  $\mu\text{m}$ .

### 5.5.3. ADAMTS3 loss drives Rho GTPase signalling

Fibronectin can mediate intracellular signalling through integrins to promote cancer cell invasion (231). Having identified that invasion is dependent on loss of ADAMTS3-mediated fibronectin cleavage, we next hypothesised that fibronectin may induce effects through activation of integrins. To study this, we performed quantitative proteomic analysis on protein lysates from the primary myoepithelial cells of two donors following ADAMTS3 knockdown. Similarities in enriched pathways were observed, with 32 pathways identified across both patients (Fig 5.13A). Pathways associated with invasion were upregulated upon ADAMTS3 loss, including Rho GTPase signalling, a major pathway downstream of integrin activation (Fig 5.13B).

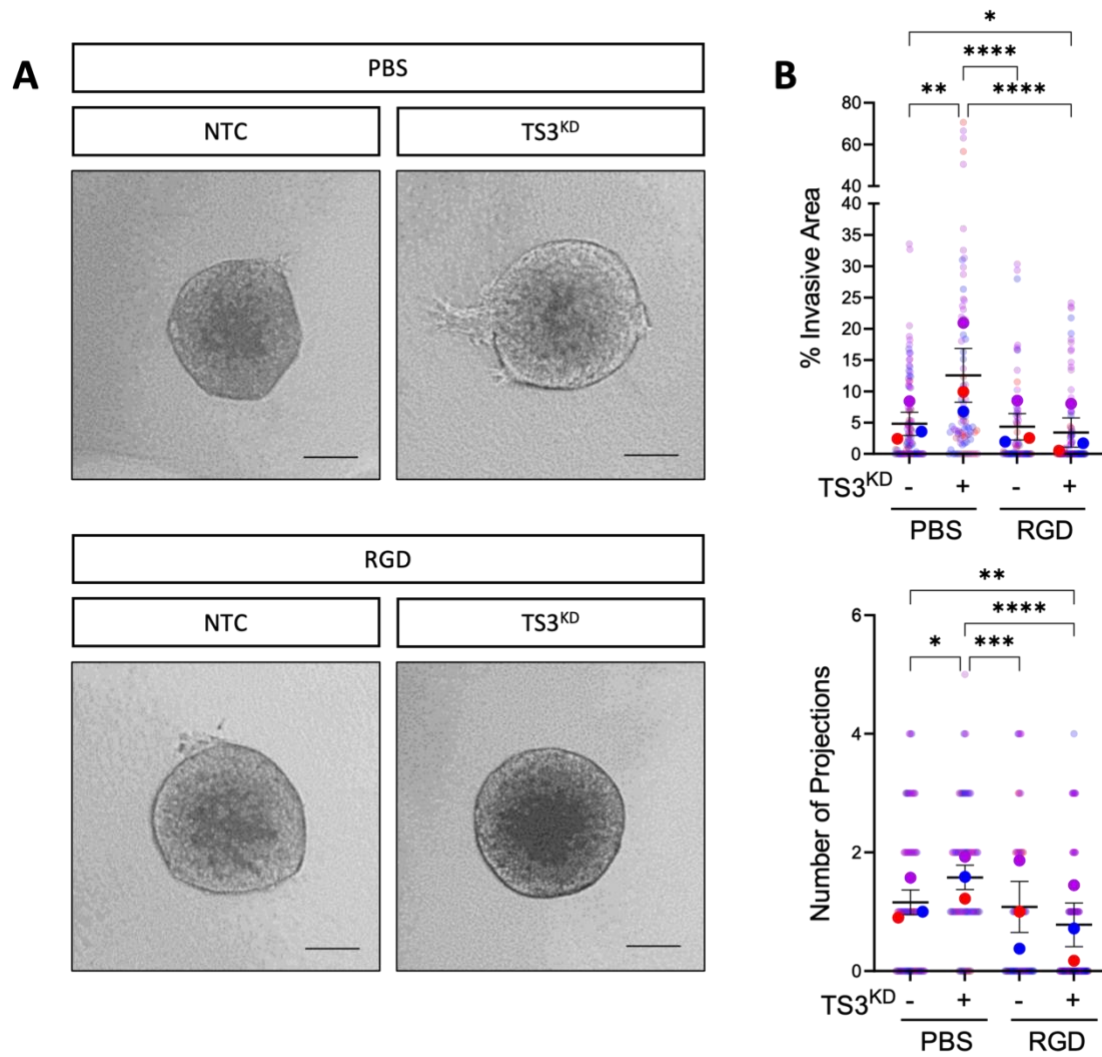
### 5.5.4. Treatment with RGD peptide blocks invasion upon ADAMTS3 loss

As fibronectin is able to mediate outside-in integrin signalling through engagement via its RGD motif, we next tested whether fibronectin was inducing its invasive effects through RGD integrins. To achieve this, we treated ADAMTS3-knockdown spheroids with an RGD peptide to inhibit signalling through integrins  $\alpha v\beta 3$ ,  $\alpha v\beta 5$  and  $\alpha 5\beta 1$  (302). Treatment with the RGD peptide significantly diminished invasion following myoepithelial loss of ADAMTS3 compared to PBS control, confirming the requirement for RGD integrins in facilitating invasion upon ADAMTS3 loss (Fig 5.14).



**Figure 5.13 Analysis of enriched pathways following ADAMTS3 knockdown.**

Primary myoepithelial cells were transfected with non-targeting control or ADAMTS3 siRNA. Protein lysates were extracted 72 h post transfection and subjected to mass spectrometry proteomic analysis. **(A)** Venn diagrams showing the number of overlapping enriched pathways across primary myoepithelial cells isolated from two patients. **(B)** Pathway enrichment analysis from quantitative proteomics data upon ADAMTS3 knockdown in primary myoepithelial cells. NES = normalised enrichment score. For this experiment, peptide mapping to parental proteins was performed by Prof. U. auf dem Keller, while S. Gibson generated pie charts and performed the pathway enrichment analysis.



**Figure 5.14 Inhibition of RGD integrins blocks invasion upon ADAMTS3 loss.**

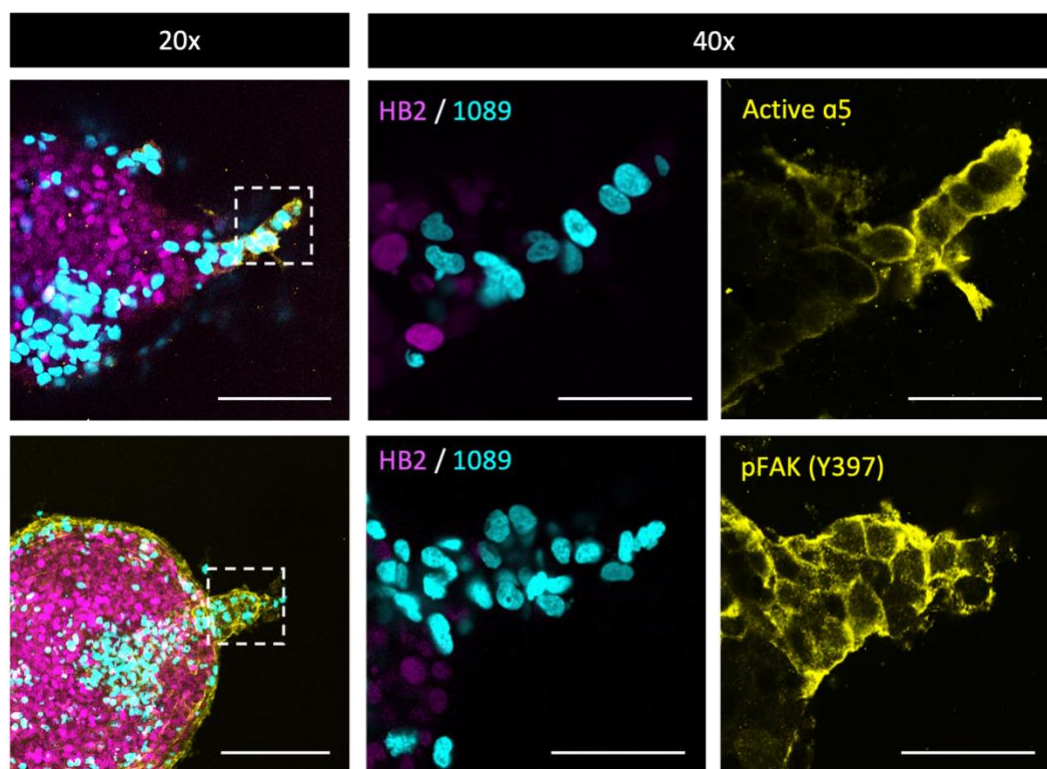
Prior to co-culture, 1089 cells were transiently transfected with non-targeting control (NTC) or ADAMTS3 (TS3<sup>KD</sup>) siRNA. Transfected 1089 cells were then combined with HB2 cells to form spheres which were embedded in collagen gels and treated with either PBS or 10  $\mu$ M of RGD peptide. **(A)** Representative light micrographs of spheres at 4 d, with either myoepithelial NTC or TS3<sup>KD</sup>. **(B)** Summary graphs showing % invasive area and number of projections. Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$  (Kruskal-Wallis test with multiple comparisons) Scale bar = 100  $\mu$ m.

#### **5.5.5. Invading spheroids display increased integrin $\alpha 5\beta 1$ and FAK activation**

Analysis of our myoepithelial RNAseq data (GSE224401) suggests that 1089 myoepithelial cells express the canonical fibronectin receptor, integrin  $\alpha 5\beta 1$  (Appendix 3.5). Interestingly, degradomic analysis of enriched peptides following ADAMTS3 loss included a significant increase in the integrin  $\alpha 5$  peptide neo-N-termini (Fig 5.7A). As integrin  $\alpha 5\beta 1$  requires post-translational endoproteolytic cleavage for integrin  $\beta 1$  association (303-305), this suggested that enhanced levels of active integrin  $\alpha 5$  were being incorporated into the membrane upon ADAMTS3 loss. To confirm this hypothesis, we visualised activated integrin  $\alpha 5\beta 1$  by immunofluorescence, using a SNAKA51  $\alpha 5$  antibody clone (306). This revealed an increase in the level of activated integrin  $\alpha 5\beta 1$ , with subsequent phosphorylation of FAK at Y397, localised to the invasive front of invading spheroids (Fig 5.15).

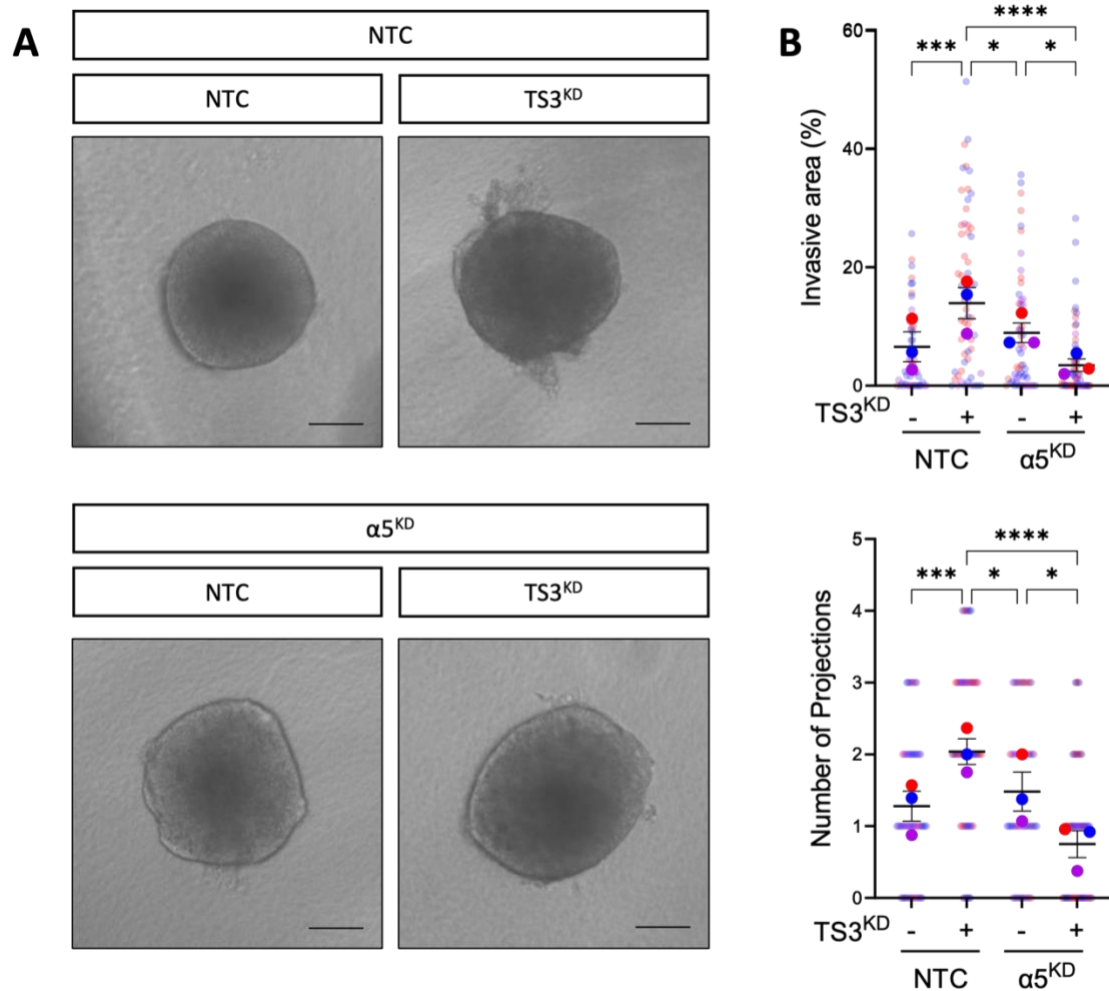
#### **5.5.6. Knockdown of myoepithelial-integrin $\alpha 5$ blocks invasion upon ADAMTS3 loss**

Finally, to confirm the specific requirement of integrin  $\alpha 5$  for invasion, we performed dual knockdown of both myoepithelial ADAMTS3 and integrin  $\alpha 5$ , confirmed by qPCR (Appendix 3.6), and assessed spheroid invasion. Upon simultaneous integrin  $\alpha 5$  and ADAMTS3 knockdown, the percentage of invasive area was significantly reduced, compared to knockdown of ADAMTS3 alone, confirming that invasion mediated through ADAMTS3-loss requires integrin  $\alpha 5$  (Fig 5.16). Taken together, these data suggest that ADAMTS3 loss drives myoepithelial-led invasion, through activation of the fibronectin/integrin  $\alpha 5$  axis.



**Figure 5.15 Enhanced integrin  $\alpha 5$  activation is observed following ADAMTS3 knockdown.**

Prior to co-culture, 1089 cells were transiently transfected with ADAMTS3 (TS3<sup>KD</sup>) siRNA. Transfected 1089 cells were then combined with HB2 cells to form spheres which were embedded in collagen gels. Representative fluorescence images of invading TS3<sup>KD</sup> spheroids at 4 d post treatment with HB2 (magenta) 1089 (cyan) and active integrin  $\alpha 5$  or pFAK (Y397) (yellow). Scale bar = 100  $\mu\text{m}$  (or 50  $\mu\text{m}$  for inserts).



**Figure 5.16 Integrin  $\alpha 5$  is required for invasion upon ADAMTS3 loss.**

Prior to co-culture, 1089 cells were transiently transfected with non-targeting control (NTC), ADAMTS3 (TS3<sup>KD</sup>) and/or integrin  $\alpha 5$  ( $\alpha 5^{\text{KD}}$ ) siRNA. Transfected 1089 cells were then combined with HB2 cells to form spheres which were embedded in collagen gels. **(A)** Representative light micrographs of spheres at 4 d, with either myoepithelial NTC, TS3<sup>KD</sup> and/or  $\alpha 5^{\text{KD}}$ . **(B)** Summary graphs showing % invasive area and number of projections. Data are presented as mean  $\pm$  SEM where each dot represents one sphere with biological replicates indicated by different colours. Average of biological replicates indicated as larger-sized points. \* $p < 0.05$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$  (Kruskal-Wallis test with multiple comparisons) Scale bar = 100  $\mu\text{m}$ .

## 5.6. Discussion

### Results summary

- ADAMTS3 expression is lost following upregulation of integrin  $\beta 6$  in myoepithelial cells.
- Loss of ADAMTS3 drives myoepithelial-led invasion in spheroid models, while ADAMTS3 overexpression has an inhibitory effect.
- Degradomic analysis implicates fibronectin as a novel ADAMTS3 substrate.
- ADAMTS3 knockdown enhances fibronectin levels, which subsequently activates signalling through the integrin  $\alpha 5 \beta 1$  signalling axis to promote invasion.

While the progression of DCIS to IDC is a critical step in cancer progression, the molecular mechanisms which underpin this transition remain unclear. We and others have shown that upregulation of integrin  $\alpha \nu \beta 6$ , a myoepithelial marker of high-risk DCIS, can promote DCIS progression through TGF $\beta$ -dependent upregulation of MMPs 9 and 13 (72, 127). While metzincins are known for their tumour-promoting roles, it is becoming increasingly recognised that they are in fact multifaceted proteases, with some metzincins behaving as natural tumour-suppressors. In this chapter, we show that ADAMTS3 acts as a tumour suppressor through the regulation of fibronectin and, by extension, integrin mediated invasion. Loss of ADAMTS3 in a pro-tumoural context, such as myoepithelial expression of the integrin  $\beta 6$ , thus represents a decisive step in disease progression by affecting matrix composition (Fig 5.17).

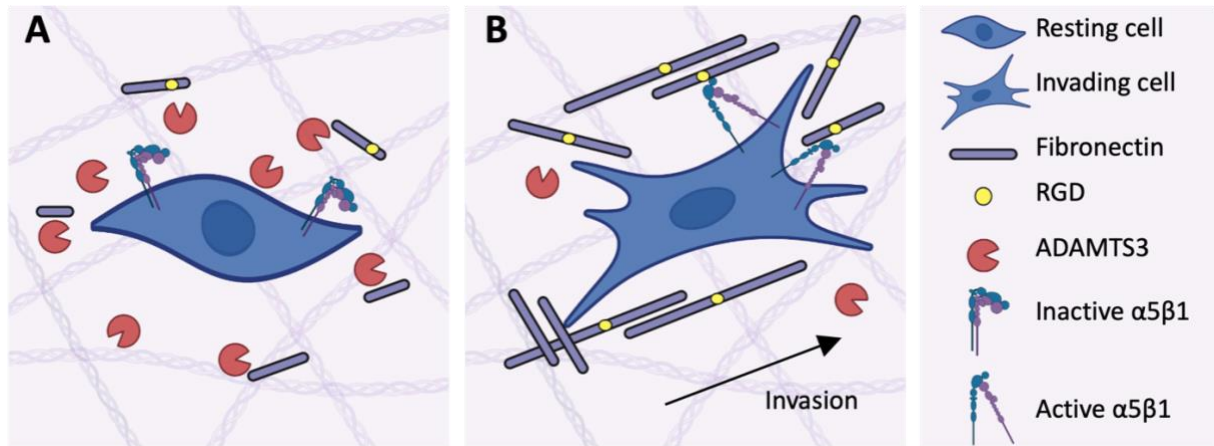
In this chapter, we posit ADAMTS3 as a tumour-suppressive protease, where knockdown of myoepithelial ADAMTS3 in heterocellular spheroids promotes myoepithelial-led invasion of luminal cells into their surrounding matrix. As a procollagen N-proteinase, ADAMTS3 is known principally for its ability to cleave the propeptide of fibrillar collagen precursors to facilitate collagen processing (291). However, recent studies indicate that ADAMTS3 may also possess functions unrelated to



collagen biology (291), with one example including cleavage and activation of pro-VEGF-C to facilitate VEGFR signalling during lymphangiogenesis (292, 307). These findings are supported by other studies where examination of novel procollagen N-proteinase (ADAMTS2, 3 and 14) substrates identified additional extracellular matrix components, with an emphasis on matrix-associated proteins related to TGF $\beta$  signalling (294).

Here, we confirm fibronectin to be a ADAMTS3 substrate, with increased fibronectin cleavage observed upon expression of wildtype ADAMTS3, but crucially unaffected with a protease-inactive point mutant. We also show that increased fibronectin upon ADAMTS3 loss promotes myoepithelial-led invasion, which can be abrogated upon fibronectin depletion. Interestingly, in our degradomic analysis, the most significantly enriched peptide upon ADAMTS3 loss belonged to the fibronectin receptor subunit, integrin  $\alpha 5$ . As integrin  $\alpha 5$  requires post-translational endoproteolytic cleavage for  $\beta 1$  integrin association (303-305), this suggested that enhanced levels of integrin  $\alpha 5$  were being incorporated into the membrane to increase transmission of extracellular signals. This was confirmed by immunofluorescence, which showed activation of integrin  $\alpha 5$  and phosphorylation of the key effector kinase FAK (Y397) at the invasive front of invading spheroids, with subsequent integrin  $\alpha 5$  knockdown blocking invasion. Our findings highlighting the tumour-promoting role of fibronectin can be supported by CAF studies, where production of fibronectin-rich matrices promote directional cancer-cell migration through RGD integrins,  $\alpha 5\beta 1$  and  $\alpha v\beta 3$  (128, 202).

In summary, we present a mechanism where integrin  $\beta 6$  expressing myoepithelial cells downregulate ADAMTS3 to reduce fibronectin cleavage. This subsequently drives increased fibronectin/integrin  $\alpha 5\beta 1$  signalling, to mediate myoepithelial-led invasion of luminal cells into the matrix (Fig 5.17). Our findings highlight the complexity of protease activity and call attention to identifying additional gatekeeper proteases that suppress disease progression.



**Figure 5.17 Proposed tumour-suppressive mechanism of myoepithelial ADAMTS3.**

**(A)** In healthy myoepithelial cells, where ADAMTS3 is expressed, fibronectin can be cleaved and degraded, resulting in less activation of integrin  $\alpha 5 \beta 1$ . **(B)** In DCIS, myoepithelial-ADAMTS3 expression is reduced, resulting in less cleavage and degradation of fibronectin. Increased levels of fibronectin mediate signalling through integrin  $\alpha 5 \beta 1$  and promote myoepithelial-led invasion of luminal cells.

## **Chapter 6. Final Discussion**

With reports describing DCIS tumour cells to be genetically indistinct from their invasive counterparts, attention has recently shifted to the role of the DCIS microenvironment in facilitating invasive progression (53, 55-58). With loss of the myoepithelial layer presenting as a key pathological determinant in disease progression, there is extensive interest in understanding the fate of myoepithelial cells and their role in modulating the transition of DCIS to IDC.

While 3D culture systems provide an advantageous platform to study cancer progression, there is a paucity of models that enable cellular interactions to be examined in the context of DCIS. In this study, we address this need by presenting several 3D *in vitro* models that incorporate both luminal and myoepithelial cell compartments into physiologically relevant matrix (245). Our developed heterocellular spheroids deposit a BM, which they subsequently degrade and invade through, providing functional insights into luminal-myoepithelial crosstalk, as well as matrix interactions during DCIS progression (245). Using these models, we present a mechanism where integrin  $\beta 6$ -expressing myoepithelial cells adopt a tumour-promoting phenotype in DCIS. Upon upregulation of integrin  $\beta 6$ , expression of myoepithelial metzincins, MMP13 and ADAMTS3, become altered to facilitate matrix remodelling and luminal cell invasion. While TGF $\beta$ /EP300-dependent upregulation of MMP13 drives enhanced collagen proteolysis, ADAMTS3 loss enhances invasion through increasing fibronectin levels, which activate integrin  $\alpha 5\beta 1$  on myoepithelial cells.

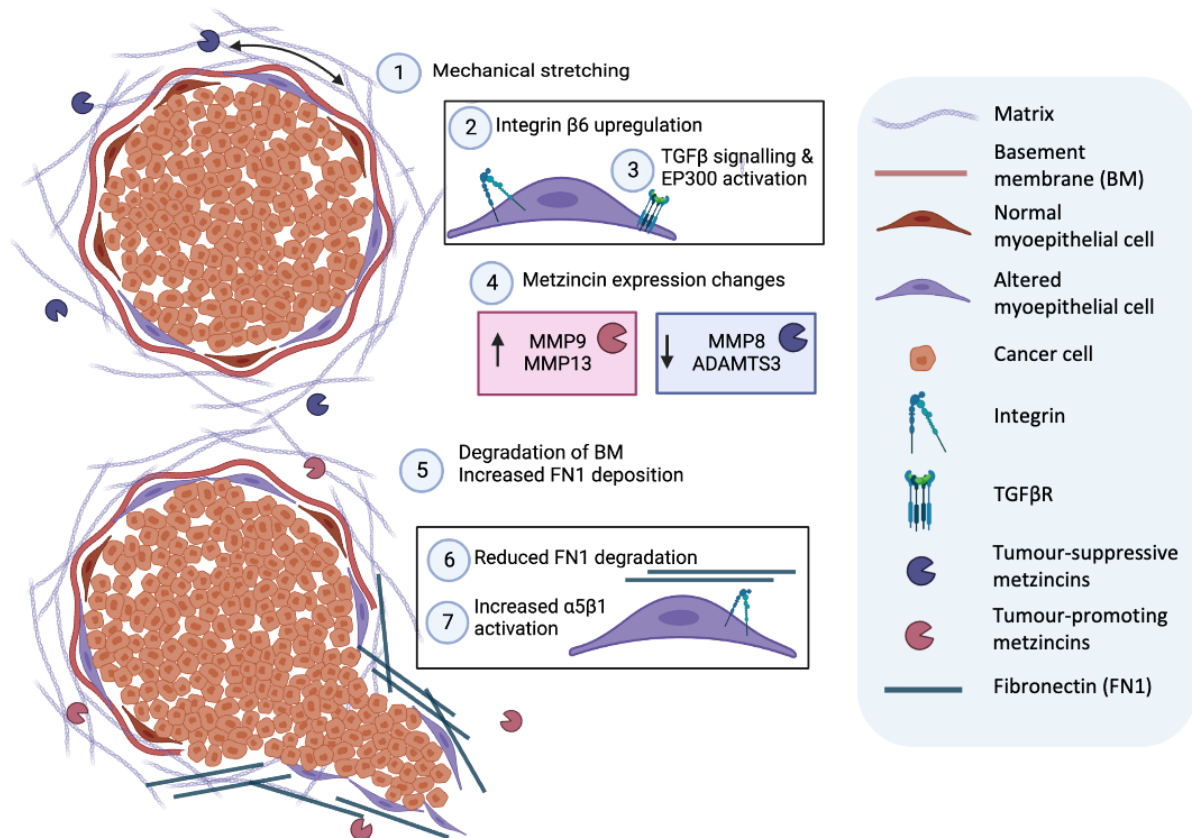
Importantly, expression of myoepithelial MMP13 was observed in clinical high-grade cases of DCIS, consistent with our *in vitro* data implicating MMP13 as an active driver of invasion. However, while we did not test for reduced expression of myoepithelial ADAMTS3 expression in our cohort of DCIS cases, our data are consistent with clinical findings from other studies, where increased periductal fibronectin is observed in high-grade DCIS cases (127), and may contribute to their faster progression to invasive disease (308). Interestingly, a positive correlation between myoepithelial integrin  $\beta 6$

expression and periductal fibronectin has also been reported in the analysis of matched DCIS ducts (127). While integrin  $\beta 6$ -expressing myoepithelial cells increase their expression of fibronectin (127), we present a complementary mechanism where reduced ADAMTS3-mediated cleavage can also maintain levels of fibronectin in the stroma, to facilitate invasive progression.

Our data are also consistent with other studies where stable expression of integrin  $\beta 6$  in myoepithelial cells promote invasion through upregulation of tumour-promoting MMP9 (72, 127), and downregulation of tumour-suppressive MMP8 (160). Thus, data from our study, combined with previous studies, provides a working hypothesis for a mechanism where myoepithelial cells drive the progression of DCIS to IDC through altered metzincin expression (Fig 6.1). First, uncontrolled proliferation of neoplastic cells results in filling and expansion of the breast duct, causing mechanical stretching of surrounding myoepithelial cells (127). Through mechanostimulation, myoepithelial cells then upregulate integrin  $\beta 6$  expression (127), which increases activation of latent TGF $\beta$  and subsequent signalling through both canonical and non-canonical TGF $\beta$  pathways (72, 127). This results in a positive feedback loop, promoting further TGF $\beta$  signalling which drives upregulation of MMPs 9 and 13. In the case of MMP13, this is effected epigenetically, through TGF $\beta$ -dependent activation of the histone acetyl transferase, EP300. Upon secretion, these proteases facilitate invasion via degradation of the BM and surrounding collagen matrix. In addition to activating metzincins, enhanced TGF $\beta$  signalling also promotes expression and localisation of myoepithelial-derived fibronectin to the periductal regions (127), further enhancing an invasion-permissive environment (Fig 6.1).

In contrast to the upregulation of tumour-promoting metzincins, integrin  $\beta 6$  expression also drives a loss in those that are tumour-suppressive, such as MMP8 and ADAMTS3. As MMP8 acts as a tumour suppressor by promoting hemidesmosome formation via inhibition of MMP9 activity, its loss further

contributes to inducing invasion (160). In the case of ADAMTS3, its loss maintains levels of fibronectin, through reduced ADAMTS3-mediated cleavage and degradation. Increased levels of fibronectin subsequently promote signalling through integrin  $\alpha 5\beta 1$  and Rho GTPases to initiate additional activation of pathways associated with invasion and migration (Fig 6.1).



**Figure 6.1** Proposed tumour-promoting mechanism of myoepithelial cells in DCIS progression.

**1)** Mechanostimulation of the myoepithelium upon ductal filling drives **2)** upregulation of integrin  $\beta 6$ , which facilitates **3)** increased TGF $\beta$  signalling. **4)** This results in changes to metzincin expression, with an upregulation of tumour-promoting metzincins (red) and a downregulation of tumour-suppressive metzincins (blue). **5)** Increased proteolysis results in degradation of the basement membrane (BM), while increased fibronectin (FN1) deposition and **6)** its reduced cleavage and degradation **7)** increases signalling through the integrin  $\alpha 5\beta 1$  axis.

### Limitations and future work

While our results provide insight into the mechanisms that underpin the myoepithelial invasion-promoting phenotype, our study also presents limitations. For example, while we describe the tumour-promoting fate of myoepithelial cells during the initial transition to invasion, how they are lost at the invasive stage remains to be elucidated. Interestingly, during disease progression, myoepithelial cells progressively lose their defining markers. For example, in mouse models of DCIS, p63, calponin,  $\alpha$ SMA, are progressively lost and have been shown to act as respective early, intermediate and late indicators of a compromised myoepithelium (296). This can be further supported by our degradomics data following ADAMTS3 siRNA-mediated knockdown, which identified increased cleavage of the myoepithelial marker P-cadherin, in an invasive setting. This suggests that, in invasion, myoepithelial cells adopt a phenotype similar to that of CAFs, rendering them currently indistinguishable from activated stromal fibroblasts.

Secondly, while it is clear that integrin  $\beta$ 6-dependent upregulation of MMP13 occurs via increased TGF $\beta$  signalling, the mechanism whereby ADAMTS3 becomes lost in DCIS-associated myoepithelial cells is unclear. One potential hypothesis is that integrin  $\beta$ 6-dependent upregulation of fibronectin (127), an ADAMTS3 substrate, initiates a negative feedback loop to downregulate myoepithelial ADAMTS3 expression, however additional investigation is required to confirm this. Regarding its transcriptional regulation, previous osteosarcoma studies suggest that ADAMTS3 expression can be dampened via activation of the transcription factor SP1 (309). Interestingly, SP1 was reported to interact with EP300 directly in multiple myeloma (310), and drive TGF $\beta$ -dependent activation of EMT in pancreatic cancer cells (311).

Finally, we were unable to determine whether expression of MMP13 and ADAMTS3 could act as prognostic biomarkers of DCIS progression or recurrence. Thus, future work would consist of designing



a myoepithelial biomarker validation panel (Table 6.1), where relevance as prognostic markers could be validated in patient cohorts. The UK/ANZ study could provide a potential cohort for biomarker validation, as it consists of a unique series of DCIS with long term follow up data. In this study, 1694 women were randomly assigned to four treatment arms; tamoxifen, radiotherapy alone, tamoxifen alone or no adjuvant treatment, and the endpoint assessed was disease recurrence, either as DCIS or IDC. Using this cohort, association between expression of the designed biomarker panel and recurrence would be calculated. This could involve initial testing for their prognostic significance as individual biomarkers, prior to combination as a refined panel. This has been successfully implemented as an approach for the validation of integrin  $\beta 6$  expression as a marker of disease recurrence (72, 206).

**Table 6.1 Myoepithelial biomarker panel for validation.**

Markers positively associated with recurrence	Markers negatively associated with recurrence
Integrin $\beta$ 6	MMP8
MMP9	ADAMTS3
MMP13	
EP300	
Fibronectin	

**Concluding remarks**

With the microenvironment being increasingly recognised as a mediator of disease progression, it is likely that, in addition to conventional histopathological factors, tumour-extrinsic markers will also be required, to develop effective treatment plans. Ultimately, further understanding of the microenvironmental drivers that underpin DCIS progression will be key in orchestrating the successful translation of putative prognostic markers into clinical biomarker studies.

First, to investigate mechanisms of DCIS *in vitro*, careful selection of model systems is necessary to accurately reflect the physiological state of the disease, with complex 3D models presenting an ideal platform for these studies (277). Secondly, *in vivo* models, such as the MIND model (312), will offer enhanced biological relevance compared with standard mammary fat pad models, which may not be reflective of the true microenvironment in DCIS. Finally, integrating extensively interrogated preclinical data with clinical outcome data from active surveillance trials for DCIS can form the foundation for subsequent biomarker studies (43-46). Ultimately, these approaches will allow for a comprehensive understanding of the biological mechanisms involved and facilitate the development of non-invasive tools to differentiate between indolent and progressive DCIS. This differentiation will be crucial in addressing the issue of overtreatment, where some cases of DCIS may not require aggressive interventions.

## References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, et al. Global cancer statistics. *CA Cancer J Clin.* 2011;61(2):69-90.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424.
3. CancerResearchUK. CRUK Breast cancer statistics [Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer>.
4. CancerResearchUK. CRUK Breast cancer risk factors [Available from: <https://www.cancerresearchuk.org/about-cancer/breast-cancer/risks-causes/risk-factors>.
5. Statistics OfN. Cancer survival in England [Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalinengland/stageatdiagnosisandchildhoodpatientsfollowedupto2018>.
6. Rønnev-Jessen L, Petersen OW, Bissell MJ. Cellular changes involved in conversion of normal to malignant breast: importance of the stromal reaction. *Physiol Rev.* 1996;76(1):69-125.
7. Gusterson BA, Warburton MJ, Mitchell D, Ellison M, Neville AM, et al. Distribution of myoepithelial cells and basement membrane proteins in the normal breast and in benign and malignant breast diseases. *Cancer Res.* 1982;42(11):4763-70.
8. Hu M, Yao J, Carroll DK, Weremowicz S, Chen H, et al. Regulation of in situ to invasive breast carcinoma transition. *Cancer cell.* 2008;13(5):394-406.
9. Allred DC, Mohsin SK, Fuqua SA. Histological and biological evolution of human premalignant breast disease. *Endocr Relat Cancer.* 2001;8(1):47-61.
10. Bombonati A, Sgroi DC. The molecular pathology of breast cancer progression. *J Pathol.* 2011;223(2):307-17.
11. Mancino AT. Diseases of the Breast. *Ann Surg.* 2001;233(4):594.
12. Wärnberg F, Casalini P, Nordgren H, Bergkvist L, Holmberg L, et al. Ductal carcinoma in situ of the breast: a new phenotype classification system and its relation to prognosis. *Breast Cancer Res Treat.* 2002;73(3):215-21.
13. CancerResearchUK. About breast cancer staging and grades [Available from: <https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/about-breast-cancer-staging-grades>.
14. Guedj M, Marisa L, de Reynies A, Orsetti B, Schiappa R, et al. A refined molecular taxonomy of breast cancer. *Oncogene.* 2012;31(9):1196-206.
15. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, et al. Comprehensive molecular portraits of human breast tumours. *Nature.* 2012;490(7418):61-70.
16. Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO, et al. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol.* 2010;28(10):1684-91.
17. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A.* 2001;98(19):10869-74.
18. Dai X, Li T, Bai Z, Yang Y, Liu X, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. *Am J Cancer Res.* 2015;5(10):2929-43.
19. CancerResearchUK. In situ breast carcinoma diagnosis and treatment statistic [Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/in-situ-breast-carcinoma-diagnosis-and-treatment>.
20. Wapnir IL, Dignam JJ, Fisher B, Mamounas EP, Anderson SJ, et al. Long-Term Outcomes of Invasive Ipsilateral Breast Tumor Recurrences After Lumpectomy in NSABP B-17 and B-24 Randomized Clinical Trials for DCIS. *JNCI: Journal of the National Cancer Institute.* 2011;103(6):478-88.

21. Fisher B, Dignam J, Wolmark N, Mamounas E, Costantino J, et al. Lumpectomy and radiation therapy for the treatment of intraductal breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-17. *J Clin Oncol*. 1998;16(2):441-52.
22. Lari SA, Kuerer HM. Biological Markers in DCIS and Risk of Breast Recurrence: A Systematic Review. *J Cancer*. 2011;2:232-61.
23. Nelson NJ. DCIS Prognostic Markers: A Few New Candidates Emerge. *JNCI: Journal of the National Cancer Institute*. 2010;102(9):588-90.
24. Casasent AK, Almekinders MM, Mulder C, Bhattacharjee P, Collyar D, et al. Learning to distinguish progressive and non-progressive ductal carcinoma in situ. *Nature Reviews Cancer*. 2022;22(12):663-78.
25. Ryser MD, Weaver DL, Zhao F, Worni M, Grimm LJ, et al. Cancer Outcomes in DCIS Patients Without Locoregional Treatment. *JNCI: Journal of the National Cancer Institute*. 2019;111(9):952-60.
26. Elshof LE, Schaapveld M, Schmidt MK, Rutgers EJ, van Leeuwen FE, et al. Subsequent risk of ipsilateral and contralateral invasive breast cancer after treatment for ductal carcinoma in situ: incidence and the effect of radiotherapy in a population-based cohort of 10,090 women. *Breast Cancer Res Treat*. 2016;159(3):553-63.
27. Rakovitch E, Nofech-Mozes S, Narod SA, Hanna W, Thiruchelvam D, et al. Can we select individuals with low risk ductal carcinoma in situ (DCIS)? A population-based outcomes analysis. *Breast Cancer Research and Treatment*. 2013;138(2):581-90.
28. Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer*. 2005;103(12):2481-4.
29. Maxwell AJ, Clements K, Hilton B, Dodwell DJ, Evans A, et al. Risk factors for the development of invasive cancer in unresected ductal carcinoma in situ. *Eur J Surg Oncol*. 2018;44(4):429-35.
30. Maxwell AJ, Hilton B, Clements K, Dodwell D, Dulson-Cox J, et al. Unresected screen-detected ductal carcinoma in situ: Outcomes of 311 women in the Forget-Me-Not 2 study. *Breast*. 2022;61:145-55.
31. van Seijen M, Lips EH, Thompson AM, Nik-Zainal S, Futreal A, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. *British Journal of Cancer*. 2019;121(4):285-92.
32. Allegra CJ, Aberle DR, Ganschow P, Hahn SM, Lee CN, et al. National Institutes of Health State-of-the-Science Conference statement: diagnosis and management of ductal carcinoma in situ September 22–24, 2009. *Journal of the National Cancer Institute*. 2010;102(3):161-9.
33. Cancer Grand Challenges [Available from: <https://cancergrandchallenges.org/teams/precision>].
34. Thompson AM, Clements K, Cheung S, Pinder SE, Lawrence G, et al. Management and 5-year outcomes in 9938 women with screen-detected ductal carcinoma in situ: the UK Sloane Project. *Eur J Cancer*. 2018;101:210-9.
35. Bijker N, Peterse JL, Duchateau L, Julien JP, Fentiman IS, et al. Risk factors for recurrence and metastasis after breast-conserving therapy for ductal carcinoma-in-situ: analysis of European Organization for Research and Treatment of Cancer Trial 10853. *J Clin Oncol*. 2001;19(8):2263-71.
36. Silverstein MJ, Poller DN, Waisman JR, Colburn WJ, Barth A, et al. Prognostic classification of breast ductal carcinoma-in-situ. *Lancet*. 1995;345(8958):1154-7.
37. Kunkiel M, Niwińska A. Assessment of the usefulness of prognostic Van Nuys Prognostic Index in the treatment in ductal carcinoma in situ in 15-year observation. *Scientific Reports*. 2021;11(1):22645.
38. Silverstein MJ. Ductal Carcinoma In Situ of the Breast: Controversial Issues. *Oncologist*. 1998;3(2):94-103.
39. Knopfmacher A, Fox J, Lo Y, Shapiro N, Fineberg S. Correlation of histopathologic features of ductal carcinoma in situ of the breast with the oncotype DX DCIS score. *Mod Pathol*. 2015;28(9):1167-73.

40. Solin LJ, Gray R, Baehner FL, Butler SM, Hughes LL, et al. A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. *J Natl Cancer Inst.* 2013;105(10):701-10.
41. Rakovitch E, Nofech-Mozes S, Hanna W, Baehner FL, Saskin R, et al. A population-based validation study of the DCIS Score predicting recurrence risk in individuals treated by breast-conserving surgery alone. *Breast Cancer Res Treat.* 2015;152(2):389-98.
42. Kerlikowske K, Molinaro AM, Gauthier ML, Berman HK, Waldman F, et al. Biomarker expression and risk of subsequent tumors after initial ductal carcinoma in situ diagnosis. *J Natl Cancer Inst.* 2010;102(9):627-37.
43. Hwang ES, Hyslop T, Lynch T, Frank E, Pinto D, et al. The COMET (Comparison of Operative versus Monitoring and Endocrine Therapy) trial: a phase III randomised controlled clinical trial for low-risk ductal carcinoma in situ (DCIS). *BMJ Open.* 2019;9(3):e026797.
44. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JM, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer.* 2015;51(16):2296-303.
45. Elshof LE, Tryfonidis K, Slaets L, van Leeuwen-Stok AE, Skinner VP, et al. Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ - The LORD study. *Eur J Cancer.* 2015;51(12):1497-510.
46. Angarita FA, Brumer R, Castelo M, Esnaola NF, Edge SB, et al. De-Escalating the Management of In Situ and Invasive Breast Cancer. *Cancers (Basel).* 2022;14(19).
47. Castro NP, Osorio CA, Torres C, Bastos EP, Mourao-Neto M, et al. Evidence that molecular changes in cells occur before morphological alterations during the progression of breast ductal carcinoma. *Breast Cancer Res.* 2008;10(5):R87.
48. Ma XJ, Salunga R, Tuggle JT, Gaudet J, Enright E, et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A.* 2003;100(10):5974-9.
49. Clark SE, Warwick J, Carpenter R, Bowen RL, Duffy SW, et al. Molecular subtyping of DCIS: heterogeneity of breast cancer reflected in pre-invasive disease. *Br J Cancer.* 2011;104(1):120-7.
50. Bergholtz H, Kumar S, Wärnberg F, Lüders T, Kristensen V, et al. Comparable cancer-relevant mutation profiles in synchronous ductal carcinoma in situ and invasive breast cancer. *Cancer Rep (Hoboken).* 2020;3(3):e1248.
51. Agahozo MC, Sieuwerts AM, Doebar SC, Verhoef EI, Beaufort CM, et al. PIK3CA mutations in ductal carcinoma in situ and adjacent invasive breast cancer. *Endocr Relat Cancer.* 2019;26(5):471-82.
52. Martínez-Sáez O, Chic N, Pascual T, Adamo B, Vidal M, et al. Frequency and spectrum of PIK3CA somatic mutations in breast cancer. *Breast Cancer Research.* 2020;22(1):45.
53. Muggerud AA, Hallett M, Johnsen H, Kleivi K, Zhou W, et al. Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer. *Mol Oncol.* 2010;4(4):357-68.
54. Sharma M, Beck AH, Webster JA, Espinosa I, Montgomery K, et al. Analysis of stromal signatures in the tumor microenvironment of ductal carcinoma in situ. *Breast cancer research and treatment.* 2010;123(2):397-404.
55. Nelson AC, Machado HL, Schwertfeger KL. Breaking through to the Other Side: Microenvironment Contributions to DCIS Initiation and Progression. *J Mammary Gland Biol Neoplasia.* 2018;23(4):207-21.
56. Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res.* 2009;11(1):R7.
57. Lesurf R, Aure MR, Mørk HH, Vitelli V, Lundgren S, et al. Molecular Features of Subtype-Specific Progression from Ductal Carcinoma In Situ to Invasive Breast Cancer. *Cell Rep.* 2016;16(4):1166-79.
58. Risom T, Glass DR, Averbukh I, Liu CC, Baranski A, et al. Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma. *Cell.* 2022;185(2):299-310.e18.

59. Adriance MC, Inman JL, Petersen OW, Bissell MJ. Myoepithelial cells: good fences make good neighbors. *Breast Cancer Res.* 2005;7(5):190-7.
60. Gudjonsson T, Adriance MC, Sternlicht MD, Petersen OW, Bissell MJ. Myoepithelial cells: their origin and function in breast morphogenesis and neoplasia. *J Mammary Gland Biol Neoplasia.* 2005;10(3):261-72.
61. Barsky SH. Myoepithelial mRNA expression profiling reveals a common tumor-suppressor phenotype. *Exp Mol Pathol.* 2003;74(2):113-22.
62. Shao ZM, Nguyen M, Alpaugh ML, O'Connell JT, Barsky SH. The human myoepithelial cell exerts antiproliferative effects on breast carcinoma cells characterized by p21WAF1/CIP1 induction, G2/M arrest, and apoptosis. *Exp Cell Res.* 1998;241(2):394-403.
63. Zou Z, Anisowicz A, Hendrix MJ, Thor A, Neveu M, et al. Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells. *Science.* 1994;263(5146):526-9.
64. Runswick SK, O'Hare MJ, Jones L, Streuli CH, Garrod DR. Desmosomal adhesion regulates epithelial morphogenesis and cell positioning. *Nat Cell Biol.* 2001;3(9):823-30.
65. Sirka OK, Shamir ER, Ewald AJ. Myoepithelial cells are a dynamic barrier to epithelial dissemination. *The Journal of cell biology.* 2018;217(10):3368-81.
66. Barsky SH, Karlin NJ. Myoepithelial cells: autocrine and paracrine suppressors of breast cancer progression. *J Mammary Gland Biol Neoplasia.* 2005;10(3):249-60.
67. Cowell CF, Weigelt B, Sakr RA, Ng CK, Hicks J, et al. Progression from ductal carcinoma in situ to invasive breast cancer: revisited. *Mol Oncol.* 2013;7(5):859-69.
68. Polyak K, Hu M. Do myoepithelial cells hold the key for breast tumor progression? *J Mammary Gland Biol Neoplasia.* 2005;10(3):231-47.
69. Allinen M, Beroukhi R, Cai L, Brennan C, Lahti-Domenici J, et al. Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell.* 2004;6(1):17-32.
70. Gudjonsson T, Ronnov-Jessen L, Villadsen R, Rank F, Bissell MJ, et al. Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition. *J Cell Sci.* 2002;115(Pt 1):39-50.
71. Lo PK, Zhang Y, Yao Y, Wolfson B, Yu J, et al. Tumor-associated myoepithelial cells promote the invasive progression of ductal carcinoma in situ through activation of TGF $\beta$  signaling. *J Biol Chem.* 2017;292(27):11466-84.
72. Allen MD, Thomas GJ, Clark S, Dawoud MM, Vallath S, et al. Altered microenvironment promotes progression of preinvasive breast cancer: myoepithelial expression of  $\alpha$ 5 $\beta$ 1 integrin in DCIS identifies high-risk patients and predicts recurrence. *Clin Cancer Res.* 2014;20(2):344-57.
73. Kim HS, Jung M, Choi SK, Woo J, Piao YJ, et al. IL-6-mediated cross-talk between human preadipocytes and ductal carcinoma in situ in breast cancer progression. *Journal of Experimental & Clinical Cancer Research.* 2018;37(1):200.
74. Andarawewa KL, Motrescu ER, Chenard MP, Gansmuller A, Stoll I, et al. Stromelysin-3 is a potent negative regulator of adipogenesis participating to cancer cell-adipocyte interaction/crosstalk at the tumor invasive front. *Cancer Res.* 2005;65(23):10862-71.
75. Delort L, Cholet J, Decombat C, Vermerie M, Dumontet C, et al. The Adipose Microenvironment Dysregulates the Mammary Myoepithelial Cells and Could Participate to the Progression of Breast Cancer. *Front Cell Dev Biol.* 2020;8:571948.
76. Picon-Ruiz M, Pan C, Drews-Elger K, Jang K, Besser AH, et al. Interactions between Adipocytes and Breast Cancer Cells Stimulate Cytokine Production and Drive Src/Sox2/miR-302b-Mediated Malignant Progression. *Cancer Research.* 2016;76(2):491-504.
77. Gernapudi R, Yao Y, Zhang Y, Wolfson B, Roy S, et al. Targeting exosomes from preadipocytes inhibits preadipocyte to cancer stem cell signaling in early-stage breast cancer. *Breast Cancer Res Treat.* 2015;150(3):685-95.
78. Chamberlin T, D'Amato JV, Arendt LM. Obesity reversibly depletes the basal cell population and enhances mammary epithelial cell estrogen receptor  $\alpha$  expression and progenitor activity. *Breast Cancer Res.* 2017;19(1):128.

79. Iyengar NM, Gucalp A, Dannenberg AJ, Hudis CA. Obesity and Cancer Mechanisms: Tumor Microenvironment and Inflammation. *J Clin Oncol*. 2016;34(35):4270-6.
80. Carter JM, Hoskin TL, Pena MA, Brahmbhatt R, Winham SJ, et al. Macrophagic “Crown-like Structures” Are Associated with an Increased Risk of Breast Cancer in Benign Breast Disease. *Cancer Prevention Research*. 2018;11(2):113-9.
81. Wilcz-Villega E, Carter E, Ironside A, Xu R, Mataloni I, et al. Macrophages induce malignant traits in mammary epithelium via IKK $\epsilon$ /TBK1 kinases and the serine biosynthesis pathway. *EMBO Mol Med*. 2020;12(2):e10491.
82. Iyengar NM, Zhou XK, Gucalp A, Morris PG, Howe LR, et al. Systemic Correlates of White Adipose Tissue Inflammation in Early-Stage Breast Cancer. *Clinical Cancer Research*. 2016;22(9):2283-9.
83. Maliniak ML, Cheriyan AM, Sherman ME, Liu Y, Gogineni K, et al. Detection of crown-like structures in breast adipose tissue and clinical outcomes among African-American and White women with breast cancer. *Breast Cancer Research*. 2020;22(1):65.
84. Almekinders MMM, Schaapveld M, Thijssen B, Visser LL, Bismeljer T, et al. Breast adipocyte size associates with ipsilateral invasive breast cancer risk after ductal carcinoma in situ. *npj Breast Cancer*. 2021;7(1):31.
85. Visser LL, Elshof LE, Schaapveld M, van de Vijver K, Groen EJ, et al. Clinicopathological Risk Factors for an Invasive Breast Cancer Recurrence after Ductal Carcinoma In Situ—A Nested Case–Control Study. *Clinical Cancer Research*. 2018;24(15):3593-601.
86. Ibrahim EM, Al-Foheidi ME, Al-Mansour MM, Kazkaz GA. The prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancer: a meta-analysis. *Breast Cancer Res Treat*. 2014;148(3):467-76.
87. Adams S, Gray RJ, Demaria S, Goldstein L, Perez EA, et al. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. *J Clin Oncol*. 2014;32(27):2959-66.
88. Toss MS, Abidi A, Lesche D, Joseph C, Mahale S, et al. The prognostic significance of immune microenvironment in breast ductal carcinoma in situ. *British Journal of Cancer*. 2020;122(10):1496-506.
89. Toss MS, Miligy I, Al-Kawaz A, Alsleem M, Khout H, et al. Prognostic significance of tumor-infiltrating lymphocytes in ductal carcinoma in situ of the breast. *Modern Pathology*. 2018;31(8):1226-36.
90. Pruneri G, Lazzeroni M, Bagnardi V, Tiburzio GB, Rotmensz N, et al. The prevalence and clinical relevance of tumor-infiltrating lymphocytes (TILs) in ductal carcinoma in situ of the breast. *Ann Oncol*. 2017;28(2):321-8.
91. Hendry S, Pang JB, Byrne DJ, Lakhani SR, Cummings MC, et al. Relationship of the Breast Ductal Carcinoma In Situ Immune Microenvironment with Clinicopathological and Genetic Features. *Clin Cancer Res*. 2017;23(17):5210-7.
92. Kim M, Chung YR, Kim HJ, Woo JW, Ahn S, et al. Immune microenvironment in ductal carcinoma in situ: a comparison with invasive carcinoma of the breast. *Breast Cancer Research*. 2020;22(1):32.
93. Gil Del Alcazar CR, Huh SJ, Ekram MB, Trinh A, Liu LL, et al. Immune Escape in Breast Cancer During In Situ to Invasive Carcinoma Transition. *Cancer Discov*. 2017;7(10):1098-115.
94. Chen X-Y, Thike AA, Md Nasir ND, Koh VCY, Bay BH, et al. Higher density of stromal M2 macrophages in breast ductal carcinoma in situ predicts recurrence. *Virchows Archiv*. 2020;476(6):825-33.
95. Linde N, Casanova-Acebes M, Sosa MS, Mortha A, Rahman A, et al. Macrophages orchestrate breast cancer early dissemination and metastasis. *Nature Communications*. 2018;9(1):21.
96. Campbell MJ, Baehner F, O'Meara T, Ojukwu E, Han B, et al. Characterizing the immune microenvironment in high-risk ductal carcinoma in situ of the breast. *Breast Cancer Res Treat*. 2017;161(1):17-28.



97. Guidi AJ, Fischer L, Harris JR, Schnitt SJ. Microvessel Density and Distribution in Ductal Carcinoma In Situ of the Breast. *JNCI: Journal of the National Cancer Institute*. 1994;86(8):614-9.
98. Zhang Z, Luo G, Tang H, Cheng C, Wang P. Prognostic Significance of High VEGF-C Expression for Patients with Breast Cancer: An Update Meta Analysis. *PLoS One*. 2016;11(11):e0165725.
99. Engels K, Fox SB, Whitehouse RM, Gatter KC, Harris AL. Distinct angiogenic patterns are associated with high-grade in situ ductal carcinomas of the breast. *J Pathol*. 1997;181(2):207-12.
100. Lee AH, Happerfield LC, Bobrow LG, Millis RR. Angiogenesis and inflammation in ductal carcinoma in situ of the breast. *J Pathol*. 1997;181(2):200-6.
101. Teo NB, Shoker BS, Jarvis C, Martin L, Sloane JP, et al. Angiogenesis and invasive recurrence in ductal carcinoma in situ of the breast. *Eur J Cancer*. 2003;39(1):38-44.
102. Wülfing P, Kersting C, Buerger H, Mattsson B, Mesters R, et al. Expression patterns of angiogenic and lymphangiogenic factors in ductal breast carcinoma in situ. *British Journal of Cancer*. 2005;92(9):1720-8.
103. Lee AH, Dublin EA, Bobrow LG. Angiogenesis and expression of thymidine phosphorylase by inflammatory and carcinoma cells in ductal carcinoma in situ of the breast. *J Pathol*. 1999;187(3):285-90.
104. Nguyen M, Lee MC, Wang JL, Tomlinson JS, Shao ZM, et al. The human myoepithelial cell displays a multifaceted anti-angiogenic phenotype. *Oncogene*. 2000;19(31):3449-59.
105. Nguyen-Ngoc K-V, Cheung KJ, Brenot A, Shamir ER, Gray RS, et al. ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium. *Proceedings of the National Academy of Sciences*. 2012;109(39):E2595-E604.
106. Hildenbrand R, Arens N. Protein and mRNA expression of uPAR and PAI-1 in myoepithelial cells of early breast cancer lesions and normal breast tissue. *Br J Cancer*. 2004;91(3):564-71.
107. Sahai E, Astsaturou I, Cukierman E, DeNardo DG, Egeblad M, et al. A framework for advancing our understanding of cancer-associated fibroblasts. *Nature Reviews Cancer*. 2020;20(3):174-86.
108. Kalluri R. The biology and function of fibroblasts in cancer. *Nat Rev Cancer*. 2016;16(9):582-98.
109. Xiang L, Song Z, Rong G. Taxotere-induced WNT16 Expression in Carcinoma-Associated Fibroblasts Might Associate with Progression and Chemoresistance of Breast Cancer. *Ann Clin Lab Sci*. 2020;50(2):205-12.
110. Costa A, Kieffer Y, Scholer-Dahirel A, Pelon F, Bourachot B, et al. Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell*. 2018;33(3):463-79.e10.
111. Holliday DL, Brouillette KT, Markert A, Gordon LA, Jones JL. Novel multicellular organotypic models of normal and malignant breast: tools for dissecting the role of the microenvironment in breast cancer progression. *Breast Cancer Res*. 2009;11(1):R3.
112. Hu M, Peluffo G, Chen H, Gelman R, Schnitt S, et al. Role of COX-2 in epithelial-stromal cell interactions and progression of ductal carcinoma in situ of the breast. *Proc Natl Acad Sci U S A*. 2009;106(9):3372-7.
113. Osuala KO, Sameni M, Shah S, Aggarwal N, Simonait ML, et al. Il-6 signaling between ductal carcinoma in situ cells and carcinoma-associated fibroblasts mediates tumor cell growth and migration. *BMC Cancer*. 2015;15:584.
114. Bernard S, Myers M, Fang WB, Zinda B, Smart C, et al. CXCL1 Derived from Mammary Fibroblasts Promotes Progression of Mammary Lesions to Invasive Carcinoma through CXCR2 Dependent Mechanisms. *J Mammary Gland Biol Neoplasia*. 2018;23(4):249-67.
115. Brummer G, Acevedo DS, Hu Q, Portsche M, Fang WB, et al. Chemokine Signaling Facilitates Early-Stage Breast Cancer Survival and Invasion through Fibroblast-Dependent Mechanisms. *Mol Cancer Res*. 2018;16(2):296-308.
116. Strell C, Paulsson J, Jin SB, Tobin NP, Mezheyski A, et al. Impact of Epithelial-Stromal Interactions on Peritumoral Fibroblasts in Ductal Carcinoma in Situ. *J Natl Cancer Inst*. 2019;111(9):983-95.

117. Witkiewicz AK, Dasgupta A, Nguyen K, Liu C, Kovatich AJ, et al. Stromal caveolin-1 levels predict early DCIS progression to invasive breast cancer. *Cancer Biology & Therapy*. 2009;8(11):1071-9.
118. Peuhu E, Jacquemet G, Scheele C, Isomursu A, Laisne MC, et al. MYO10-filopodia support basement membranes at pre-invasive tumor boundaries. *Dev Cell*. 2022;57(20):2350-64.e7.
119. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nature Communications*. 2020;11(1):5120.
120. Kaushik S, Pickup MW, Weaver VM. From transformation to metastasis: deconstructing the extracellular matrix in breast cancer. *Cancer Metastasis Rev*. 2016;35(4):655-67.
121. Acerbi I, Cassereau L, Dean I, Shi Q, Au A, et al. Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. *Integr Biol (Camb)*. 2015;7(10):1120-34.
122. Conklin MW, Eickhoff JC, Riching KM, Pehlke CA, Eliceiri KW, et al. Aligned collagen is a prognostic signature for survival in human breast carcinoma. *Am J Pathol*. 2011;178(3):1221-32.
123. Provenzano PP, Eliceiri KW, Campbell JM, Inman DR, White JG, et al. Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med*. 2006;4(1):38.
124. Koorman T, Jansen KA, Khalil A, Haughton PD, Visser D, et al. Spatial collagen stiffening promotes collective breast cancer cell invasion by reinforcing extracellular matrix alignment. *Oncogene*. 2022;41(17):2458-69.
125. Conklin MW, Gangnon RE, Sprague BL, Van Gemert L, Hampton JM, et al. Collagen Alignment as a Predictor of Recurrence after Ductal Carcinoma In Situ. *Cancer Epidemiology, Biomarkers & Prevention*. 2018;27(2):138-45.
126. Papanicolaou M, Parker AL, Yam M, Filipe EC, Wu SZ, et al. Temporal profiling of the breast tumour microenvironment reveals collagen XII as a driver of metastasis. *Nature Communications*. 2022;13(1):4587.
127. Hayward M-K, Allen MD, Gomm JJ, Goulding I, Thompson CL, et al. Mechanostimulation of breast myoepithelial cells induces functional changes associated with DCIS progression to invasion. *npj Breast Cancer*. 2022;8(1):109.
128. Attieh Y, Clark AG, Grass C, Richon S, Pocard M, et al. Cancer-associated fibroblasts lead tumor invasion through integrin- $\beta$ 3-dependent fibronectin assembly. *J Cell Biol*. 2017;216(11):3509-20.
129. Tarullo SE, Hill RC, Hansen KC, Behbod F, Borges VF, et al. Postpartum breast cancer progression is driven by semaphorin 7a-mediated invasion and survival. *Oncogene*. 2020;39(13):2772-85.
130. García-Mendoza MG, Inman DR, Ponik SM, Jeffery JJ, Sheerar DS, et al. Neutrophils drive accelerated tumor progression in the collagen-dense mammary tumor microenvironment. *Breast Cancer Res*. 2016;18(1):49.
131. Schedin P, Keely PJ. Mammary gland ECM remodeling, stiffness, and mechanosignaling in normal development and tumor progression. *Cold Spring Harb Perspect Biol*. 2011;3(1):a003228.
132. Rubashkin MG, Cassereau L, Bainer R, DuFort CC, Yui Y, et al. Force engages vinculin and promotes tumor progression by enhancing PI3K activation of phosphatidylinositol (3, 4, 5)-triphosphate. *Cancer research*. 2014;74(17):4597-611.
133. Escalona RM, Chan E, Kannourakis G, Findlay JK, Ahmed N. The Many Facets of Metzincins and Their Endogenous Inhibitors: Perspectives on Ovarian Cancer Progression. *Int J Mol Sci*. 2018;19(2).
134. Park KC, Dharmasivam M, Richardson DR. The Role of Extracellular Proteases in Tumor Progression and the Development of Innovative Metal Ion Chelators that Inhibit their Activity. *Int J Mol Sci*. 2020;21(18).
135. López-Otín C, Matrisian LM. Emerging roles of proteases in tumour suppression. *Nature Reviews Cancer*. 2007;7(10):800-8.
136. Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. *Nature Reviews Cancer*. 2002;2(3):161-74.
137. Madzharova E, Kastl P, Sabino F, Auf dem Keller U. Post-Translational Modification-Dependent Activity of Matrix Metalloproteinases. *Int J Mol Sci*. 2019;20(12).

138. Turk B. Targeting proteases: successes, failures and future prospects. *Nature Reviews Drug Discovery*. 2006;5(9):785-99.
139. Vizovisek M, Ristanovic D, Menghini S, Christiansen MG, Schuerle S. The Tumor Proteolytic Landscape: A Challenging Frontier in Cancer Diagnosis and Therapy. *Int J Mol Sci*. 2021;22(5).
140. Duffy MJ, Mullooly M, O'Donovan N, Sukor S, Crown J, et al. The ADAMs family of proteases: new biomarkers and therapeutic targets for cancer? *Clinical Proteomics*. 2011;8(1):9.
141. Seals DF, Courtneidge SA. The ADAMs family of metalloproteases: multidomain proteins with multiple functions. *Genes Dev*. 2003;17(1):7-30.
142. Edwards DR, Handsley MM, Pennington CJ. The ADAM metalloproteinases. *Mol Aspects Med*. 2008;29(5):258-89.
143. Wang L, Hoggard JA, Korleski ED, Long GV, Ree BC, et al. Multiple non-catalytic ADAMs are novel integrin  $\alpha 4$  ligands. *Molecular and Cellular Biochemistry*. 2018;442(1):29-38.
144. Zhong S, Khalil RA. A Disintegrin and Metalloproteinase (ADAM) and ADAM with thrombospondin motifs (ADAMTS) family in vascular biology and disease. *Biochem Pharmacol*. 2019;164:188-204.
145. Kelwick R, Desanlis I, Wheeler GN, Edwards DR. The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family. *Genome Biol*. 2015;16(1):113.
146. Sternlicht MD, Werb Z. How matrix metalloproteinases regulate cell behavior. *Annu Rev Cell Dev Biol*. 2001;17:463-516.
147. Strongin AY, Collier I, Bannikov G, Marmer BL, Grant GA, et al. Mechanism of cell surface activation of 72-kDa type IV collagenase. Isolation of the activated form of the membrane metalloprotease. *J Biol Chem*. 1995;270(10):5331-8.
148. Pei D, Weiss SJ. Furin-dependent intracellular activation of the human stromelysin-3 zymogen. *Nature*. 1995;375(6528):244-7.
149. Langton KP, Barker MD, McKie N. Localization of the functional domains of human tissue inhibitor of metalloproteinases-3 and the effects of a Sorsby's fundus dystrophy mutation. *J Biol Chem*. 1998;273(27):16778-81.
150. Ko YC, Langley KE, Mendiaz EA, Parker VP, Taylor SM, et al. The C-terminal domain of tissue inhibitor of metalloproteinases-2 is required for cell binding but not for antimetalloproteinase activity. *Biochem Biophys Res Commun*. 1997;236(1):100-5.
151. Vandooren J, Itoh Y. Alpha-2-Macroglobulin in Inflammation, Immunity and Infections. *Front Immunol*. 2021;12:803244.
152. Lochter A, Galosy S, Muschler J, Freedman N, Werb Z, et al. Matrix metalloproteinase stromelysin-1 triggers a cascade of molecular alterations that leads to stable epithelial-to-mesenchymal conversion and a premalignant phenotype in mammary epithelial cells. *J Cell Biol*. 1997;139(7):1861-72.
153. Mitsiades N, Yu WH, Poulaki V, Tsokos M, Stamenkovic I. Matrix metalloproteinase-7-mediated cleavage of Fas ligand protects tumor cells from chemotherapeutic drug cytotoxicity. *Cancer Res*. 2001;61(2):577-81.
154. Wang WS, Chen PM, Wang HS, Liang WY, Su Y. Matrix metalloproteinase-7 increases resistance to Fas-mediated apoptosis and is a poor prognostic factor of patients with colorectal carcinoma. *Carcinogenesis*. 2006;27(5):1113-20.
155. Liu H, Zhang T, Li X, Huang J, Wu B, et al. Predictive value of MMP-7 expression for response to chemotherapy and survival in patients with non-small cell lung cancer. *Cancer Sci*. 2008;99(11):2185-92.
156. Dufour A, Sampson NS, Zucker S, Cao J. Role of the hemopexin domain of matrix metalloproteinases in cell migration. *J Cell Physiol*. 2008;217(3):643-51.
157. García-Pardo A, Opdenakker G. Nonproteolytic functions of matrix metalloproteinases in pathology and insights for the development of novel therapeutic inhibitors. *Metalloproteinases Med*. 2015;2:19-28.

158. Chetty C, Lakka SS, Bhoopathi P, Rao JS. MMP-2 alters VEGF expression via  $\alpha$ V $\beta$ 3 integrin-mediated PI3K/AKT signaling in A549 lung cancer cells. *Int J Cancer*. 2010;127(5):1081-95.
159. Chattopadhyay S, Shubayev VI. MMP-9 controls Schwann cell proliferation and phenotypic remodeling via IGF-1 and ErbB receptor-mediated activation of MEK/ERK pathway. *Glia*. 2009;57(12):1316-25.
160. Sarper M, Allen MD, Gomm J, Haywood L, Decock J, et al. Loss of MMP-8 in ductal carcinoma in situ (DCIS)-associated myoepithelial cells contributes to tumour promotion through altered adhesive and proteolytic function. *Breast Cancer Res*. 2017;19(1):33.
161. Montel V, Kleeman J, Agarwal D, Spinella D, Kawai K, et al. Altered metastatic behavior of human breast cancer cells after experimental manipulation of matrix metalloproteinase 8 gene expression. *Cancer Res*. 2004;64(5):1687-94.
162. Gorrin-Rivas MJ, Arai S, Furutani M, Mizumoto M, Mori A, et al. Mouse macrophage metalloelastase gene transfer into a murine melanoma suppresses primary tumor growth by halting angiogenesis. *Clin Cancer Res*. 2000;6(5):1647-54.
163. Dong Z, Kumar R, Yang X, Fidler IJ. Macrophage-derived metalloelastase is responsible for the generation of angiostatin in Lewis lung carcinoma. *Cell*. 1997;88(6):801-10.
164. Miller MA, Sullivan RJ, Lauffenburger DA. Molecular Pathways: Receptor Ectodomain Shedding in Treatment, Resistance, and Monitoring of Cancer. *Clin Cancer Res*. 2017;23(3):623-9.
165. Kenny PA, Bissell MJ. Targeting TACE-dependent EGFR ligand shedding in breast cancer. *J Clin Invest*. 2007;117(2):337-45.
166. Zheng X, Jiang F, Katakowski M, Zhang ZG, Lu QE, et al. ADAM17 promotes breast cancer cell malignant phenotype through EGFR-PI3K-AKT activation. *Cancer Biol Ther*. 2009;8(11):1045-54.
167. Sahin U, Weskamp G, Kelly K, Zhou HM, Higashiyama S, et al. Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR ligands. *J Cell Biol*. 2004;164(5):769-79.
168. Najy AJ, Day KC, Day ML. The ectodomain shedding of E-cadherin by ADAM15 supports ErbB receptor activation. *J Biol Chem*. 2008;283(26):18393-401.
169. Arribas J, Bech-Serra JJ, Santiago-Josefat B. ADAMs, cell migration and cancer. *Cancer Metastasis Rev*. 2006;25(1):57-68.
170. Nath D, Slocombe PM, Webster A, Stephens PE, Docherty AJ, et al. Meltrin gamma (ADAM-9) mediates cellular adhesion through  $\alpha$ (6) $\beta$ (1) integrin, leading to a marked induction of fibroblast cell motility. *J Cell Sci*. 2000;113 ( Pt 12):2319-28.
171. Chen J, Zhang C, Xu X, Zhu X, Dai D. Downregulation of A disintegrin and metalloproteinase with thrombospondin motif type 1 by DNA hypermethylation in human gastric cancer. *Mol Med Rep*. 2015;12(2):2487-94.
172. Freitas VM, do Amaral JB, Silva TA, Santos ES, Mangone FR, et al. Decreased expression of ADAMTS-1 in human breast tumors stimulates migration and invasion. *Molecular Cancer*. 2013;12(1):2.
173. Du W, Wang S, Zhou Q, Li X, Chu J, et al. ADAMTS9 is a functional tumor suppressor through inhibiting AKT/mTOR pathway and associated with poor survival in gastric cancer. *Oncogene*. 2013;32(28):3319-28.
174. Choi GC, Li J, Wang Y, Li L, Zhong L, et al. The metalloprotease ADAMTS8 displays antitumor properties through antagonizing EGFR-MEK-ERK signaling and is silenced in carcinomas by CpG methylation. *Mol Cancer Res*. 2014;12(2):228-38.
175. Llamazares M, Obaya AJ, Moncada-Pazos A, Heljasvaara R, Espada J, et al. The ADAMTS12 metalloproteinase exhibits anti-tumorigenic properties through modulation of the Ras-dependent ERK signalling pathway. *J Cell Sci*. 2007;120(Pt 20):3544-52.
176. Vilorio CG, Obaya AJ, Moncada-Pazos A, Llamazares M, Astudillo A, et al. Genetic inactivation of ADAMTS15 metalloprotease in human colorectal cancer. *Cancer Res*. 2009;69(11):4926-34.
177. Dunn JR, Reed JE, du Plessis DG, Shaw EJ, Reeves P, et al. Expression of ADAMTS-8, a secreted protease with antiangiogenic properties, is downregulated in brain tumours. *Br J Cancer*. 2006;94(8):1186-93.

178. Lee NV, Sato M, Annis DS, Loo JA, Wu L, et al. ADAMTS1 mediates the release of antiangiogenic polypeptides from TSP1 and 2. *Embo j.* 2006;25(22):5270-83.
179. Freitas VM, do Amaral JB, Silva TA, Santos ES, Mangone FR, et al. Decreased expression of ADAMTS-1 in human breast tumors stimulates migration and invasion. *Mol Cancer.* 2013;12:2.
180. Porter S, Scott SD, Sassoon EM, Williams MR, Jones JL, et al. Dysregulated expression of adamalysin-thrombospondin genes in human breast carcinoma. *Clinical Cancer Research.* 2004;10(7):2429-40.
181. Hehlhans S, Haase M, Cordes N. Signalling via integrins: implications for cell survival and anticancer strategies. *Biochim Biophys Acta.* 2007;1775(1):163-80.
182. Cooper J, Giancotti FG. Integrin Signaling in Cancer: Mechanotransduction, Stemness, Epithelial Plasticity, and Therapeutic Resistance. *Cancer Cell.* 2019;35(3):347-67.
183. Desgrosellier JS, Cheresh DA. Integrins in cancer: biological implications and therapeutic opportunities. *Nat Rev Cancer.* 2010;10(1):9-22.
184. Hamidi H, Ivaska J. Every step of the way: integrins in cancer progression and metastasis. *Nature Reviews Cancer.* 2018;18(9):533-48.
185. Kechagia JZ, Ivaska J, Roca-Cusachs P. Integrins as biomechanical sensors of the microenvironment. *Nature Reviews Molecular Cell Biology.* 2019;20(8):457-73.
186. Hynes RO. Integrins: bidirectional, allosteric signaling machines. *cell.* 2002;110(6):673-87.
187. Abram CL, Lowell CA. The ins and outs of leukocyte integrin signaling. *Annu Rev Immunol.* 2009;27:339-62.
188. Raab-Westphal S, Marshall JF, Goodman SL. Integrins as Therapeutic Targets: Successes and Cancers. *Cancers (Basel).* 2017;9(9).
189. Seguin L, Desgrosellier JS, Weis SM, Cheresh DA. Integrins and cancer: regulators of cancer stemness, metastasis, and drug resistance. *Trends Cell Biol.* 2015;25(4):234-40.
190. Yom CK, Noh DY, Kim WH, Kim HS. Clinical significance of high focal adhesion kinase gene copy number and overexpression in invasive breast cancer. *Breast Cancer Res Treat.* 2011;128(3):647-55.
191. Desgrosellier JS, Lesperance J, Seguin L, Gozo M, Kato S, et al. Integrin  $\alpha\beta 3$  drives slug activation and stemness in the pregnant and neoplastic mammary gland. *Dev Cell.* 2014;30(3):295-308.
192. Ramirez NE, Zhang Z, Madamanchi A, Boyd KL, O'Rear LD, et al. The  $\alpha_2\beta_1$  integrin is a metastasis suppressor in mouse models and human cancer. *J Clin Invest.* 2011;121(1):226-37.
193. Munshi HG, Stack MS. Reciprocal interactions between adhesion receptor signaling and MMP regulation. *Cancer Metastasis Rev.* 2006;25(1):45-56.
194. Li X, Yang Y, Hu Y, Dang D, Regezi J, et al.  $\alpha\beta 6$ -Fyn signaling promotes oral cancer progression. *Journal of Biological Chemistry.* 2003;278(43):41646-53.
195. Gu X, Niu J, Dorahy DJ, Scott R, Agrez MV. Integrin  $\alpha(v)\beta 6$ -associated ERK2 mediates MMP-9 secretion in colon cancer cells. *Br J Cancer.* 2002;87(3):348-51.
196. Ahmed N, Pansino F, Clyde R, Murthi P, Quinn MA, et al. Overexpression of  $\alpha(v)\beta 6$  integrin in serous epithelial ovarian cancer regulates extracellular matrix degradation via the plasminogen activation cascade. *Carcinogenesis.* 2002;23(2):237-44.
197. Baum O, Hlushchuk R, Forster A, Greiner R, Cl  zardin P, et al. Increased invasive potential and up-regulation of MMP-2 in MDA-MB-231 breast cancer cells expressing the  $\beta 3$  integrin subunit. *Int J Oncol.* 2007;30(2):325-32.
198. Brooks PC, Str  mblad S, Sanders LC, von Schalscha TL, Aimes RT, et al. Localization of matrix metalloproteinase MMP-2 to the surface of invasive cells by interaction with integrin  $\alpha v \beta 3$ . *Cell.* 1996;85(5):683-93.
199. Yang GY, Xu KS, Pan ZQ, Zhang ZY, Mi YT, et al. Integrin  $\alpha v \beta 6$  mediates the potential for colon cancer cells to colonize in and metastasize to the liver. *Cancer Sci.* 2008;99(5):879-87.
200. Morini M, Mottolese M, Ferrari N, Ghorzo F, Buglioni S, et al. The  $\alpha 3 \beta 1$  integrin is associated with mammary carcinoma cell metastasis, invasion, and gelatinase B (MMP-9) activity. *Int J Cancer.* 2000;87(3):336-42.

201. Rolli M, Fransvea E, Pilch J, Saven A, Felding-Habermann B. Activated integrin  $\alpha v \beta 3$  cooperates with metalloproteinase MMP-9 in regulating migration of metastatic breast cancer cells. *Proc Natl Acad Sci U S A*. 2003;100(16):9482-7.
202. Erdogan B, Ao M, White LM, Means AL, Brewer BM, et al. Cancer-associated fibroblasts promote directional cancer cell migration by aligning fibronectin. *J Cell Biol*. 2017;216(11):3799-816.
203. Cox TR, Bird D, Baker AM, Barker HE, Ho MW, et al. LOX-mediated collagen crosslinking is responsible for fibrosis-enhanced metastasis. *Cancer Res*. 2013;73(6):1721-32.
204. Voloshenyuk TG, Landesman ES, Khoutorova E, Hart AD, Gardner JD. Induction of cardiac fibroblast lysyl oxidase by TGF- $\beta 1$  requires PI3K/Akt, Smad3, and MAPK signaling. *Cytokine*. 2011;55(1):90-7.
205. Meecham A, Marshall JF. The ITGB6 gene: its role in experimental and clinical biology. *Gene*. 2020;5:100023.
206. Allen MD, Marshall JF, Jones JL.  $\alpha v \beta 6$  Expression in myoepithelial cells: a novel marker for predicting DCIS progression with therapeutic potential. *Cancer research*. 2014;74(21):5942-7.
207. Moore KM, Thomas GJ, Duffy SW, Warwick J, Gabe R, et al. Therapeutic targeting of integrin  $\alpha v \beta 6$  in breast cancer. *J Natl Cancer Inst*. 2014;106(8).
208. Bates RC. The  $\alpha v \beta 6$  integrin as a novel molecular target for colorectal cancer. 2005.
209. Zhuang Z, Zhou R, Xu X, Tian T, Liu Y, et al. Clinical significance of integrin  $\alpha v \beta 6$  expression effects on gastric carcinoma invasiveness and progression via cancer-associated fibroblasts. *Med Oncol*. 2013;30(3):580.
210. Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *Lancet Oncol*. 2011;12(1):21-9.
211. Ikushima H, Miyazono K. TGF $\beta$  signalling: a complex web in cancer progression. *Nat Rev Cancer*. 2010;10(6):415-24.
212. Padua D, Massagué J. Roles of TGF $\beta$  in metastasis. *Cell Res*. 2009;19(1):89-102.
213. Annes JP, Munger JS, Rifkin DB. Making sense of latent TGF $\beta$  activation. *J Cell Sci*. 2003;116(Pt 2):217-24.
214. Massagué J. TGF $\beta$  signalling in context. *Nature Reviews Molecular Cell Biology*. 2012;13(10):616-30.
215. Derynck R, Turley SJ, Akhurst RJ. TGF $\beta$  biology in cancer progression and immunotherapy. *Nature Reviews Clinical Oncology*. 2021;18(1):9-34.
216. Levy L, Hill CS. Alterations in components of the TGF-beta superfamily signaling pathways in human cancer. *Cytokine Growth Factor Rev*. 2006;17(1-2):41-58.
217. Yagi K, Furuhashi M, Aoki H, Goto D, Kuwano H, et al. c-myc is a downstream target of the Smad pathway. *J Biol Chem*. 2002;277(1):854-61.
218. Datto MB, Li Y, Panus JF, Howe DJ, Xiong Y, et al. Transforming growth factor beta induces the cyclin-dependent kinase inhibitor p21 through a p53-independent mechanism. *Proc Natl Acad Sci U S A*. 1995;92(12):5545-9.
219. Tang B, Yoo N, Vu M, Mamura M, Nam JS, et al. Transforming growth factor-beta can suppress tumorigenesis through effects on the putative cancer stem or early progenitor cell and committed progeny in a breast cancer xenograft model. *Cancer Res*. 2007;67(18):8643-52.
220. Walker RA, Dearing SJ. Transforming growth factor beta 1 in ductal carcinoma in situ and invasive carcinomas of the breast. *Eur J Cancer*. 1992;28(2-3):641-4.
221. de Jong JS, van Diest PJ, van der Valk P, Baak JP. Expression of growth factors, growth inhibiting factors, and their receptors in invasive breast cancer. I: An inventory in search of autocrine and paracrine loops. *J Pathol*. 1998;184(1):44-52.
222. Dalal BI, Keown PA, Greenberg AH. Immunocytochemical localization of secreted transforming growth factor-beta 1 to the advancing edges of primary tumors and to lymph node metastases of human mammary carcinoma. *The American journal of pathology*. 1993;143(2):381.

223. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* 2014;15(3):178-96.
224. Xu J, Lamouille S, Derynck R. TGF- $\beta$ -induced epithelial to mesenchymal transition. *Cell Research.* 2009;19(2):156-72.
225. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, et al. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell.* 2003;3(6):537-49.
226. Batlle E, Massagué J. Transforming Growth Factor- $\beta$  Signaling in Immunity and Cancer. *Immunity.* 2019;50(4):924-40.
227. Takimoto T, Wakabayashi Y, Sekiya T, Inoue N, Morita R, et al. Smad2 and Smad3 are redundantly essential for the TGF-beta-mediated regulation of regulatory T plasticity and Th1 development. *J Immunol.* 2010;185(2):842-55.
228. McKarns SC, Schwartz RH, Kaminski NE. Smad3 is essential for TGF-beta 1 to suppress IL-2 production and TCR-induced proliferation, but not IL-2-induced proliferation. *J Immunol.* 2004;172(7):4275-84.
229. Thomas DA, Massagué J. TGF-beta directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer Cell.* 2005;8(5):369-80.
230. Lin TC, Yang CH, Cheng LH, Chang WT, Lin YR, et al. Fibronectin in Cancer: Friend or Foe. *Cells.* 2019;9(1).
231. Spada S, Tocci A, Di Modugno F, Nisticò P. Fibronectin as a multiregulatory molecule crucial in tumor matrisome: from structural and functional features to clinical practice in oncology. *J Exp Clin Cancer Res.* 2021;40(1):102.
232. Pankov R, Yamada KM. Fibronectin at a glance. *J Cell Sci.* 2002;115(Pt 20):3861-3.
233. Martino MM, Hubbell JA. The 12th-14th type III repeats of fibronectin function as a highly promiscuous growth factor-binding domain. *Faseb j.* 2010;24(12):4711-21.
234. Dallas SL, Sivakumar P, Jones CJ, Chen Q, Peters DM, et al. Fibronectin regulates latent transforming growth factor-beta (TGF beta) by controlling matrix assembly of latent TGF beta-binding protein-1. *J Biol Chem.* 2005;280(19):18871-80.
235. Koukoulis GK, Howedy AA, Korhonen M, Virtanen I, Gould VE. Distribution of tenascin, cellular fibronectins and integrins in the normal, hyperplastic and neoplastic breast. *J Submicrosc Cytol Pathol.* 1993;25(2):285-95.
236. Ioachim E, Charchanti A, Briasoulis E, Karavasilis V, Tsanou H, et al. Immunohistochemical expression of extracellular matrix components tenascin, fibronectin, collagen type IV and laminin in breast cancer: their prognostic value and role in tumour invasion and progression. *Eur J Cancer.* 2002;38(18):2362-70.
237. Attieh Y, Vignjevic DM. The hallmarks of CAFs in cancer invasion. *Eur J Cell Biol.* 2016;95(11):493-502.
238. Schaffner F, Ray AM, Dontenwill M. Integrin  $\alpha 5 \beta 1$ , the Fibronectin Receptor, as a Pertinent Therapeutic Target in Solid Tumors. *Cancers (Basel).* 2013;5(1):27-47.
239. Gopal S, Veracini L, Grall D, Butori C, Schaub S, et al. Fibronectin-guided migration of carcinoma collectives. *Nature Communications.* 2017;8(1):14105.
240. Saito S, Yamaji N, Yasunaga K, Saito T, Matsumoto S, et al. The fibronectin extra domain A activates matrix metalloproteinase gene expression by an interleukin-1-dependent mechanism. *J Biol Chem.* 1999;274(43):30756-63.
241. Meng XN, Jin Y, Yu Y, Bai J, Liu GY, et al. Characterisation of fibronectin-mediated FAK signalling pathways in lung cancer cell migration and invasion. *Br J Cancer.* 2009;101(2):327-34.
242. O'Hare MJ, Bond J, Clarke C, Takeuchi Y, Atherton AJ, et al. Conditional immortalization of freshly isolated human mammary fibroblasts and endothelial cells. *Proc Natl Acad Sci U S A.* 2001;98(2):646-51.
243. Nash CE, Mavria G, Baxter EW, Holliday DL, Tomlinson DC, et al. Development and characterisation of a 3D multi-cellular in vitro model of normal human breast: a tool for cancer initiation studies. *Oncotarget.* 2015;6(15):13731-41.

244. Bartek J, Bartkova J, Kyprianou N, Lalani E-N, Staskova Z, et al. Efficient immortalization of luminal epithelial cells from human mammary gland by introduction of simian virus 40 large tumor antigen with a recombinant retrovirus. *Proceedings of the National Academy of Sciences*. 1991;88(9):3520-4.
245. Carter EP, Gopsill JA, Gomm JJ, Jones JL, Grose RP. A 3D in vitro model of the human breast duct: a method to unravel myoepithelial-luminal interactions in the progression of breast cancer. *Breast Cancer Research*. 2017;19(1):1-10.
246. Gomm JJ, Browne PJ, Coope RC, Liu QY, Buluwela L, et al. Isolation of pure populations of epithelial and myoepithelial cells from the normal human mammary gland using immunomagnetic separation with Dynabeads. *Analytical biochemistry*. 1995;226(1):91-9.
247. Gomm JJ, Coope RC, Browne PJ, Coombes RC. Separated human breast epithelial and myoepithelial cells have different growth factor requirements in vitro but can reconstitute normal breast lobuloalveolar structure. *Journal of cellular physiology*. 1997;171(1):11-9.
248. Hines WC, Yaswen P, Bissell MJ. Modelling breast cancer requires identification and correction of a critical cell lineage-dependent transduction bias. *Nature communications*. 2015;6(1):1-11.
249. Meerbrey KL, Hu G, Kessler JD, Roarty K, Li MZ, et al. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc Natl Acad Sci U S A*. 2011;108(9):3665-70.
250. Reece-Hoyes JS, Walhout AJM. Gateway Recombinational Cloning. *Cold Spring Harb Protoc*. 2018;2018(1):pdb.top094912.
251. Haring M, Offermann S, Danker T, Horst I, Peterhansel C, et al. Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods*. 2007;3:11.
252. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *methods*. 2001;25(4):402-8.
253. Varet H, Brillet-Guéguen L, Coppée J-Y, Dillies M-A. SARTools: a DESeq2-and EdgeR-based R pipeline for comprehensive differential analysis of RNA-Seq data. *PloS one*. 2016;11(6):e0157022.
254. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics*. 2007;23(23):3251-3.
255. Krämer A, Green J, Pollard Jr J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*. 2014;30(4):523-30.
256. Kleifeld O, Doucet A, Prudova A, auf dem Keller U, Gioia M, et al. Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates. *Nat Protoc*. 2011;6(10):1578-611.
257. Madzharova E, Sabino F, Auf dem Keller U. Exploring Extracellular Matrix Degradomes by TMT-TAILS N-Terminomics. *Methods in molecular biology*. 2019;1944:115-26.
258. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*. 2018;47(D1):D607-D13.
259. Sflomos G, Dormoy V, Metsalu T, Jeitziner R, Battista L, et al. A preclinical model for ER $\alpha$ -positive breast cancer points to the epithelial microenvironment as determinant of luminal phenotype and hormone response. *Cancer cell*. 2016;29(3):407-22.
260. Debnath J, Muthuswamy SK, Brugge JS. Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods*. 2003;30(3):256-68.
261. Petersen OW, Rønnev-Jessen L, Howlett AR, Bissell MJ. Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. *Proc Natl Acad Sci U S A*. 1992;89(19):9064-8.
262. DiRenzo J, Signoretti S, Nakamura N, Rivera-Gonzalez R, Sellers W, et al. Growth factor requirements and basal phenotype of an immortalized mammary epithelial cell line. *Cancer research*. 2002;62(1):89-98.

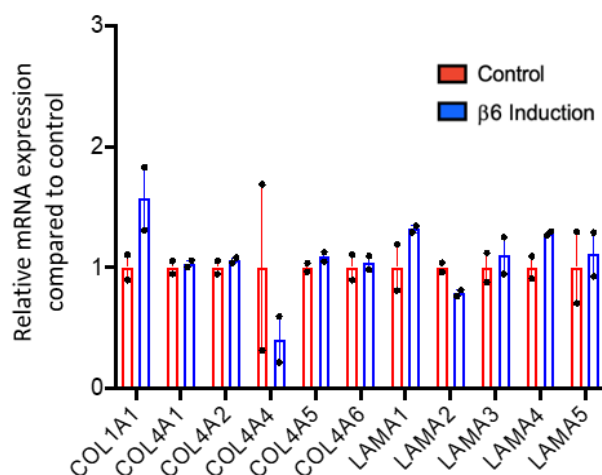


263. Yusuf R, Frenkel K. Morphologic transformation of human breast epithelial cells MCF-10A: dependence on an oxidative microenvironment and estrogen/epidermal growth factor receptors. *Cancer Cell International*. 2010;10(1):1-16.
264. Wisdom BJ, Jr., Gunwar S, Hudson MD, Noelken ME, Hudson BG. Type IV collagen of Engelbreth-Holm-Swarm tumor matrix: identification of constituent chains. *Connect Tissue Res*. 1992;27(4):225-34.
265. Friedl P, Locker J, Sahai E, Segall JE. Classifying collective cancer cell invasion. *Nature cell biology*. 2012;14(8):777-83.
266. Friedl P, Wolf K. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nature Reviews Cancer*. 2003;3(5):362-74.
267. Wang X, Enomoto A, Asai N, Kato T, Takahashi M. Collective invasion of cancer: Perspectives from pathology and development. *Pathology international*. 2016;66(4):183-92.
268. Mørk C, van Deurs B, Petersen OW. Regulation of vimentin expression in cultured human mammary epithelial cells. *Differentiation*. 1990;43(2):146-56.
269. De Wever O, Demetter P, Mareel M, Bracke M. Stromal myofibroblasts are drivers of invasive cancer growth. *International journal of cancer*. 2008;123(10):2229-38.
270. Cheung KJ, Gabrielson E, Werb Z, Ewald AJ. Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell*. 2013;155(7):1639-51.
271. Labernadie A, Kato T, Brugués A, Serra-Picamal X, Derzsi S, et al. A mechanically active heterotypic E-cadherin/N-cadherin adhesion enables fibroblasts to drive cancer cell invasion. *Nature cell biology*. 2017;19(3):224-37.
272. Chang J, Chaudhuri O. Beyond proteases: Basement membrane mechanics and cancer invasion. *Journal of Cell Biology*. 2019;218(8):2456-69.
273. Wolf K, Te Lindert M, Krause M, Alexander S, Te Riet J, et al. Physical limits of cell migration: control by ECM space and nuclear deformation and tuning by proteolysis and traction force. *Journal of Cell Biology*. 2013;201(7):1069-84.
274. Arihiro K, Kaneko M, Fujii S, Inai K, Yokosaki Y. Significance of  $\alpha 9 \beta 1$  and  $\alpha v \beta 6$  integrin expression in breast carcinoma. *Breast cancer*. 2000;7(1):19-26.
275. Reader CS, Vallath S, Steele CW, Haider S, Brentnall A, et al. The integrin  $\alpha v \beta 6$  drives pancreatic cancer through diverse mechanisms and represents an effective target for therapy. *The Journal of pathology*. 2019;249(3):332-42.
276. Thomas GJ, Hart IR, Speight PM, Marshall JF. Binding of TGF- $\beta 1$  latency-associated peptide (LAP) to  $\alpha v \beta 6$  integrin modulates behaviour of squamous carcinoma cells. *British journal of cancer*. 2002;87(8):859-67.
277. Carter EP, Roozitalab R, Gibson SV, Grose RP. Tumour microenvironment 3D-modelling: Simplicity to complexity and back again. *Trends in Cancer*. 2021;7(11):1033-46.
278. Guo W, Giancotti FG. Integrin signalling during tumour progression. *Nature Reviews Molecular Cell Biology*. 2004;5(10):816-26.
279. Engel CK, Pirard B, Schimanski S, Kirsch R, Habermann J, et al. Structural basis for the highly selective inhibition of MMP-13. *Chem Biol*. 2005;12(2):181-9.
280. Annes JP, Chen Y, Munger JS, Rifkin DB. Integrin  $\alpha v \beta 6$ -mediated activation of latent TGF- $\beta$  requires the latent TGF- $\beta$  binding protein-1. *The Journal of cell biology*. 2004;165(5):723-34.
281. Shi M, Zhu J, Wang R, Chen X, Mi L, et al. Latent TGF- $\beta$  structure and activation. *Nature*. 2011;474(7351):343-9.
282. Sheppard D. Integrin-mediated activation of latent transforming growth factor  $\beta$ . *Cancer and Metastasis Reviews*. 2005;24(3):395-402.
283. Conery AR, Centore RC, Neiss A, Keller PJ, Joshi S, et al. Bromodomain inhibition of the transcriptional coactivators CBP/EP300 as a therapeutic strategy to target the IRF4 network in multiple myeloma. *Elife*. 2016;5.

284. Tang W, Zhou M, Dorsey TH, Prieto DA, Wang XW, et al. Integrated proteotranscriptomics of breast cancer reveals globally increased protein-mRNA concordance associated with subtypes and survival. *Genome medicine*. 2018;10(1):1-14.
285. Luo X-L, Lin L, Hu H, Hu F-L, Lin Y, et al. Development and characterization of mammary intraductal (MIND) spontaneous metastasis models for triple-negative breast cancer in syngeneic mice. *Scientific reports*. 2020;10(1):1-11.
286. Nielsen BS, Rank F, López JM, Balbin M, Vizoso F, et al. Collagenase-3 expression in breast myofibroblasts as a molecular marker of transition of ductal carcinoma in situ lesions to invasive ductal carcinomas. *Cancer research*. 2001;61(19):7091-100.
287. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell*. 2018;173(2):291-304.
288. Buck MB, Fritz P, Dippon J, Zugmaier G, Knabbe C. Prognostic significance of transforming growth factor  $\beta$  receptor II in estrogen receptor-negative breast cancer patients. *Clinical Cancer Research*. 2004;10(2):491-8.
289. Lee M, Partridge NC. Parathyroid hormone activation of matrix metalloproteinase-13 transcription requires the histone acetyltransferase activity of p300 and PCAF and p300-dependent acetylation of PCAF. *Journal of Biological Chemistry*. 2010;285(49):38014-22.
290. Boumah CE, Lee M, Selvamurugan N, Shimizu E, Partridge NC. Runx2 recruits p300 to mediate parathyroid hormone's effects on histone acetylation and transcriptional activation of the matrix metalloproteinase-13 gene. *Molecular Endocrinology*. 2009;23(8):1255-63.
291. Bekhouche M, Colige A. The procollagen N-proteinases ADAMTS2, 3 and 14 in pathophysiology. *Matrix Biol*. 2015;44-46:46-53.
292. Janssen L, Dupont L, Bekhouche M, Noel A, Leduc C, et al. ADAMTS3 activity is mandatory for embryonic lymphangiogenesis and regulates placental angiogenesis. *Angiogenesis*. 2016;19(1):53-65.
293. Jha SK, Rauniyar K, Karpanen T, Leppänen VM, Brouillard P, et al. Efficient activation of the lymphangiogenic growth factor VEGF-C requires the C-terminal domain of VEGF-C and the N-terminal domain of CCBE1. *Sci Rep*. 2017;7(1):4916.
294. Bekhouche M, Leduc C, Dupont L, Janssen L, Delolme F, et al. Determination of the substrate repertoire of ADAMTS2, 3, and 14 significantly broadens their functions and identifies extracellular matrix organization and TGF- $\beta$  signaling as primary targets. *Faseb j*. 2016;30(5):1741-56.
295. Abba MC, Gong T, Lu Y, Lee J, Zhong Y, et al. A Molecular Portrait of High-Grade Ductal Carcinoma In Situ. *Cancer Res*. 2015;75(18):3980-90.
296. Russell TD, Jindal S, Agunbiade S, Gao D, Troxell M, et al. Myoepithelial cell differentiation markers in ductal carcinoma in situ progression. *Am J Pathol*. 2015;185(11):3076-89.
297. Kanjanapradit K, Wangsawibul S. P-cadherin and CD10 Expression to Distinguish between Ductal Carcinoma in Situ and Invasive Ductal Carcinoma of the Breast. 2019. 2019;37(1):8.
298. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
299. Artym VV, Yamada KM, Mueller SC. ECM degradation assays for analyzing local cell invasion. *Methods in molecular biology*. 2009;522:211-9.
300. Summerbell ER, Mouw JK, Bell JSK, Knippler CM, Pedro B, et al. Epigenetically heterogeneous tumor cells direct collective invasion through filopodia-driven fibronectin micropatterning. *Sci Adv*. 2020;6(30):eaaz6197.
301. Efthymiou G, Saint A, Ruff M, Rekad Z, Ciais D, et al. Shaping Up the Tumor Microenvironment With Cellular Fibronectin. *Front Oncol*. 2020;10:641.
302. Kapp TG, Rechenmacher F, Neubauer S, Maltsev OV, Cavalcanti-Adam EA, et al. A Comprehensive Evaluation of the Activity and Selectivity Profile of Ligands for RGD-binding Integrins. *Scientific Reports*. 2017;7(1):39805.
303. Paule S, Aljofan M, Simon C, Rombauts LJF, Nie G. Cleavage of endometrial  $\alpha$ -integrins into their functional forms is mediated by proprotein convertase 5/6. *Human Reproduction*. 2012;27(9):2766-74.

304. Bergeron E, Basak A, Decroly E, Seidah NG. Processing of alpha4 integrin by the proprotein convertases: histidine at position P6 regulates cleavage. *The Biochemical journal*. 2003;373(Pt 2):475-84.
305. Lissitzky JC, Luis J, Munzer JS, Benjannet S, Parat F, et al. Endoproteolytic processing of integrin pro-alpha subunits involves the redundant function of furin and proprotein convertase (PC) 5A, but not paired basic amino acid converting enzyme (PACE) 4, PC5B or PC7. *The Biochemical journal*. 2000;346 Pt 1(Pt 1):133-8.
306. Clark K, Pankov R, Travis MA, Askari JA, Mould AP, et al. A specific alpha5beta1-integrin conformation promotes directional integrin translocation and fibronectin matrix formation. *J Cell Sci*. 2005;118(Pt 2):291-300.
307. Jha SK, Rauniyar K, Karpanen T, Leppänen V-M, Brouillard P, et al. Efficient activation of the lymphangiogenic growth factor VEGF-C requires the C-terminal domain of VEGF-C and the N-terminal domain of CCBE1. *Scientific Reports*. 2017;7(1):4916.
308. Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer*. 2005;103(12):2481-4.
309. Aydemir AT, Alper M, Kockar F. SP1-mediated downregulation of ADAMTS3 gene expression in osteosarcoma models. *Gene*. 2018;659:1-10.
310. Jin Z, Zhou S, Ye H, Jiang S, Yu K, et al. The mechanism of SP1/p300 complex promotes proliferation of multiple myeloma cells through regulating IQGAP1 transcription. *Biomed Pharmacother*. 2019;119:109434.
311. Jungert K, Buck A, von Wichert G, Adler G, König A, et al. Sp1 is required for transforming growth factor-beta-induced mesenchymal transition and migration in pancreatic cancer cells. *Cancer Res*. 2007;67(4):1563-70.
312. Hong Y, Limback D, Elsarraj HS, Harper H, Haines H, et al. Mouse-INtraDuctal (MIND): an in vivo model for studying the underlying mechanisms of DCIS malignancy. *J Pathol*. 2022;256(2):186-201.

## Appendix

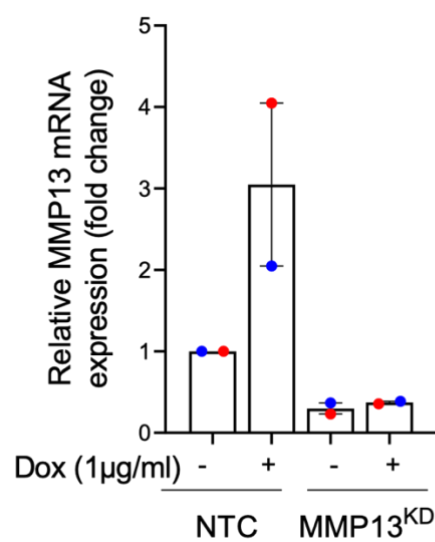


### Appendix 1.1 Matrix expression changes following integrin β6 induction in 1089<sup>iβ6</sup> cells.

Relative expression of laminin and collagen genes compared to control from RNAseq normalised counts in 1089<sup>iβ6</sup> cells. Data are presented as mean ± SEM. Data are representative of two independent experiments.

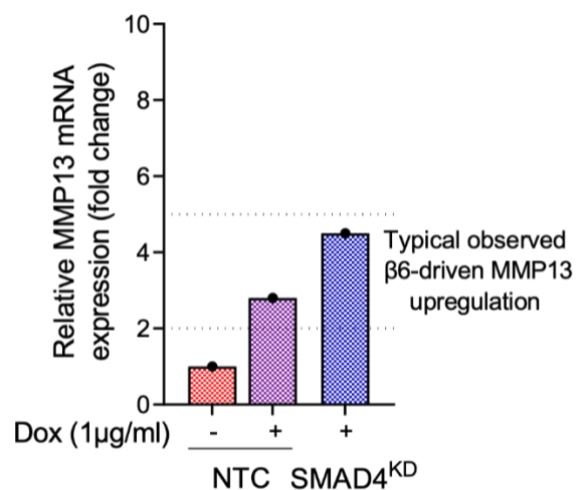
**Appendix 1.2 HB2 monoculture spheroids.**

HB2 monoculture spheroids are made via methylcellulose hanging drops. Spheroids are embedded into a 4 mg/ml collagen gel. Representative image of HB2 monoculture sphere 4 days post embedding.



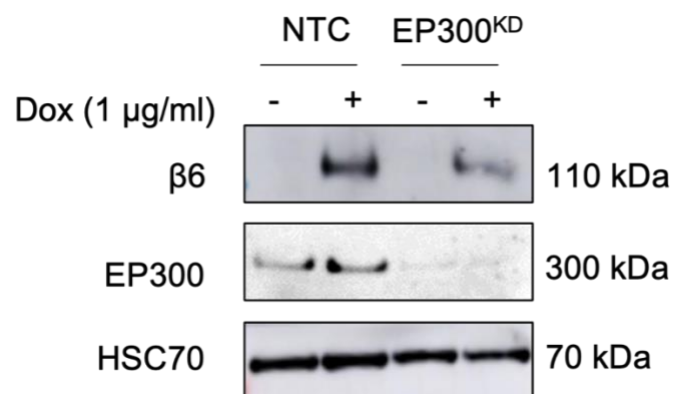
### Appendix 2.1 Validation of MMP13 siRNA knockdown in 1089<sup>iB6</sup> cells.

MMP13 mRNA expression in doxycycline (1 µg/ml) treated 1089<sup>iB6</sup> cells 72 h post transfection with either non-targeting control (NTC) or MMP13 siRNA (MMP13<sup>KD</sup>).



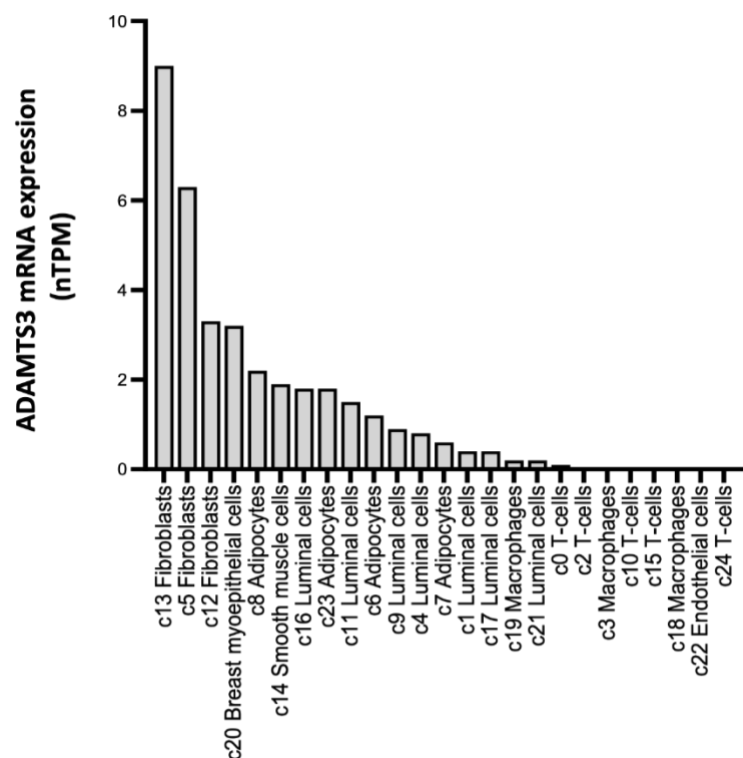
#### Appendix 2.2 MMP13 expression following SMAD4 siRNA knockdown in 1089<sup>β6</sup> cells.

Relative MMP13 mRNA expression 72 h following siRNA knock-down of SMAD4 (SMAD4<sup>KD</sup>) in doxycycline (1 µg/ml) treated 1089<sup>β6</sup> cells compared to non targeting control (NTC).

**Appendix 2.3 Validation of EP300 siRNA knockdown in 1089<sup>iβ6</sup> cells.**

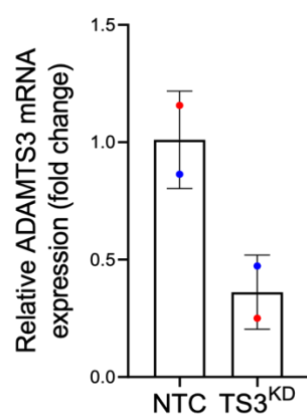
Western blot showing integrin β6 and EP300 expression in doxycycline (1 µg/ml) treated 1089<sup>iβ6</sup> cells 72 h post transfection with either NTC or EP300 siRNA (EP300<sup>KD</sup>).





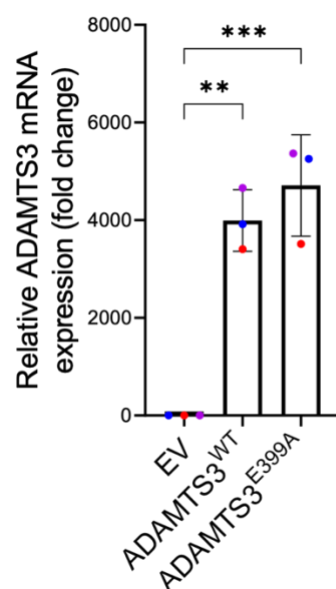
### Appendix 3.1 ADAMTS3 expression in cell types of the human breast.

ADAMTS3 mRNA expression, represented as normalised transcripts per million (nTPM), in different cellular compartments of the healthy breast. Data from ProteinAtlas single-cell RNAseq analysis of the healthy breast (GSE164898).



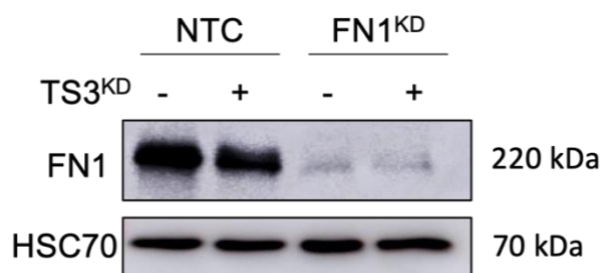
### Appendix 3.2 Validation of ADAMTS3 siRNA knockdown in 1089 cells.

ADAMTS3 mRNA expression in 1089 cells 72 h post transfection with either non-targeting control (NTC) or ADAMTS3 siRNA (TS3<sup>KD</sup>).

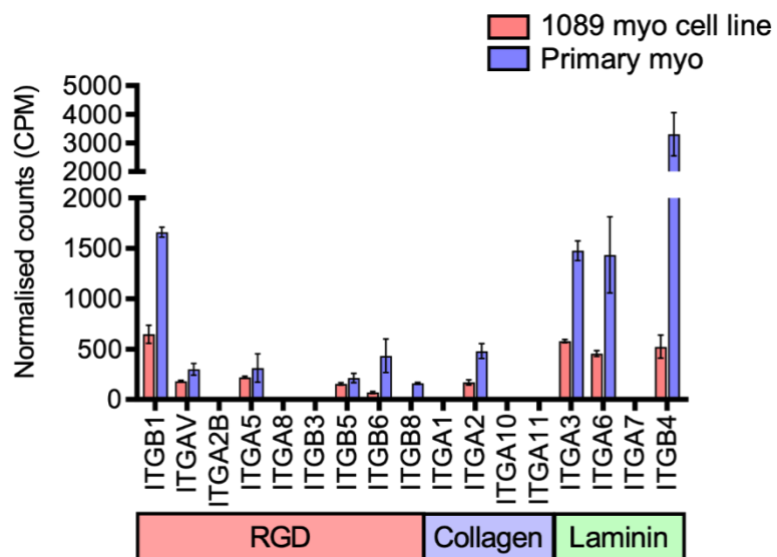


### Appendix 3.3 Validation of ADAMTS3 expression constructs in 1089 cells.

Relative ADAMTS3 mRNA expression in the 1089 myoepithelial cell line, transduced with empty vector (EV), ADAMTS3 wildtype (ADAMTS3<sup>WT</sup>) and inactive ADAMTS3 (ADAMTS3<sup>E399A</sup>) constructs. \*\* $p < 0.01$ , \*\*\* $p < 0.001$  (One-way ANOVA with multiple comparisons).

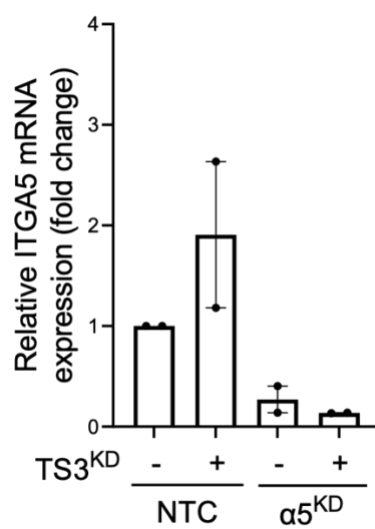
**Appendix 3.4 Validation of fibronectin siRNA knockdown in 1089 cells.**

Western blot showing fibronectin (FN1) expression in 1089 myoepithelial cells 72 h post-transfection with non-targeting control (NTC), ADAMTS3 (TS3<sup>KD</sup>) and/or FN1 siRNA (FN1<sup>KD</sup>).



### Appendix 3.5 Integrin expression in 1089 and primary myoepithelial cells.

mRNA expression of RGD, collagen and laminin integrins represented as normalised counts per million (CPM) across 1089 cell line and primary myoepithelial cells.



### Appendix 3.6 Validation of integrin $\alpha 5$ siRNA knockdown in 1089 cells.

Relative integrin  $\alpha 5$  mRNA expression in 1089 myoepithelial cells 72 h post transfection with non-targeting control (NTC), ADAMTS3 (TS3<sup>KD</sup>) and/or  $\alpha 5$  siRNA ( $\alpha 5^{KD}$ ).