

# Investigating the role of versican in immune exclusion in triple negative breast cancer

Priyanka Karsan Hirani

Thesis submitted in partial fulfilment of the requirements of the Degree of

**Doctor of Philosophy (PhD)**

at Queen Mary University of London

Centre for Tumour Microenvironment

Barts Cancer Institute

Queen Mary University of London,

Charterhouse Square

EC1M 6BQ

This work was supported by Against Breast Cancer and Barts Charity  
(MGU0499)

## Statement of originality

I, Priyanka Karsan Hirani, 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 Queen Mary University of London has the right to use plagiarism detection 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: Priyanka Karsan Hirani

Date: 27/10/2023

## Details of Collaborations

- CS isomer analysis was completed by Dr Kim Alonge at the University of Washington
- Bulk RNA-seq was completed by Dr Ying Liu and analysed by Dr Eleni Maniati
- TNBC tissue peptide processing and mass spectrometry analysis was completed by the Mass spectrometry facility (Dr Vinothini Rajeeve)
- Mass spectrometry analysis of VCAN enrichment samples was completed by the Mass spectrometry facility (Dr Vinothini Rajeeve)
- 0GAG, 2GAGS507G and 2GAGS525G plasmids were provided by Dr Salvatore Santamaria
- Patient tissue samples and cells were provided by the Breast Cancer Now tissue bank
- Gels were paraffin embedded and sections of tissues and gels were cut by the Pathology facility (Nadia Rahman).

## Details of Presentations and Publications

### Presentations:

- BCI PhD day – July 2023 – Poster
- BSMB Early Careers Meeting – June 2023 – Short talk – Prize for best short talk
- AACR Annual Meeting – April 2023 – Poster
- UK Proteoglycan Meeting – January 2023 – Poster
- Crick Cancer Conference – October 2022 – Poster
- Matrix Biology Europe – September 2022 – Short talk
- BCI PhD day – September 2022 – Poster
- EACR: The structural microenvironment – February 2022 – Poster
- UK Proteoglycan Meeting – January 2022 – Poster/Flash talk – Prize for best poster figures
- BCI PhD day – September 2021 – Poster
- BSMB winter meeting – September 2021 – Poster
- London PhD Network Conference – June 2021 – Poster
- UK Proteoglycan Meeting – January 2021 – Flash talk

### Publications:

- Puttock EH, Tyler EJ, Manni M, Maniati E, Butterworth C, Burger Ramos M, Peerani E, **Hirani P**, Gauthier V, Liu Y, Maniscalco G, Rajeeve V, Cutillas P, Trevisan C, Pozzobon M, Lockley M, Rastrick J, Läubli H, White A, Pearce OMT. Extracellular matrix educates an immunoregulatory tumor macrophage phenotype found in ovarian cancer metastasis. *Nat Commun.* 2023 May 15;14(1):2514. doi: 10.1038/s41467-023-38093-5. PMID: 37188691; PMCID: PMC10185550.
- **Hirani P**, Gauthier V, Allen CE, Wight TN, Pearce OMT. Targeting Versican as a Potential Immunotherapeutic Strategy in the Treatment of Cancer. *Front Oncol.* 2021 Aug 30;11:712807. doi: 10.3389/fonc.2021.712807. PMID: 34527586; PMCID: PMC8435723.
- Maniati E, Berlato C, Gopinathan G, Heath O, Kotantaki P, Lakhani A, McDermott J, Pegrum C, Delaine-Smith RM, Pearce OMT, **Hirani P**, Joy JD, Szabova L, Perets R, Sansom OJ, Drapkin R, Bailey P, Balkwill FR. Mouse Ovarian Cancer Models Recapitulate the Human Tumor Microenvironment



and Patient Response to Treatment. Cell Rep. 2020 Jan 14;30(2):525-540.e7. doi: 10.1016/j.celrep.2019.12.034. PMID: 31940494; PMCID: PMC6963791.

Paper in review:

- **Hirani P**, McDermott J, Rajeeve V, Cutillas P, Jones LJ, Pennington D, Wight TN, Santamaria S, Alonge K, Pearce OMT. The proteoglycan versican associates with tumour immune phenotype and limits T-cell trafficking via post-translational modifications of chondroitin sulphate.

## Abstract

Triple negative breast cancer has the highest T cell infiltrate in comparison to other subtypes of breast cancer. To try to improve the anti-tumour response of these T cells, immunotherapy has been trialled, however clinical trials showed poor results. The response to immunotherapy in solid tumours is limited and this has been attributed to the presence of the extracellular matrix (ECM). The ECM can interact with T cells biochemically or physically, affecting their trafficking in the tumour. This can cause the restriction of T cells in the stroma limiting their contact with the tumour epithelial cells, leading to an immune excluded phenotype. Identifying key components of the ECM that are associated with the restriction of immune cells can provide potential targets that could be degraded to improve anti-tumour immunity.

From previous work in the lab a signature of molecules were identified which were associated with immunosuppression. In the initial analysis of these molecules in a subset of TNBC tissues, versican (VCAN) was identified as an ECM component that associates with immune cell infiltration into the tumour epithelium. VCAN is a proteoglycan which has the glycosaminoglycan chondroitin sulphate (CS) attached to the peptide backbone. Through its multiple domains and glycan post-translational modifications, VCAN has been shown to have a role in inflammation and cancer progression.

To study how VCAN may affect the trafficking of T cells, I first looked at how VCAN expression associated with immune excluded tissues. It was observed that VCAN levels were higher in the epithelial zone of excluded tissues compared to inflamed tissues. CS levels were then explored within the tissues where the sulphation patterns on CS in the stroma led to the discovery of CS-C being higher in excluded tissues and CS-A being higher in inflamed tissues. To observe this effect in-vitro, VCAN was enriched from TNBC and fibroblast cell line secretions. The effect of CS was tested through chondroitinase (CSase) treatment of VCAN enriched protein in a transwell model. An increase in invasion was observed following CSase treatment of protein with high levels of CS-C.

To conclude, from the study I identified that within TNBC tissues the excluded immune phenotype associates with epithelial zone expressed VCAN which has a different CS sulphation pattern compared to inflamed tissues, and this difference in sulphation inhibits T-cell trafficking in in vitro models, which can be overcome through enzymatic digestion of the CS. Therefore, targeting VCAN by degrading CS

may provide a way to drive excluded tumours into an inflamed and therapy responsive phenotype. Such an approach could be coupled with immunotherapy such as cell-based T-cell therapies.

## Acknowledgements

Firstly, I would like to thank my supervisor Dr Oliver Pearce for the guidance and support you have offered throughout this PhD. You have always been there to encourage me at every stage of the project and inspire me with ideas whenever I got stuck. Your passion for the project continued to motivate me and your patience taught me to never give up. It has been a privilege to complete my PhD under your mentorship.

I would like to thank everyone in the Pearce Lab past and present. It has been an honour working with you all over the last 4 years. I am grateful for all the help, advice and insights you have all offered and for the memories in and outside of the lab. I would also like to thank everyone in the Centre of the Tumour Microenvironment. From all the office laughs to the struggles in the lab, you have made this a memorable journey.

I would like to give a special thanks to Dr Beatrice Malacrida, Dr Ganga Gopinathan and Dr Joash Joy for being the best mentors, friends and role models. You have all been there at the ups and downs of this project, reassured me when things went wrong and celebrated all my achievements. The last 4 years would have not been the same without you.

I would also like to thank to all the collaborators on this project. Dr Kimberly Alonge, for taking the time to allow me to spend some time in her lab, teaching me new techniques and sharing your knowledge. Prof Thomas Wight, for the support and enthusiasm you have shown for this project. And Dr Salvatore Santamaria, for all the advice and support.

I would like to give a special mention to the patients who have kindly donated tissue samples making this work possible.

And lastly I would like to thank my family to whom I would like to dedicate this thesis to. My Mum and Dad who have always been supportive and encouraging. I will always be grateful for their unconditional love and for inspiring me to keep going. Neelam and Dhillon, for their continuous support even from the other side of the world.

॥ जय सियराम ॥

# Table of Contents

Chapter 1: Introduction .....	22
1.1. The extracellular matrix .....	23
1.1.1 Main components of the ECM .....	23
1.1.2 Tumour ECM remodelling .....	24
1.2. Matrix and tumour immunity .....	25
1.2.1 ECM and immune phenotypes .....	26
1.2.2 ECM and immune cell migration .....	27
1.3. Tumour immune phenotype .....	28
1.3.1 Tumour immunity cycle .....	28
1.3.2 Tumour immune phenotype classifications .....	29
1.3.3 Determination of tumour immune phenotypes .....	30
1.3.4 Immune phenotype and immunotherapy response .....	32
1.4. Targeting the ECM for immunotherapy .....	33
1.5. VCAN .....	34
1.5.1 VCAN structure and isoforms .....	34
1.5.2 Chondroitin sulphate .....	35
1.5.3 VCAN interactions with the ECM .....	37
1.5.4 VCAN expression .....	39
1.6. VCAN and cancer .....	41
1.6.1 Cells expressing VCAN .....	41
1.6.2 Impact on tumour prognosis .....	42
1.6.3 VCAN isoforms in cancer .....	43
1.7. VCAN and immunity .....	44
1.7.1 VCAN interactions with chemokines .....	44
1.7.2 VCAN interactions with immune cells .....	45
1.8. Breast cancer .....	46
1.8.1 Risk factors .....	46

1.8.2 Subtypes of breast cancer .....	48
1.9. Triple negative breast cancer .....	49
1.9.1 Subtypes of TNBC .....	49
1.9.2 Treatment .....	51
1.9.3 Immunotherapy .....	52
1.10. Hypothesis and Aims .....	54
Chapter 2: Materials and Methods .....	55
2.1. Cell lines and culture .....	56
2.2. Immunohistochemistry .....	58
2.2.1 Single stain .....	58
2.2.2 Dual stain .....	59
2.3. Masson's trichrome .....	60
2.4. RNAscope .....	61
2.4.1 Single .....	61
2.4.2 Duplex .....	62
2.5. Image analysis .....	62
2.5.1 Stain deconvolution .....	62
2.5.2 Co-localisation analysis .....	63
2.5.3 Image overlay .....	63
2.5.4 Positive cell counting .....	63
2.5.5 Stain area detection .....	63
2.5.6 Area classification .....	64
2.5.7 Tissue phenotype analysis .....	64
2.5.8 Immune exclusion analysis .....	64
2.6. Gel models .....	65
2.6.1 Collagen gel model .....	65
2.6.2 Embedded gel model .....	65
2.7. RNA extraction .....	66
2.8. qRT-PCR .....	67

2.8.1 cDNA transformation.....	67
2.8.2 qRTPCR .....	67
2.8.3 Agarose gel electrophoresis for DNA .....	69
2.9. VCAN enrichment.....	69
2.9.1 Affinity chromatography .....	70
2.9.2 Ion exchange .....	71
2.10. Dot blot.....	71
2.11. Protein concentration using BCA.....	72
2.12. Gel electrophoresis .....	72
2.13. PBMC isolation from leukocyte cones .....	72
2.13.1 T cell isolation from PBMC.....	73
2.14. T cell activation.....	73
2.15. Flow cytometry .....	74
2.16. Transwell assay.....	74
2.17. Gel invasion assay .....	75
2.18. Mass spectrometry .....	75
2.19. CS Mass spectrometry .....	77
2.19.1 Tissue sections .....	77
2.19.2 From cell culture .....	77
2.20. Cell bacterial transformation.....	78
2.20.1 Plasmid expansion.....	78
2.20.2 Enzyme restriction of plasmids.....	79
2.20.3 Antibiotic survival test .....	80
2.20.4 Transient transfection .....	80
2.21. Gene data analysis.....	81
2.22. Statistical analysis .....	81
Chapter 3: Characterising the tumour immune infiltrate and matrix expression in TNBC tissues.....	82
3.1. Introduction .....	83

3.1.1. Chapter highlights at a glance: .....	83
3.2. Identifying TIPs in TNBC tissues .....	84
3.3. Expression of matrix index molecules in TNBC .....	93
3.4. Expression of VCAN and its associated matrix molecules in comparison to CD8+ and CD68+ immune cells. ....	100
3.5. Immune exclusion associates with VCAN expression in the epithelial zone .....	105
3.6. Tumour cells and fibroblasts express VCAN in TNBC .....	114
3.7. Discussion.....	122
Chapter 4: The expression and arrangement of chondroitin sulphate proteoglycans in the TME .....	127
4.1. Introduction .....	128
4.1.1. Chapter highlights at a glance:.....	128
4.2. Localisation of CS in TNBC .....	129
4.3. CS proteoglycans in TNBC.....	134
4.4. Association of BGN and DCN with tissue immune phenotypes.....	138
4.5. Effect of CS isomer on immune phenotype.....	144
4.6. Discussion.....	150
Chapter 5: Enrichment of VCAN from a TNBC and a human mammary fibroblast cell line.....	154
5.1. Introduction .....	155
5.1.1 Chapter highlights at a glance:.....	155
5.2. VCAN enrichment via HA-affinity chromatography .....	156
5.3. Optimising media for VCAN enrichment .....	162
5.4. Variations in elutions from different cell lines .....	166
5.5. Ion exchange chromatography .....	170
5.6. Determining level of VCAN enrichment .....	176
5.7. Proteoglycans in enrichments.....	182
5.8. Discussion.....	184



Chapter 6: Exploring the structural variations of VCAN and their role in T cell trafficking. ....	187
6.1. Introduction .....	188
6.1.1. Chapter highlights at a glance: .....	188
6.2. Exploring VCAN expression in 3D .....	189
6.3. Building a tumour-stroma 3D model. ....	191
6.4. Effect of VCAN and T cell invasion .....	198
6.4.1 Invasion in a transwell model .....	198
6.4.2 Invasion in a gel .....	208
6.5. Effect of VCAN degradation towards T cell invasion .....	211
6.6. Discussion.....	219
Chapter 7: Discussion.....	222
7.1. Exploring TIPs in TNBC.....	223
7.2. VCAN expression around the epithelium associates with TIP .....	223
7.3. CS proteoglycans and immune exclusion .....	224
7.4. Isoforms of CS associate with TIP .....	225
7.5. Removal of CS-C improves T cell invasion.....	226
7.6. VCAN degradation affects T cell invasion.....	227
7.7. Conclusions.....	227
7.8. Impact of study .....	228
7.9. Future work .....	229
Bibliography .....	231
Appendices.....	250

# Table of Figures

Figure 1.1. Tumour ECM.....	25
Figure 1.2. The tumour immunity cycle. ....	29
Figure 1.3. Tumour immune phenotypes. ....	30
Figure 1.4. Breaking down the matrix to improve immune cell infiltration. ....	33
Figure 1.5. Structure of VCAN.....	35
Figure 1.6. Formation of CS linkage to serine. ....	36
Figure 1.7. Structure of CS isomers.....	37
Figure 1.8. Binding partners of VCAN at the different domains. ....	39
Figure 1.9 Chemokine binding to VCAN facilitating extravasation. ....	46
Figure 1.10. Subtypes of breast cancer. ....	49
Figure 1.11. TNBC subtypes. ....	51
Figure 2.1. TNBC cell lines HCC38 and MDAMB468 and immortalised mammary fibroblast HMF3S. .....	57
Figure 2.2. Fibroblasts from healthy patients. ....	57
Figure 2.3. Amplified products from qRTPCR.....	69
Figure 2.4. Enzyme restriction of V1 plasmid.....	79
Figure 2.5. Drug dilution test of cells. ....	80
Figure 3.1. Analysing the tumour epithelial border. ....	85
Figure 3.2. Spread of immune cells across regions. ....	86
Figure 3.3. Exclusion analysis using the Kather et al method.....	88
Figure 3.4. Exclusion analysis with adapted method. ....	91
Figure 3.5. Immune phenotypes within tissues. ....	93
Figure 3.6. Expression of overexpressed matrix index markers in TNBC. ....	96
Figure 3.7. Co-localisation analysis of matrix index proteins. ....	98
Figure 3.8. Dual RNAscope of CTSB and VCAN.....	99
Figure 3.9. Phenotyping tissues for matrix and immune cells.....	101
Figure 3.10. Comparison of matrix markers and CD8+ T cells and CD68+ macrophages. ....	102
Figure 3.11. Tumour epithelium-stroma ratio of immune cells and the matrix.....	104
Figure 3.12. VCAN expression and immune phenotype. ....	107
Figure 3.13. Representative images of immune phenotypes and VCAN expression. ....	108
Figure 3.14. CD44 expression and immune phenotype.....	110
Figure 3.15. Comparisons of matrix markers in CD8 excluded and inflamed tissues. ....	111
Figure 3.16 HA and VCAN in immune phenotypes.....	113
Figure 3.17. Comparing VCAN expression in tumour cells and fibroblasts. ....	115
Figure 3.18. RNAscope of excluded and inflamed tissues. ....	116
Figure 3.19. Expression of VCAN isoforms in tissues and cell lines. ....	118

Figure 3.20. Effect of cell cross-talk on VCAN expression. ....	121
Figure 3.21. Summary of excluded and inflamed tissue phenotypes. ....	126
Figure 4.1. CS expression across the tumour epithelial border. ....	129
Figure 4.2. Co-localisation of CS and VCAN in excluded tissues. ....	131
Figure 4.3. Co-localisation of CS and VCAN in inflamed tissues. ....	132
Figure 4.4. Correlation between VCAN and CS in the different regions of excluded and inflamed tissues.....	133
Figure 4.5. Expression of CS proteoglycans in TNBC. ....	135
Figure 4.6. CS proteoglycan expression in tissues and cells in mono-culture and co-culture, .....	137
Figure 4.7. Tissue immune phenotype analysis in second subset of tissues.....	139
Figure 4.8. VCAN, BGN and DCN in the different TIPs. ....	141
Figure 4.9. BGN is not associated with TIP.....	142
Figure 4.10. DCN expression is not associated with phenotype.....	143
Figure 4.11 Chromatogram of CS-MS.....	144
Figure 4.12. CS isomers A and C are differentially expressed in excluded and inflamed tissues....	147
Figure 4.13. CS isomers in cell lines. ....	149
Figure 4.14. Summary of CSPG expression in excluded and inflamed tissues and CS isomer. ....	153
Figure 5.1. Binding of HA to carboxylink beads. ....	157
Figure 5.2. IF on beads for HA. ....	158
Figure 5.3. Protein extraction and isolation procedure. ....	159
Figure 5.4. Dot blots of elutions. ....	161
Figure 5.5. Comparing cell viability and VCAN production in different media. ....	163
Figure 5.6. Enrichment process of VCAN using PFM and salt gradient with HA-affinity chromatography. ....	165
Figure 5.7. Comparison of elutions with HCC38 and HMF3S. ....	167
Figure 5.8. Mass spectrometry analysis of enrichments from HMF3S and HCC38.....	169
Figure 5.9. IEX chromatography for the enrichment of VCAN.....	171
Figure 5.10. Comparing IEX and HA- affinity chromatography.....	173
Figure 5.11. Mass spectrometry analysis comparing VCAN positive elutions from HA- affinity chromatography (HA) and IEX.....	175
Figure 5.12. BGN detections in HA-affinity chromatography and IEX. ....	176
Figure 5.13. Mascot scores of proteins in the crude and enrichments. ....	177
Figure 5.14. empAI score of proteins in the crude and enrichments.....	178
Figure 5.15. Fold change in the empAI scores of the proteins. ....	179
Figure 5.16. Quantification of VCAN by comparing to standards. ....	181
Figure 5.17. Elutions of different CS proteoglycans. ....	183
Figure 6.1. VCAN isoform expression in cells grown in 2D and 3D. ....	189
Figure 6.2. VCAN isoform expression at different time points. ....	190

Figure 6.3. Effect of treatment of HMF3S cells towards VCAN expression. ....	190
Figure 6.4. Schematic of the tumour-stroma 3D model. ....	191
Figure 6.5. Addition of crude and enriched protein to tumour gel.....	192
Figure 6.6. Procedure for growing embedded gels. ....	193
Figure 6.7. Location of tumour cells, fibroblasts and VCAN within the embedded gels. ....	194
Figure 6.8. VCAN expression in the different gel conditions. ....	195
Figure 6.9. Invasion of T cells into the embedded gel model. ....	197
Figure 6.10. T cell migration to the bottom well of the transwell. ....	200
Figure 6.11. Cells on the transwell membrane. ....	202
Figure 6.12. Effect of protein level, enzymes and HA on T cell invasion.....	203
Figure 6.13. Effect of cell specific protein and enzymes on T cell invasion. ....	204
Figure 6.14. Invasion with different T cell donors.....	206
Figure 6.15. CS isomer analysis of the crude and enriched samples of HCC38 and HMF3S. ....	207
Figure 6.16. Effect of CSase towards invasion through a collagen gel. ....	209
Figure 6.17 Effect of HMF3S co-culture with HCC38 and MDAMB468 towards invasion through a collagen gel.....	210
Figure 6.18. Comparison of versikine and VCAN in the different TIPs. ....	212
Figure 6.19. Versikine analysis in the different TIPs. ....	213
Figure 6.20. ADAMTS enzyme expression in the cell lines and in co-cultures. ....	213
Figure 6.21. Effect of CS expression on VCAN to T cell invasion. ....	215
Figure 6.22. Effect of VCAN plasmid transfection on T cell invasion. ....	217
Figure 6.23. Effect of VCAN expression on T cell invasion.....	218
Figure 6.24. Summary of T cell invasion and CS. ....	221
Figure 7.1. Effect of VCAN structure towards tumour immune phenotype.....	228

## List of Tables

Table 2-1. Immunohistochemistry antibodies. ....	60
Table 2-2. Reagents for collagen gels of different stiffness. ....	65
Table 2-3. Primer sequences for each gene detected with qRT-PCR. ....	68
Table 2-4. Buffer composition for chromatography. ....	69
Table 2-5. Antibody dilutions for immunoblotting. ....	72
Table 2-6. Mass spectrometry analysis parameters. ....	76
Table 5-1. Buffers composition. ....	160
Table 5-2. Buffers for salt gradient elutions. Composition of different buffers for elutions. ....	161
Table 5-3. Elution buffers for updated salt gradient. ....	165
Table 5-4. Elution buffers used for IEX chromatography. ....	171

## Abbreviations

3D	Three dimensional
$\alpha$ -SMA	Alpha- smooth muscle actin
ACN	Acetonitrile
ADAMTS	A disintegrin and metalloproteinase with thrombospondin motif
Batf3	Basic leucine zipper atf-like transcription factor 3
BGN	Biglycan
BL	Basal-like
C	Crude
CAF	Cancer associated fibroblast
COL11A1	Collagen Type XI $\alpha$ 1
COMP	Cartilage oligomeric matrix protein
CS	Chondroitin sulphate
CSase	Chondroitinase
CSPG	Chondroitin sulphate proteoglycan
CTSB	Cathepsin B
Ct	Cycling threshold
DAB	3,3'-diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
DCN	Decorin
DDR1	Discoidin domain receptor 1
DTT	Dithiothreitol
EB	Elution buffer
ECM	Extracellular matrix
EDC	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	Ethylenediaminetetraacetic acid

EGF-like	Epidermal growth factor-like
EGFR	Epidermal growth factor receptor
EmPAI	Exponentially modified protein abundance index
ER	Oestrogen receptor
FAP	Fibroblast activation protein
FBS	Foetal bovine serum
FGF	Fibroblast growth factor
FN1	Fibronectin
FT	Flow-through
Gal	Galactose
GalNAc	N-acetylgalactosamine
GHAP	Glial hyaluronan binding protein
GlcA	Glucuronic acid
GluC	Endoproteinase gluc
HA	Hyaluronan
HER2	Human epidermal growth factor receptor 2
HRP	Horseradish peroxidase
IEX	Ion exchange chromatography
IG-like	Immunoglobular-like
IL-	Interleukin-
IM	Immunomodulatory
IRF-8	Interferon regulatory factor 8
ITH	Intratumoural heterogeneity
LAR	Luminal androgen receptor
LOX	Lysyl oxidases
LPS	Lipopolysaccharide
M	Mesenchymal

MES	2-(N-morpholino)ethanesulfonic acid
MMP	Matrix metalloproteases
MS water	Mass spectrometry grade water
MSL	Mesenchymal stem like
NK	Natural killer
P/S	Penicillin/streptomycin
PanCK	Pancytokeratin
PARP	Poly ADP-ribose polymerase
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline – Tween 20
PD-1	Programmed cell death protein 1
PDGF	Platelet derived growth factor
PD-L1	Programmed cell death protein ligand 1
PFM	Protein free media
PR	Progesterone receptor
qRTPCR	Quantitative reverse transcription polymerase chain reaction
RPMI	Roswell park memorial institute
RT	Room temperature
SDF-1	Stromal cell derived factor-1
SFM	Serum free media
TAM	Tissue associated macrophage
TFA	Trifluoroacetic acid
TGFβ	Transforming growth factor β
TIP	Tumour immune phenotype
TLR	Toll-like receptor
TMA	Tissue microarray



TME	Tumour microenvironment
TNBC	Triple negative breast cancer
TNF $\alpha$	Tumour necrosis factor- $\alpha$
T-reg	Regulatory T cells
VCAN	Versican
VEGF	Vascular endothelial growth factor
W	Wash

# Chapter 1: Introduction

## 1.1. The extracellular matrix

The extracellular matrix (ECM) is a collective term for the non-cellular component of tissues. It is a dynamic network of secreted molecules that forms the organisation of a tissue. The structural role of the ECM is to provide mechanical support to the cells and the tissue. The functional role of the ECM is the interactions it forms with cells to play a role in cell signalling, growth, differentiation, and migration <sup>1,2</sup>.

### 1.1.1 Main components of the ECM

The group of proteins that constitute the ECM have been defined under the term 'matrisome'. The matrisome consists of two main groups of proteins: ECM (core matrix) and ECM associated proteins. The core matrix is formed of 278 proteins. These proteins can be classed into three different groups – collagens, proteoglycans, and glycoproteins <sup>3,4</sup>.

Collagen is the most abundant protein within the ECM, making up 30% of the total protein. Within humans there are 28 different types of collagen fibres. The cross-linking of collagen fibrils by lysyl oxidases (LOX) leads to the assembly of the collagen backbone for the tissues. This backbone forms the main structural component of the tissues providing the mechanical strength and dictating the organisation of a tissue. The majority of the interstitial collagen is secreted by either resident or recruited fibroblasts <sup>1-3,5</sup>.

Proteoglycans can be found interspersed between the collagen fibres. They are formed by a core protein which has attachment sites for glycosaminoglycans (GAGs). GAGs are linear chains of repeating disaccharide units. There are four main types of GAGs which include heparin sulphate, chondroitin sulphate (CS), keratan sulphate and hyaluronan (HA) <sup>3</sup>. The structure and linkage of the disaccharides will be discussed later in the chapter. HA is the only GAG that cannot be sulphated, and it is also present without a protein backbone. Proteoglycans can be secreted like versican (VCAN) and aggrecan, be present on the basement membrane like perlecan or be present on the cell membrane such as syndecan. GAG chains are acidic and hydrophilic with a negative charge enabling the formation of a hydrated gel like structure <sup>3,5</sup>.

Glycoproteins are proteins which have short, branched oligosaccharides attached to the peptide compared to the long repeating sugars on proteoglycans. There have been around 200 complex glycoproteins identified within the ECM. The most

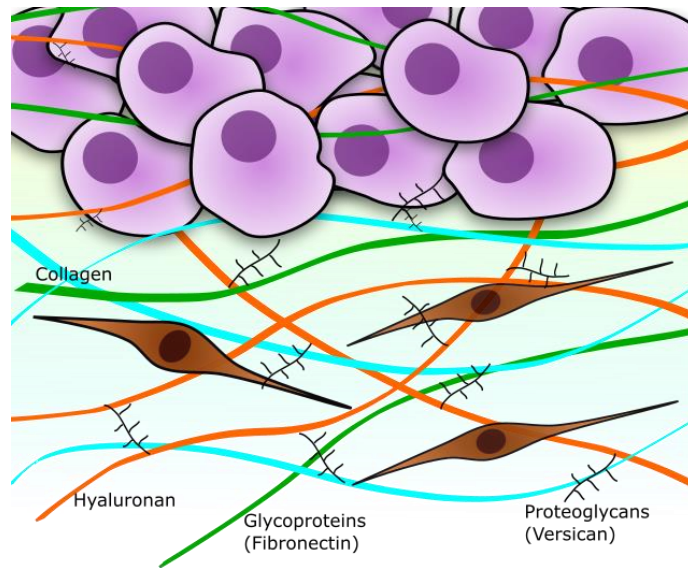
common glycoproteins are fibronectin (FN1) and laminin. Both these proteins have important roles in cell attachment and migration <sup>1,3</sup>.

Matrix associated proteins can be divided into three groups – secreted factors, ECM regulators and ECM-affiliated proteins. Secreted factors include cytokines, chemokines and growth factors which can bind to the core matrix proteins. FN1 has been shown to bind to a variety of growth factors such as vascular endothelial growth factor (VEGF) and platelet derived growth factor (PDGF). The binding of these secreted proteins to the ECM can lead to the formation of gradients that regulate cell binding, migration, and growth. ECM regulators have the ability to modify the ECM. This can include enzymes involved in cross-linking and protein degradation. The ECM-affiliated proteins are proteins with a close association to core matrix proteins. Examples include mucins, lectins and syndecans <sup>3,6</sup>.

### 1.1.2 Tumour ECM remodelling

During tumourigenesis there is a remodelling of the tumour ECM to assist in the growth and survival of the tumour. During wound healing, the remodelling of the ECM is a feature of homeostasis and is tightly regulated. However, during diseases such as cancer and fibrosis this process is dysregulated. There are three main mechanisms of ECM remodelling. This includes the deposition of ECM proteins, the modification of the ECM and the degradation of proteins <sup>7</sup>.

The deposition of proteins within the ECM is stimulated by inflammatory and growth factors such as transforming growth factor  $\beta$  (TGF $\beta$ ) and epidermal growth factor <sup>8</sup>. These growth factors can lead to the differentiation of fibroblasts to become cancer associated fibroblasts (CAFs) which in turn produce high levels of ECM proteins. Collagen I is highly present during tumour desmoplasia. Changes in the levels of specific types of collagen fibres such as collagen XV and collagen XIX have been observed during the progression of breast cancers from *in situ* to invasive <sup>1</sup>. Different collagen fibres have different roles in tumourigenesis with some being anti and some being pro-tumour, therefore the proportion of the different fibres is important within each tumour <sup>9</sup>. Proteoglycans and glycoproteins are also secreted by both tumour cells and CAFs. HA secretion is highly upregulated in tumours with the size of the chain affecting whether it has pro or anti-tumourigenic functions <sup>10</sup>. (Figure 1.1)



**Figure 1.1. Tumour ECM.** Proteins are secreted into the ECM by tumour cells and fibroblasts during tumourigenesis.

Modifications are introduced to the secreted proteins by an increased level of cross-linking and changes to the alignment of the fibres. These changes can lead to an increase in stiffness. Data have shown the stiffness of a tumour associates with disease progression and prognosis <sup>1,7</sup>.

The degradation of proteins is important for the invasion of tumour cells and tumour growth. Degradation enzymes such as matrix metalloproteases break down the basement membrane and can create migratory paths through the stroma. This process is also used by immune cells to invade into the tumour <sup>7</sup>.

Treatments targeting tumour remodelling have been trialled, however have not shown much success. These include targeting CAFs that produce high levels of ECM <sup>11</sup>, inhibiting enzymes like LOX to reduce cross-linking <sup>12</sup> and stiffness and also the introduction of ECM degrading enzymes <sup>7,13</sup>.

## 1.2. Matrix and tumour immunity

The ECM can have a direct or indirect effect on tumour immunity. Direct interactions via receptors can impact immune cell phenotypes and functions whilst indirect interactions occur through chemokines and cytokines.

### 1.2.1 ECM and immune phenotypes

The impact of the ECM on macrophage phenotype was first observed in 1983 where growing primary monocytes on collagen led to their inability to kill cancer cells<sup>14</sup>. ECM proteins can act as toll-like receptor (TLR) ligands and effect the differentiation of immune cells. The glycoprotein tenascin-C is a ligand of TLR4 and can initiate a pro-inflammatory phenotype on myeloid cells expressing TLR4<sup>15</sup>. This leads to tumour associated macrophages (TAMs) skewing to an M2 pro-tumour phenotype. VCAN also has binding sites for TLRs. TLR2 and TLR6 expressing macrophages can be activated by VCAN to produce tumour necrosis factor- $\alpha$  (TNF $\alpha$ )<sup>16</sup>. HA and collagen expression was found to upregulate CD44 expression in the macrophage cell line THP-1, where the upregulation of M2-like markers such as CD163 and CCL22 were also observed<sup>17</sup>. ECM dependent responses have been shown by Huleihel et al., where they show that bone marrow derived macrophages have an interleukin-4 (IL-4) activation like (M2-like) response towards the small intestinal submucosa matrix but an interferon- $\gamma$  and lipopolysaccharide (LPS) (M1-like) like response from the urinary bladder matrix<sup>18</sup>.

Similarly, dendritic cells can be activated by TLR ligands on the matrix. Immature dendritic cells cultured on FN1 or laminin for 48 hours led to a decrease in maturation markers like HLA-DR, CD83 and CD86. Stimulating the cells with LPS and CD40L did lead to the formation of mature phenotype showing that they can recover from that state<sup>19</sup>. Versikine, produced by the proteolysis of VCAN can activate IRF8 (interferon regulatory factor 8) in myeloid cells to select for Batf3 (Basic Leucine Zipper ATF-Like Transcription Factor 3) dendritic cells. The intact VCAN prevents this differentiation via the TLR2 pathway<sup>20</sup>.

The activation of T cells and differentiation into different subsets is also affected by matrix accumulation and alignment. Aligned collagen fibres and also high density collagen showed to reduce the level of IL-2 expressed by both primary T cells and Jurkat cells<sup>21</sup>. The encapsulation of resting CD4+ T cells in a three-dimensional (3D) hydrogel led to silencing of CD69 and low levels of CD25. The hydrogels were made to recapitulate both normal and tumour tissue and in both of these gels, programmed cell death protein 1 (PD-1) was not expressed by the cells, indicating that the stiffness does not affect the phenotype of resting T cells. A stiff matrix was shown to increase the level of CD25 expression in pre-activated CD4+ and CD8+ T cells<sup>22</sup>. A signature of the matrix associating 22 matrix molecules with disease

progression significantly correlated with the immune signatures for regulatory T cells (T-reg) and Th2 <sup>23</sup>.

### 1.2.2 ECM and immune cell migration

Matrix proteins can act as roads for the migration of immune cells. Collagen 1 fibrils have been found to assist in the migration of dendritic cells, in a process that is independent of integrins <sup>24</sup>. The migration of T cells and natural killer (NK) cells in this manner can also occur through an amoeboid movement, but this is affected by the porosity of the matrix and nuclear deformability. The migration into a collagen gel was associated with the expression of discoidin domain receptor 1 (DDR1). Either the blockage of DDR1 or the blockade of binding sites of collagen, led to a reduction in migration of T cells, neutrophils and monocytes <sup>25</sup>.

On the contrary, the ECM can be a barrier for immune cells migration. Fibre density also affects the migration with the loosely organised stroma having more immune cells present in comparison to the densely fibrous tumour. To counteract this stiffness, immune cells are dependent on the expression of proteases such as matrix metalloproteases (MMPs) to break down the fibres. This has been observed with T cells, where their migration into a collagen gel is dependent on MMP and integrin  $\beta 1$  expression <sup>26</sup>. The alignment of fibres also impacts migration, where a reduction in alignment in a 3D collagen gel led to a reduction in T cell migration <sup>25</sup>. The production of tenascin C by tumour cells can be used to paralyse T cells and prevent infiltration <sup>27</sup>.

As well as affecting immune cell migration directly, indirect effects via chemokine and cytokine binding have also been observed. Proteoglycans are the key matrix molecules associated with haptotactic chemokine and gradient formation <sup>28</sup>. This is due to the negative charge possessed by the GAGs attached to the proteins. The binding of chemokines to GAGs occurs in a specific manner with the ability of chemokines to bind to GAGs being dependent on the type of GAG and even the sulphation of the GAG. Mutations to substitute the positively charged amino acids in the binding sites of chemokines eliminates their ability to bind to GAGs.

Chemokines bind to GAGs in a covalent manner leading to a strong bond that can help maintain the gradient. The type of chemokine is specific for the recruitment of different immune cells. CCL3, CCL4, CCL5, CXCL9 and CXCL10 are associated with T cell recruitment with CXCL9 and CXCL10 being able to recruit CD8+ T cells to limit tumour growth. Contrastingly the accumulation of CXCL12 in tumours was

found to be correlated with T-reg cells aiding tumour growth <sup>29</sup>. Some chemokines such as CXCL4 bind to GAGs as opposed to chemokine receptors to function. This has been seen in the bloodstream and associated with extravasation <sup>30</sup>. Binding of proinflammatory cytokines to GAGs on the vascular has been shown to enhance immune cell infiltration and blocking these interactions has reduce leucocyte accumulation <sup>31</sup>.

### 1.3. Tumour immune phenotype

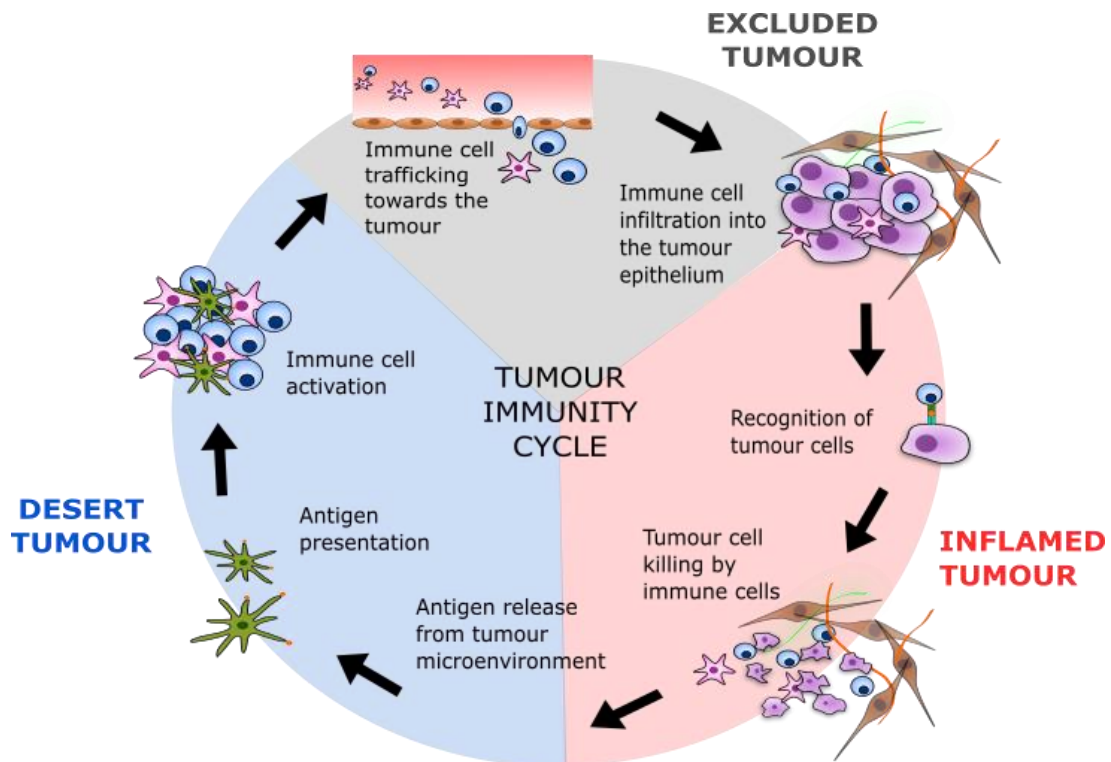
The tumour immune phenotype (TIP) classifies the abundance of immune cells in the tissue and the localisation of these immune cells with respect to the tumour epithelium. The phenotype can vary for different immune cells as they are impacted by various signals and interactions within the tumour microenvironment (TME). Understanding of the TIP for each tissue can provide a better understanding to how immune cells are interacting with the TME and provide indications for treatment response.

#### 1.3.1 Tumour immunity cycle

The tumour immunity cycle illustrated in Figure 1.2 was brought forward by Chen and Mellman and details how a tumour can trigger an immune reaction causing the trafficking of immune cells to the tumour eventually leading to tumour cell killing <sup>32</sup>. Interruptions at the different stages of this cycle lead to the different TIPs.

The tumour immunity cycle is initiated by the release of antigens by tumour cells through apoptosis or necrosis. These antigens are processed by antigen presenting cells such as dendritic cells which then travel to nearby lymph nodes for the priming and activation of immune cells <sup>33</sup>. For this stage to be successful proinflammatory cytokines are also required to be released by the tumour cells <sup>34</sup>. The immune cells then traffic towards the tumour site where they extravasate the blood vessel and are directed towards the tumour bed. Contact with the tumour cells leads to cell killing<sup>35</sup>, releasing more antigens, triggering the cycle again. Each stage of the cycle can be stimulated and inhibited with various ligands and cytokines. The presence of these inhibitors leads to the incompleteness of the cycle causing the further development of the cancer <sup>32</sup>.





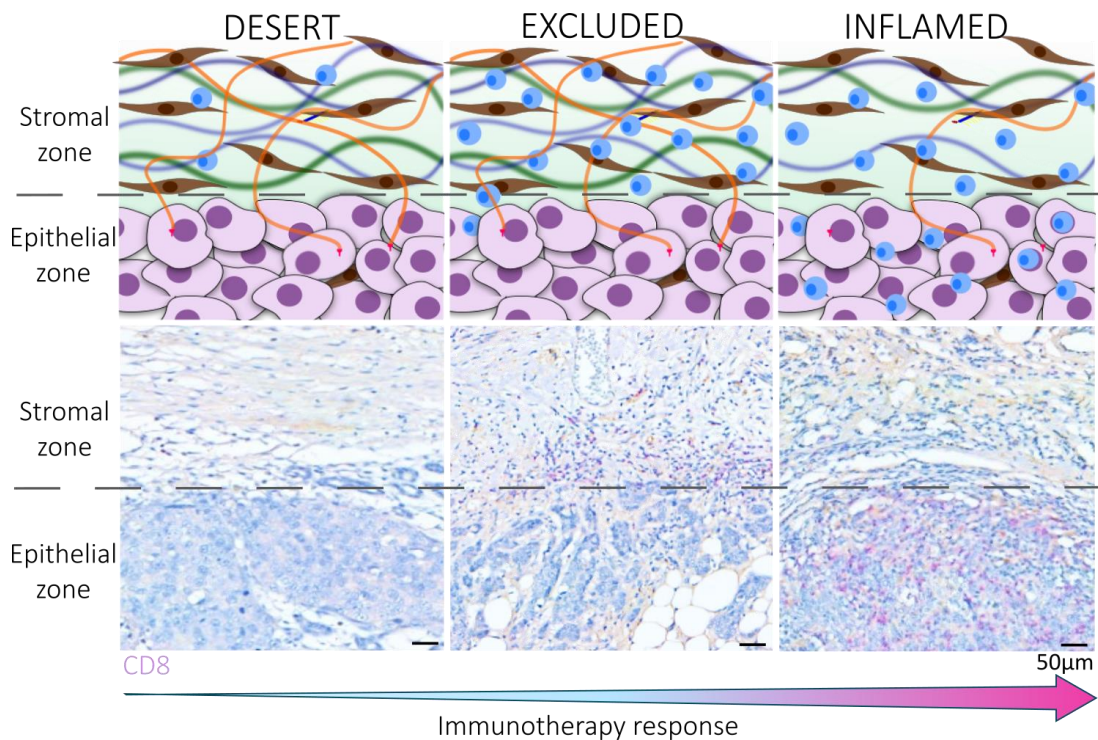
**Figure 1.2. The tumour immunity cycle.** Steps in the tumour immunity cycle outlined by Chen and Mellman. Steps highlighted in blue are where desert tumours stop. Steps highlighted in grey are where excluded tumours stop and stages in red represent inflamed tumours.

### 1.3.2 Tumour immune phenotype classifications

The initial characterisation of the TIP for tumours was that they are either hot or cold. Hot tumours referred to tumours where there is a high number of infiltrating immune cells. Cold tumours on the other hand have a very low immune cell count, leading to poor tumour immunity. Cold tumours could be a result of poor immune activation at the early stages of the cycle <sup>36</sup> or due to immune cells being unable to traffic inside the tumour <sup>37,38</sup>. This is visible in tumours such as pancreatic ductal adenocarcinoma where the density of the ECM makes it difficult for immune cells to migrate <sup>39</sup>. Poor vessel formation also limits the migration of immune cells into a tumour <sup>40</sup>.

Over the last 10 years the classifications of hot and cold tumours have become outdated with more importance being placed on where the immune cells are located. This has led to the formation of three classes: inflamed, excluded and desert. Inflamed tumours have high levels of immune infiltration within the tumour and within the epithelial beds. Desert tumours are similar to cold tumours where there is little to no immune presence within the tumour. An excluded tumour can be defined as a high immune cell presence with cells excluded to the outside of the

tumour or within the bordering stroma and having little contact with the tumour cells (Figure 1.3). These three phenotypes have been described by Chen and Mellman with the absence of an immune response and the prevention of migration within the stroma being the rate limiting steps for the desert and excluded tumours respectively (Figure 1.2) <sup>41</sup>.



**Figure 1.3. Tumour immune phenotypes.** Spread of T cells (blue) in the stroma and epithelium of the different TIPs.

### 1.3.3 Determination of tumour immune phenotypes

Multiple analysis techniques have been trialled for the determination of these phenotypes in tumours. Following immunohistochemistry of the tumours, the type of immune cell, density and localisation can be determined which will enable the identification of the phenotypes. Surgically resected tissues are used for the analysis in most cases compared to core needle biopsies, due to the size of the sample and the inability to identify tumour heterogeneity.

One of the first types of analysis of immune infiltrates and localisations was completed by Galon et al. They compared the immune cell densities in colorectal cancers. The centre of the tumour and invasive margin regions were analysed. A median was deduced for each area and was used as the threshold for determining

whether tissues had a high or low number of cells in these regions. High levels of CD3 positive cells in both regions (inflamed tissues) had the highest survival <sup>42</sup>.

The concept of an immunoscore was validated by Pages et al <sup>43</sup>. Multiple pathologists participated to determine the number of CD3+ and CD8+ T cells in the tumour and invasive margin. The number of cells were converted into percentiles with a 0-25% density scored as low, 25-70% scored as intermediate and over 70% scored as high. A significant correlation was observed between immune cell density and survival where a low immunoscore saw the lowest disease-free survival and a high immunoscore associated with a higher disease-free survival. <sup>43</sup>

A study by Kather et al. then looked at the immune topography of multiple tumour types. Within this study they increased the number of regions explored to look at both the inner (500µm within the tumour invasive front) and outer invasive margins (500µm outside the tumour invasive front) as well as the tumour core. Further into the analysis they observed a correlation between the inner invasive margin and tumour core so combined these regions. As well as multiple tumour types, the study also investigated different immune cell phenotype markers such as CD8, PD-1, FoxP3, CD68 and CD163. The cut-off values for the definition of the TIPs was determined as the median of the outer invasive and tumour core regions. Inflamed tissues had a density higher than this in the tumour core. Desert tissues were lower in both regions whilst excluded tissues were above in the outer invasive regions but lower in the tumour core. Likewise with the other studies the phenotype was associating with the survival. <sup>44</sup>

Grusso et al. looked at the spatial distribution of CD8+ T cells in triple negative breast cancer (TNBC). They studied the tumour margin and then the tumour which was subdivided into the tumour stroma and tumour epithelium. Their analysis led to the formation of four phenotypes. The fully inflamed tumours were high in tumour core and the epithelium. The stroma restricted tumours were high in the tumour core but low in the epithelium. The margin restricted tumours were low in the tumour core but high in tumour margin. The immune desert tumours were low in all regions. From this analysis they were able to assess the exclusion of CD8+ T cells at the tumour border and within the stroma. <sup>45</sup>

The analysis for immune phenotype has also been completed using only genomic and transcriptomic information. Xu et al. used data from the TCGA-BRCA project and analytical tools such as CIBERSORT and ESTIMATE to look at the level of immune and stromal cells. Single sample gene-set enrichment analysis was then

used to look at the expression level of 29 immunity associated gene signatures. From the analysis, three clusters were determined. The first cluster (ICI-A) had a high stromal score and high levels of resting immune cell populations which was considered as an immune excluded phenotype. The second cluster (ICI-B) had a lower immune score and considered desert. The final cluster (ICI-C) had an increase in anti-tumour lymphocytes identifying as an immune inflamed phenotype. The ICI-C cluster did not associate with survival. It was hypothesised that this may be due to immune evasion mechanisms.<sup>46</sup>

Overall, from the analysis techniques previously trialled it is clear that immunohistochemical analysis is important to understand where the immune cells are located. The techniques to classify the phenotypes following this analysis can vary based on the cut-off margins used. However, an association was seen with survival for the different methods. Identifying the correct method for analysis may depend on the type of tissue and regions to be analysed.

#### 1.3.4 Immune phenotype and immunotherapy response

Chen and Mellman have shown that with each immune phenotype there are mechanisms that can be associated with prevention of a natural immune response and how additional therapy is still necessary to aid a complete tumour immune response. Despite inflamed tumours having an abundant level of immune cells in the tumour, the activation states of these immune cells can be affected by cytokines such as IL-1<sup>47</sup> and co-stimulatory receptors like PD-1<sup>48</sup>.

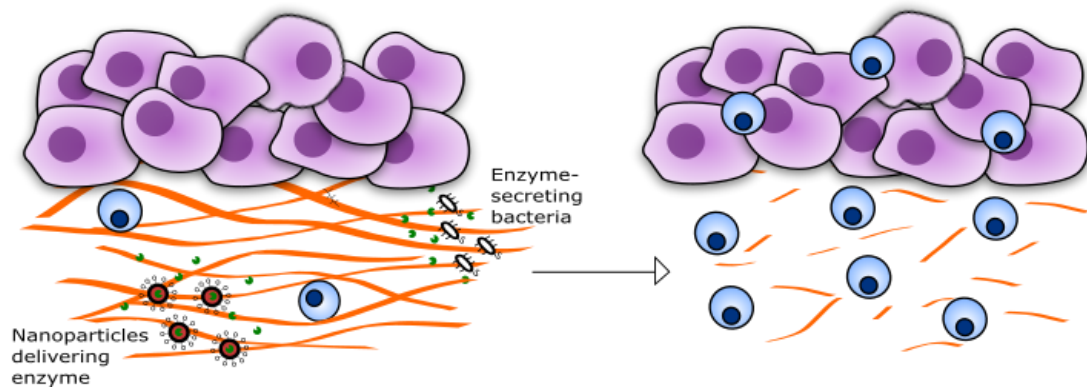
The importance of the TIP is that it has been shown to associate with immunotherapy responses. Checkpoint blockade therapy such as anti-PD-1/PD-L1 (Programmed cell death protein ligand 1) immunotherapy can prevent the dampening of T cell activation through this receptor. The inflamed phenotype tends to associate with a higher response rate to checkpoint blockade therapy<sup>41</sup>.

A study by Hammerl et al. looked at the association of immune topography to prognosis. Analysis of lymphocyte and myeloid cell localisation found that the location of lymphocytes associated with prognosis. From matching gene sets and immunohistochemistry analysis of phenotypes they were able to develop a gene-expression classifier for each phenotype. This included genes such as COL10A1, FAP, CXCL13 and CCL5. The test of the gene set on a database from an immunotherapy clinical trial showed that non-responders were enriched for the

excluded or ignored (desert) phenotype whilst responders were enriched for the inflamed phenotype. The classifier developed could be applied to other tumour types and was found to be associated with prognosis <sup>49</sup>.

#### 1.4. Targeting the ECM for immunotherapy

There have been multiple approaches to target the ECM to help improve response to immunotherapy. The targeting of the ECM involves the use of enzymes to degrade specific components of the ECM, helping to facilitate immune cell migration through the TME. The degradation of the ECM leads to gaps and pores which can be used by immune cells to travel through <sup>50,51</sup> (Figure 1.4).



**Figure 1.4. Breaking down the matrix to improve immune cell infiltration.** Enzymes can be delivered to the tumour to break down ECM fibres allowing more space for immune cells to invade the tumour,

Two of the main ECM components that have been targeted are collagen and hyaluronan. McKee et al. showed how the injection of collagenase led to an improved distribution of their oncolytic virus <sup>52</sup>. Hyaluronan has been targeted using a PEGylated hyaluronidase (PEGPH20) which has been used in trials alongside chemotherapy <sup>53,54</sup>. *In vivo* studies have also shown that the addition of hyaluronidase increased the sensitivity of tumours to anti-PD-L1 therapy <sup>55</sup>. The limitations of these treatments are the off-targets effects on other organs containing these matrix components.

To overcome the off-target effects, methods of delivering the enzymes directly to the tumour are being developed. Nanoparticles have been constructed where the enzyme is enclosed within a capsule that is degraded at the site of the tumour due to a mildly acidic environment <sup>56</sup>. Bacteria have also been used to deliver enzymes.

Attenuated *Salmonella typhimurium* bacteria have been developed to express hyaluronidase and express the enzyme specifically at the tumour site reducing the amount of HA detected in the tumour <sup>57</sup> (Figure 1.4).

Building on the new development of treatment strategies with more specific targets and enzymes could lead to more tumours changing from an excluded to inflamed phenotype.

## 1.5. VCAN

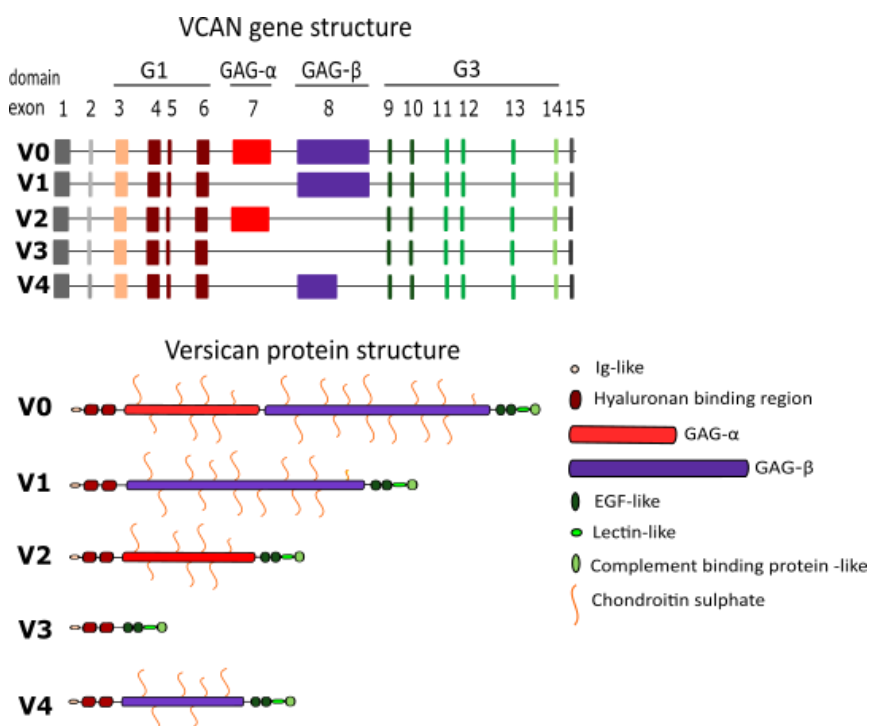
To identify potential ECM targets to improve immunotherapy, upregulated molecules from an ECM signature identified in our lab <sup>23</sup> were analysed in TNBC. Through this analysis I identified VCAN to be a potential target. Details of this analysis are outlined in Chapter 3. VCAN is a proteoglycan that is present within the ECM, with upregulation of its levels observed in cancer. The multidomain structure and different isoforms enable the protein to have a versatile role within the TME <sup>58</sup>.

### 1.5.1 VCAN structure and isoforms

VCAN is formed from 15 exons. The splicing of these exons leads to the generation of five different isoforms. These exons translate into three domains. The first and last domain are conserved between all CS proteoglycans (CSPG). The first domain is the G1 domain. This is expressed by the exons 3-6 <sup>59</sup>. This domain contains an immunoglobulin (ig)-like module and a hyaluronan binding region. The Ig-like module enhances the binding of the hyaluronan binding domain to hyaluronan. The minimum length of HA which can interact with the G1 domain is HA10 (10 disaccharides) <sup>60</sup>. The last domain is G3 formed by the exons 9-15. This domain contains the EGF-like (Epidermal growth factor-like), lectin-like and complement binding protein like modules. Both these domains are present within all isoforms. Splicing for different isoforms occurs at the 7<sup>th</sup> and 8<sup>th</sup> exon. This translates to the CS binding GAG domain which is split into the GAG- $\alpha$  and GAG- $\beta$  domain. The GAG- $\alpha$  domain is 989 amino acids long and contains 5-7 binding sites for CS binding. The GAG- $\beta$  domain is 1751 amino acids long and contains 12-15 CS binding sites <sup>59,61,62</sup> (Figure 1.5).

The V0 isoform contains all domains and therefore also known as full-length VCAN. The V1 isoform does not contain exon 7 for the GAG- $\alpha$  domain. The V0 and V1 isoforms are the largest with molecular weight of approximately 370 and 263kDa

respectively. The V2 isoform does not contain exon 8 (GAG- $\beta$  domain) and is 180kDa. V3 is the smallest isoform at 74kDa as it has neither exon 7 or 8 for GAG domains and as a result has no CS binding sites, therefore not a proteoglycan by definition<sup>59,61,63,64</sup>. The V4 isoform was recently discovered by Kischel et al. and like V1 it does not contain the GAG- $\alpha$  domain, but it also has a truncated GAG- $\beta$  domain<sup>65</sup> (Figure 1.5).



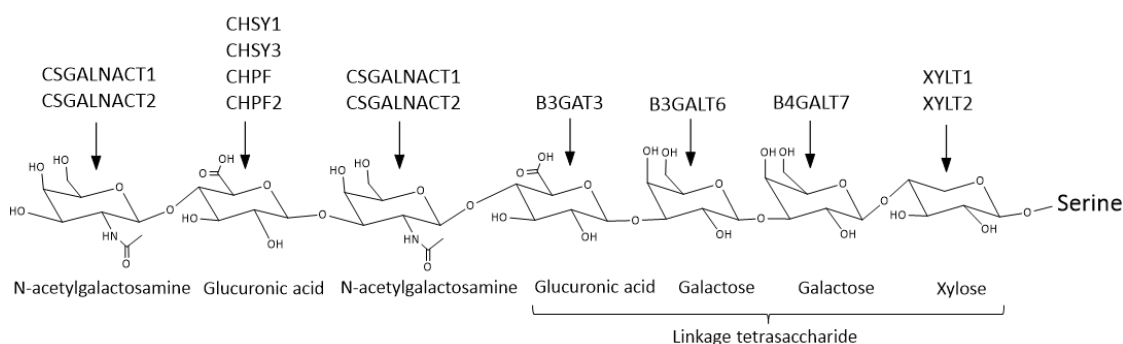
**Figure 1.5. Structure of VCAN.** The gene and protein structure of each isoform of VCAN. The colours of the exons are of the respective protein structure. Adapted from Hirani et al.<sup>58</sup>

### 1.5.2 Chondroitin sulphate

CS is attached to VCAN at serine residues on the GAG domains. Through the consensus sequence of GAG attachment sites, it is thought that the GAG- $\alpha$  domain contains 5-8 sites of CS glycosylation and the GAG- $\beta$  domain binds 12-15 CS sites<sup>61,66</sup>. These serine residues are flanked by glycine residues on the carboxyl side. Acidic residues are also present around the serine residue. The CS chains are linked to the serine via a linkage tetrasaccharide (Figure 1.6). The tetrasaccharide is formed by a xylosyltransferase enzyme attaching a xylose to the serine. The activity of xylosyltransferase (XYLT1, XYLT2) in the cell can affect the number of chains attached to the proteoglycan. The xylose becomes phosphorylated to enable the addition of the other sugars. Without the phosphorylation the chain is capped with a

sialic acid. The first galactose (Gal) is transferred by  $\beta$ -1,4-galactosyltransferase-7 (B4GALT7). The second Gal is added by  $\beta$ -1,3-galactosyltransferase-6 (B3GALT6). Glucuronic acid (GlcA) completes the tetrasaccharide and its insertion is facilitated by  $\beta$ -1,3-glucuronyltransferase-3 (B3GAT3)<sup>67</sup>. The addition of these sugars has been proposed to take place in the pre-golgi compartment. The remainder of the GAG polymerisation takes place in the golgi apparatus lumen<sup>68-70</sup>.

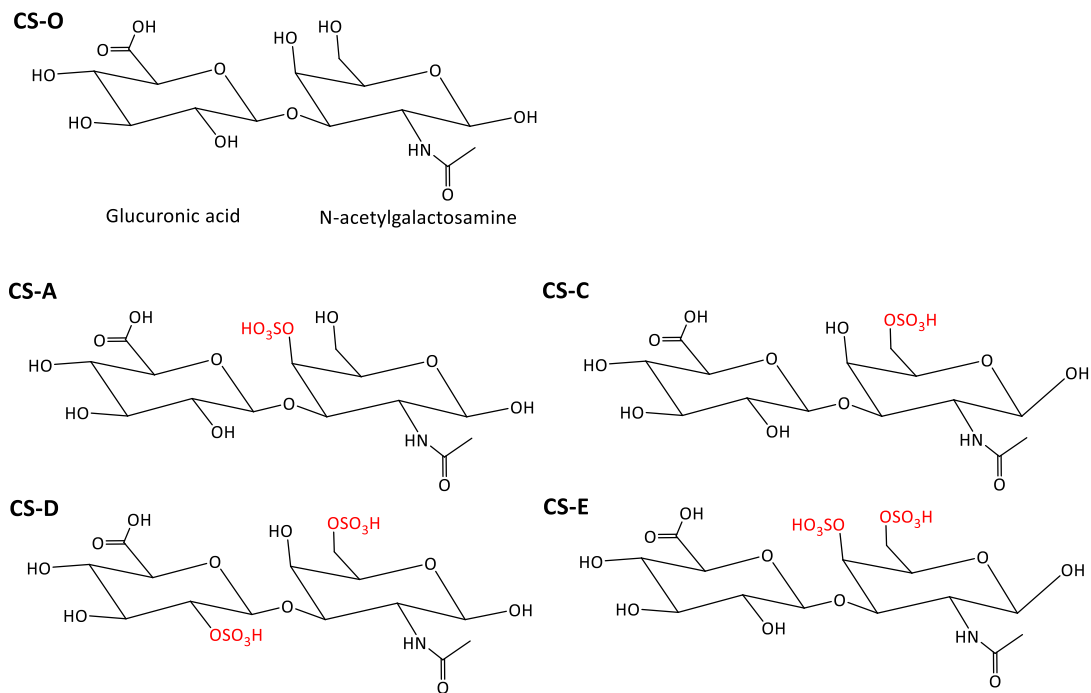
The structure of CS was first characterised by Levene and Forge in 1915<sup>71,72</sup>. The CS disaccharide consists of N-acetylgalactosamine (GalNAc) and GlcA. The initiation of the CS chain is via  $\beta$ 4-N-acetylgalactosaminyltransferase (CSGALNACT1,2) adding the N-acetylgalactosamine to the chain<sup>73</sup>. This is then followed by the elongation of the chain with the CS- $\beta$ 3-glucuronyltransferase-II and  $\beta$ 4-N-acetylgalactosaminyltransferase enzymes responsible for the transferring of GlcA and GalNAc respectively<sup>70,74</sup>. The average length of CS chains are 40 disaccharides corresponding to around 20kDa<sup>74</sup>.



**Figure 1.6. Formation of CS linkage to serine.** Enzymes shown which attach the respective sugar to the CS chain to form the linkage tetrasaccharide and CS chain.

As well as variations in the chain length of CS, there are also variations in the sulphation patterns of the disaccharides within a chain. These isomers are brought about by the sulphation of the sugars. There are five different isomers of CS as shown in Figure 1.7. CS-O has no sulphations. CS-A (C-4-S) is sulphated at the 4<sup>th</sup> carbon of the GalNAc. CS-C (C-6-S) is sulphated at the 6<sup>th</sup> carbon of the GalNAc. CS-E (C-4,6-S) is sulphated at both the 4<sup>th</sup> and 6<sup>th</sup> carbon of GalNAc. CS-D (C-2,6-S) is sulphated at the 2<sup>nd</sup> carbon of GlcA and the 6<sup>th</sup> carbon of GalNAc<sup>74</sup>. These sulphation patterns are driven by different sulfotransferases. C-4-S is formed by chondroitin 4-O-sulfotransferase, C-6-S by chondroitin 6-O sulfotransferase, C-4,6-S by GalNAc 4-O-sulfate 6-O sulfotransferase and C-2-S by uronyl 2-O sulfotransferase<sup>70,75-77</sup>.





**Figure 1.7. Structure of CS isomers.** CS is made of GalNAc linked to GlcA. Sulphation of the disaccharide give rise to the different isomers. Sulphation of the different isoforms shown in red.

The difference in the sulphation pattern has been shown to be associated with different functions. Proteins show varied levels of specificity and affinity to the different CS isomers<sup>78,79</sup>. This is independent to the number of sulphates present on the CS isomers<sup>80,81</sup>. In a study by Sato et al. they showed that nephronectin, a basement membrane protein, was able to bind to CS-E but not to any other isomer<sup>82</sup>. N-cadherin binds to CS-E but not CS-A<sup>83</sup> whilst E-cadherin binds to both CS-E and CS-A but cannot bind to CS-C<sup>84,85</sup>. This variation in binding abilities leads to tissues showing contrasting proportions of each isomer in order to facilitate function.

### 1.5.3 VCAN interactions with the ECM

The versatility in the functions of VCAN are due to the multiple binding opportunities from the different domains. The G1 domain interacts with hyaluronan (Figure 1.8). This interaction occurs via the two hyaluronan binding regions illustrated in Figure 1.5. The binding of HA to VCAN is strengthened by the presence of a link protein. The binding of VCAN to both HA and the link protein occurs at these hyaluronan binding regions<sup>60,86</sup>. The gene for link protein *HAPLN1* was found to be co-localised with the *VCAN* gene<sup>87</sup>. The binding of HA to VCAN enables the formation of

aggregates consisting of other HA ligands within both the extracellular and pericellular matrix. The binding of both these molecules has been shown to be important in development as the knockout of either of the genes led to mice dying after 10 days of gestation due to heart defects <sup>88-90</sup>.

One of the molecules that forms an aggregate with VCAN and HA is CD44. CD44 is a transmembrane glycoprotein. It is expressed on the cell surface of multiple cell types including cancer cells and immune cells. CD44 is a receptor for HA and binds to it with high affinity. Via CD44, HA and VCAN can form a pericellular matrix around the tumour cells. This was seen to increase invasiveness of ovarian cancer cells <sup>91</sup>. As well as VCAN forming a macromolecular structure with CD44 through HA, it can also bind to CD44 via the GAG domains, with hyaluronidase treatment not disturbing the VCAN-CD44 complex formation. The affinity for CS to CD44 is lower than HA to CD44, with no specificity identified for the different CS isomers <sup>92</sup>.

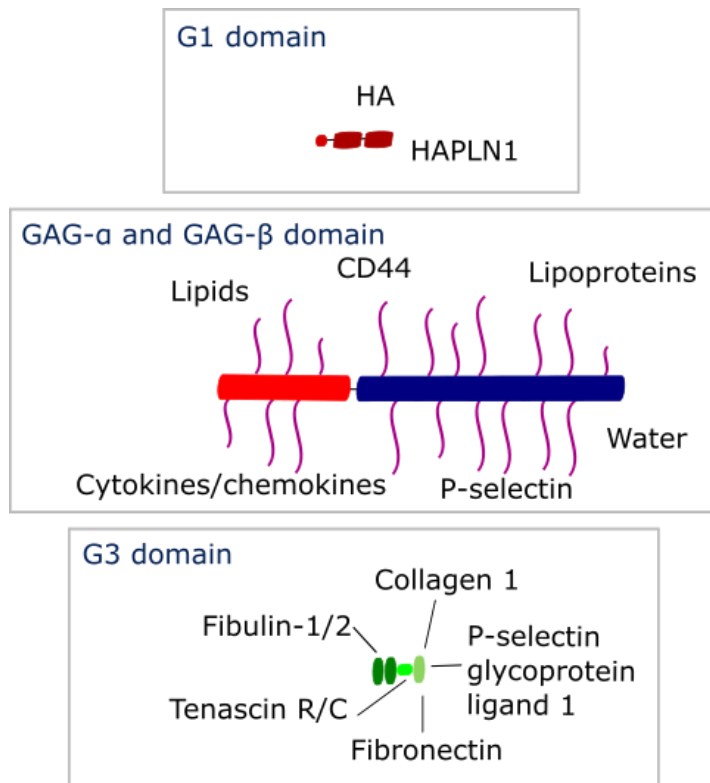
The CS chains in the GAG domain are negatively charged so can bind to positively charged molecules such as chemokines and lipoproteins (**Figure 1.8**). The binding to chemokines is later discussed in the roles VCAN has in immunity. The binding of VCAN to lipids and lipoproteins has been visible on the vascular walls and can be linked to promoting atherosclerosis. The CS-C isomer has a stronger binding capability to low density lipoproteins compared to CS-A <sup>92-94</sup>.

The G3 domain has three different modules which have varying binding properties (Figure 1.8). The first module is the EGF-like motifs. These have been shown to bind to fibulin-1 and 2. The function of VCAN and fibulin binding has been explored in heart development. The binding of VCAN and fibulin was seen to be present in a macromolecular structure with HA where the degradation of HA removed both these proteins within tissue. <sup>92,95</sup>

The second module of the G3 domain is the lectin-like module which has also been referred to as the carbohydrate recognition domain. Tenascin R and C both bind to this module in a calcium dependent manner <sup>96,97</sup>. The knockdown of tenascin R in mice led to disorganised formation of perineural nets indicating its involvement with structure along with VCAN and other lecticans <sup>92,98</sup>. P-selectin glycoprotein ligand-1 also binds at this module which is detailed later in this chapter.

The final module is the complement binding like motif. The predominant binding partners for VCAN in this region are FN1 and collagen type 1. Pull-down experiments of VCAN from conditioned media found both these proteins bound to the VCAN extracted <sup>99</sup>. VCAN has been found to effect collagen organisation <sup>100</sup>,

whilst complexes of FN1 with VEGF and VCAN were found to impact endothelial cell function <sup>92</sup>.



**Figure 1.8. Binding partners of VCAN at the different domains.**

#### 1.5.4 VCAN expression

The expression of VCAN has been shown to change from embryogenesis and development to healthy adults. The different isoforms are also observed in diseases where it has a role in multiple inflammatory diseases.

##### 1.5.4.1 Activation of VCAN transcription

The stimulation of VCAN transcription has not yet been fully understood. Associations have been made with transcription factor pathways to an increase in VCAN expression. VCAN has been identified to be stimulated by multiple growth factors such as TGFβ <sup>101</sup>, fibroblast growth factor (FGF) <sup>102</sup> and PDGF <sup>103</sup>. These growth factors can lead to the stabilisation of β-catenin through the Wnt signalling pathway or the PI3K signalling pathway <sup>104</sup>. The stabilisation leads to the downstream signalling of the transcription factor TCF4 to bind to the TCF/LEF site

on the transcription factor binding site of the VCAN promoter, leading to VCAN transcription<sup>105</sup>.

#### 1.5.4.2 In development

VCAN is first present during embryogenesis. The control of the expression and degradation of VCAN has been shown to be important to organ formation during embryonic development. During cardiac development VCAN is present within the endocardial cushion. The *heart defect* mouse has a transgene insertion within the *CSPG2* sequence. These mice show synthetic lethality due to the abnormal development of the right ventricle and no endocardial cushion<sup>89</sup>. The cleavage of VCAN by ADAMTS1 has been shown to be important in facilitating the remodelling into the atrial and ventricular septa<sup>106</sup>. Research into development of the mouse lung has shown that VCAN staining is initially present within the lung mesenchyme and basement membrane of the epithelium of the bronchiole. A reduction in VCAN later on in gestation associates with a reduction in lung tissue volume leading to structural development of the lung<sup>107</sup>.

#### 1.5.4.3 In adults

To study where VCAN is expressed in healthy adult tissues, immunohistochemical staining of V0 and V1 isoforms was completed on different tissues of multiple organs. VCAN was found to be present around all blood vessels with only the capillaries of the kidney glomeruli and liver sinus found to be negative<sup>108</sup>. Smooth muscle cells were found to be the source of VCAN in the vessels<sup>109</sup>. The V2 isoform has been found in the central nervous system. It is found in the ECM that surrounds myelinated fibres. Accumulation of VCAN has been shown at the nodes of Ranvier<sup>110</sup>.

#### 1.5.4.4 In diseases

Within cardiovascular disease the modulation of VCAN expression can be observed. During an abdominal aortic aneurysm there is a reduction in the amount of VCAN, specifically the V0 isoform. This reduction has been associated with a reduction in smooth muscle cells and an increase in VCAN degradation<sup>111</sup>. In tissues with high VCAN the levels of elastic fibres are depleted. During vascular

disease there is an overexpression of the V3 isoform as well as the inhibition of other isoform expression inducing elastic fibre formation <sup>112</sup>. An upregulation of VCAN is also observed in inflammatory diseases. This will be further discussed in the section outlining the role VCAN has with immunity.

#### 1.5.4.5 Degradation of VCAN

VCAN can be degraded by multiple proteases including MMPs <sup>113</sup>, serine protease plasmin <sup>114</sup> and the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motif) enzymes. Degradation by the ADAMTS 1, 4, 5, 9 and 15 enzymes leads to the formation of the versikine fragment. This fragment is at the amino terminal of the protein and is formed at the cleavage site of Glu<sup>441</sup> and Ala<sup>442</sup> <sup>115</sup>. This versikine fragment can be detected by the epitope DPEAAE <sup>116</sup>. The importance of the breakdown of VCAN has been shown in development and diseases.

Another fragment of VCAN degradation that has been detected is glial hyaluronan-binding protein (GHAP). GHAP is formed through the cleavage of the V0 and V2 isoform at the site Glu<sup>405</sup>-Gln<sup>406</sup>. ADAMTS 1, 4, 5 and 9 have been associated with this degradation in the adult brain. <sup>117</sup>

Analysis into the cleavage sites for the ADAMTS1, 4 and 5 enzymes shows that the Glu<sup>441</sup> and Ala<sup>442</sup> site is conserved between the enzymes <sup>118</sup>. The roles of the peptides formed from the other cleavage sites have not yet been defined.

### 1.6. VCAN and cancer

VCAN has been found to be upregulated in tumours. It has been associated with tumour cell proliferation and migration. The expression of VCAN and its roles can vary between tumours.

#### 1.6.1 Cells expressing VCAN

The cellular source of VCAN varies between tumour type and the role of VCAN also differs from cells in different tumours.

A study looking at the cellular source of proteins in pancreatic cancer showed that VCAN in the ECM was derived from cancer cells <sup>119</sup>. Tumour cell expression was

also observed in ovarian cancer <sup>120</sup>, hepatocellular carcinoma <sup>121</sup>, colon carcinoma <sup>122</sup> and glioma <sup>123</sup>.

Stromal expression of VCAN can be seen in multiple adenocarcinomas such as breast cancer <sup>124</sup>, cervical cancer <sup>125</sup>, ovarian cancer <sup>126</sup>, oral squamous cell carcinoma <sup>127</sup> and prostate cancer <sup>128</sup>. The expression of VCAN in CAFs is associated with TGF $\beta$  stimulation <sup>126</sup>. CAFs expressing VCAN in non-small cell lung cancer are both fibroblast activation protein (FAP) and  $\alpha$ -smooth muscle actin ( $\alpha$ SMA) positive <sup>129</sup>.

Myeloid VCAN expression has been identified in metastatic tumours. Bone marrow derived myeloid progenitor cells were found to secrete VCAN in the premetastatic lung, promoting tumour cell metastasis from the breast <sup>130</sup>. VCAN expression in TAMs has been identified as part of the signature identifying pro-angiogenic TAMs <sup>131</sup>. Low levels of VCAN expression has also been detected in activated B cells <sup>132</sup>.

### 1.6.2 Impact on tumour prognosis

VCAN has been linked with prognosis in multiple tumours. It was identified in a signature of 22 molecules to be associated with disease progression. This signature was found to be relevant for 13 different tumour types, including breast, ovarian, lung, colon and pancreatic cancer <sup>23</sup>. The upregulation of VCAN has been studied in multiple tumour types with association found with poor prognosis. This has been shown in bladder cancer <sup>133</sup>, endometrial cancer <sup>134</sup>, renal carcinoma <sup>135</sup> and gastric cancer <sup>136</sup>. Immunohistochemistry analysis in bladder cancer showed a significant correlation between VCAN expression and the number of tumours, level of invasion, metastasis and grade. The depletion of VCAN expression in clear cell renal cell carcinoma cell lines led to a reduction in cell migration and invasion. This was also correlated to MMP7 and CXCR4 depletion <sup>136</sup>. This indicates that the role of VCAN in tumour cell migration is important in metastasis leading to its association with poor prognosis.

On the other hand there have been cases of better prognosis and survival with greater VCAN expression. One study found that the source of VCAN expression is important to its role in prognosis. A high expression of VCAN by tumour epithelial cells in gastric cancer was seen to correlate with higher 5-year disease free survival <sup>137</sup>. This was also observed in stage 2 and 3 colon cancer where the expression of VCAN by epithelial cells at the invasive front correlated to a higher disease-free

survival <sup>138</sup>. Studies on pancreatic neuroendocrine tumours also indicated VCAN expression is linked to better disease-free survival <sup>139</sup>.

The proteolysis of VCAN is also an important indicator of prognosis. The breakdown of VCAN to versikine has been identified as a prognostic immune biomarker and could be important for the prediction of immune response <sup>140,141</sup>.

### 1.6.3 VCAN isoforms in cancer

The size of the different VCAN isoforms affect the potential binding partners of the protein as well as the effect on the organisation of the matrix. Due to this the roles of the different isoforms vary in tumours.

The V0 and V1 are the highest expressed isoforms in tumours <sup>142,143</sup>. The roles of these isoforms are similar due to their large size. The overexpression of V1 in cancer cells was shown to increase proliferation and inhibit apoptosis <sup>144</sup>. The transfection of chondrosarcoma cells with V1 led to an enhancement in cell motility and migration <sup>145</sup>. The mutation of V1 led to the inability of epidermal growth factor receptor (EGFR) signalling and therefore limited cell growth <sup>146</sup>. This was not the case for V2 mutations, indicating that the GAG- $\beta$  domain has a role in this pathway.

The overall level of V2 in tumours is much lower than other isoforms. The roles of this isoform in tumours has mostly been explored through its overexpression or in neural tissues where expression levels are generally higher <sup>147</sup>. Studies have shown V2 transfection to reduce the proliferation of tumour cells but enhance the viability <sup>148</sup>. The stimulation of V2 expression was observed when breast cancer cell lines were treated with the potassium channel blocker Amiodarone. Following treatment a reduction in cell proliferation and metastasis was also observed. This effect was found to be associated with the EGFR pathway <sup>149</sup>. The expression of the V2 isoform in glioblastoma cell lines was also found to promote angiogenesis through a process mediated by an increase in FN1 expression <sup>148</sup>.

The V3 isoform possesses no GAG domains. Studying the function in cancer provides an insight into the role of the G1 and G3 domains. Overexpression of the V3 isoform was found to interfere with the binding of CD44 with EGFR-ErbB2 via its EGF-like repeats in the G3 domain. This led to a decrease in cell proliferation <sup>150</sup>. However a study showed that the V3 isoform has dual roles in melanoma, where it reduces the proliferation of cells as observed in previous studies and also favours the tumour metastasis <sup>151</sup>.

The presence of V4 has only been published in breast cancer so little is known about the roles it plays towards tumour progression <sup>65</sup>.

## 1.7. VCAN and immunity

The interactions of VCAN with chemokines and immune cells directly enable it to play a role in inflammation. Due to these interactions VCAN has been seen as a key matrix molecule in the study of inflammatory diseases.

### 1.7.1 VCAN interactions with chemokines

The negative charge of the CS on VCAN leads to the attraction of positively charged molecules such as growth factors, chemokines, and cytokines. The binding of chemokines to GAGs within the ECM is important for the formation of a chemokine gradient to induce leukocyte migration <sup>152</sup>. For leukocyte extravasation to occur chemokines are bound to GAGs on the endothelial lumen leading to rolling of immune cells which then form an integrin ligand bond with endothelial cells to facilitate migration through the vessel wall <sup>153</sup>. (Figure 1.9)

Hirose et al. tested which chemokines VCAN had the ability to bind to using a dot blot with different chemokines spotted onto a membrane and detecting binding to the chemokines with biotinylated VCAN protein. From the membrane, VCAN was seen to bind to the CXC chemokines: CXCL10, CXCL4 and SDF-1 $\beta$  (Stromal cell derived factor 1, CXCL12), and the CC chemokines: CCL2, CCL8, CCL5, CCL20 and CCL21. The binding of VCAN to the chemokines was found to inhibit their downstream effects on T cells such as the increase in intracellular Ca<sup>2+</sup> <sup>154</sup>. The binding of CCL2 to CS was found to protect CCL2 from degradation. CCL2 leads to the recruitment of monocytes. Monocytes also produce VCAN and therefore the formation of a positive feedback loop may lead to a high influx of monocytes <sup>155</sup>.

The sulphation patterns of the CS isomers has been shown to be important to the binding of the chemokines. The testing of chemokine binding to oversulphated CS using surface plasmon resonance showed that these CS isomers could strongly bind to CXCL10, CXCL11 and CXCL13. Moderate and low binding was seen with CCL28, CXCL12, CXCL9, CCL19, CCL18, CCL16, CCL1, CCL22 and CCL21 <sup>156</sup>. The single sulphated CS-A did not bind to these chemokines. This was also shown with the exogenous addition of CS isomer CS-E, which inhibited VCAN binding to



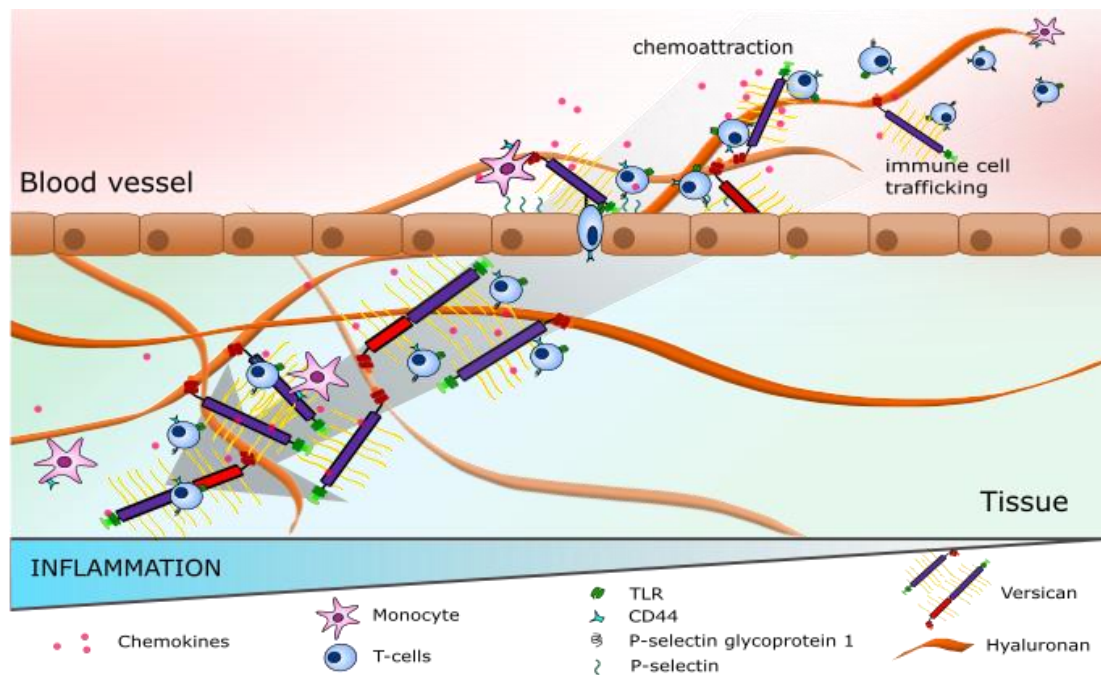
chemokines<sup>154</sup>. Oversulphation of CS has been shown to inhibit T cell chemotaxis via SDF-1<sup>156</sup>.

### 1.7.2 VCAN interactions with immune cells

As well as the indirect interactions with immune cells via chemokines, VCAN can also interact directly. One form of interaction is via the TLR pathway. Macrophages have been shown to be activated by VCAN through the TLR2-TLR6-CD14 complex, this leads to the secretion of TNF- $\alpha$  and IL-6<sup>157</sup>. VCAN can also stimulate the production of IL-6 and IL-10 by dendritic cells through TLR2 binding<sup>158</sup>. This leads to an immunosuppressive TME<sup>159</sup>.

The breakdown of VCAN to versikine has been positively associated with CD8+ T cells levels in colorectal cancer. Versikine interacts with FLT3L treated dendritic cells and increases the expression of Batf3 and IRF8, increasing the activation of T cells<sup>20,160</sup>.

VCAN can interact with receptors on lymphocytes such as CD44. This binding is through the CS chains<sup>161</sup>. This binding is associated with the extravasation of leukocytes. The binding of VCAN to HA can lead to the interference of lymphocyte binding to HA causing immunosuppression<sup>162</sup>. P-selectin glycoprotein 1 is expressed on the surface of leukocytes and can bind to the G3 domain of VCAN (Figure 1.9). This binding led to cell aggregation both *in vitro* and *in vivo*. Blood plasma has been shown to contain VCAN fragments containing G3<sup>163</sup>. This could lead to prevention of excess leukocyte extravasation<sup>164</sup>.



**Figure 1.9 Chemokine binding to VCAN facilitating extravasation.** Adapted from Hirani et al. <sup>58</sup>. VCAN binding to hyaluronan leads to the accumulation of chemokines around the proteoglycan, guiding immune cells such as monocytes and T-cells for extravasation across the endothelium into the ECM.

## 1.8. Breast cancer

Breast cancer is currently the most diagnosed cancer in women in the world, accounting for nearly 2.3 million new cancer cases in 2020. The increase in the number of cases has been attributed to the increase in levels of mammography screening and the aging population <sup>165</sup>.

### 1.8.1 Risk factors

Around 10% of all breast cancers are due to hereditary mutations. Mutations to the *BRCA1/2* genes account for approximately half of these cases. Around 70-85% of patients with a *BRCA1* associated breast cancer have TNBC <sup>166</sup>.

The *BRCA1/2* genes have major functions in DNA repair and cell cycle control <sup>167</sup>. Advantages have been taken on these mutations by the introduction of poly ADP-ribose polymerase (PARP) inhibitors, which have led to an increase in progression free survival <sup>168</sup>. A first-degree relative with breast cancer leads to the relative risk increasing by 3 for early onset breast cancer. Other breast cancer associated genes with high penetrance are *TP53*, *CDH1*, *PTEN* and *STK11* <sup>169,170</sup>. Like the *BRCA* genes, these four genes are also tumour suppressors. *TP53* also has a role in DNA

repair and cell cycle control, but can also induce apoptosis and senescence. *CDH1* regulates cellular adhesions and controls the proliferation and motility of epithelial cells. *PTEN* is involved in cell cycle control. *STK11* regulates energy metabolism and cell cycle.

Age has been found to be a risk factor with the occurrence risk increasing from 1.5 to 3-4% as age increases from 40-50 to 70. The more aggressive TNBC subtypes tend to be found in younger patients who are under the age of 40 and in patients over 70 the luminal A subtype is much higher <sup>170</sup>.

Race and ethnicity have also been found to affect the risk of breast cancer and the subtype. Within Asian developing countries the proportion of cases of patients under the age of 35 increases to 25% in comparison to 10% in developed countries <sup>166</sup>. African and African-American women were found to have the highest rates of TNBC cases and higher rates of metastatic diseases leading to lower survival rates <sup>171</sup>.

Lifestyle and environmental factors also play a role as they do with other forms of cancer. Factors such as obesity, low physical activity and alcohol have been linked to around 20% of all breast cancers worldwide. A high body mass index was found to associate with an increase in risk of relapse and a reduction in disease free survival <sup>172</sup>. The role of physical activity towards risk has not been clearly determined but hypotheses have been made that it may be linked towards pathways involving sex hormones, adiposity and chronic inflammation <sup>173</sup>. Alcohol intake has been shown to increase the levels of oestrogen leading to a hormonal imbalance, as a result alcohol intake has been associated with an increase in oestrogen receptor (ER)+ breast cancer <sup>174</sup>.

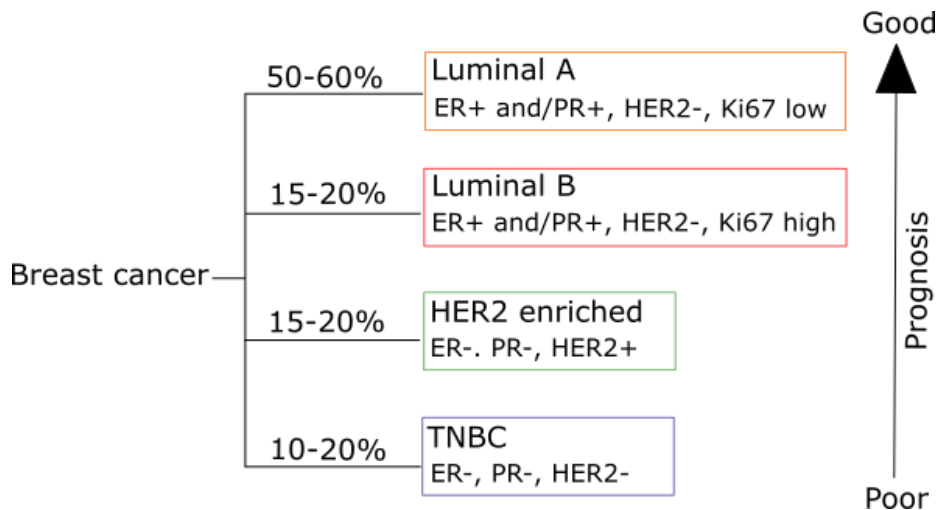
The exposure to hormones oestrogen and progesterone have been associated with the risk of breast cancer. Early pregnancy in the early 30s was found to reduce risk due to the changes in the sensitivity of glands to exposure to hormones at a later stage <sup>175</sup>. The use of hormone replacement therapy has been found to increase the risk of breast cancer, this has been associated with the delay in onset of the menopause <sup>176</sup>.

Under 1% of breast cancer cases are in men. The diagnosis in men tends to be around the age of 67. As with women the mutations in the *BRCA1/2* genes, increased oestrogen levels and family history are associated with increased risk <sup>177</sup>.

### 1.8.2 Subtypes of breast cancer

Breast tumours are highly heterogenous, therefore to aid in the prognosis and therapeutics, tumour subtypes are applied to the tumours. There are four main subtypes: Luminal A, Luminal B, human epidermal growth factor receptor 2 (HER2) positive and TNBC (Basal-like) (Figure 1.10). These subtypes were identified by Perou et al. where they identified these subgroups based on the gene expression patterns from 42 patients <sup>178</sup>. The diagnosis of subtypes is completed through the staining of a biopsy and the immunohistochemical expression of receptors and proliferation marker ki-67. These receptors are the ER, progesterone receptor (PR) and HER2.

The luminal A subtype is positive for ER+, PR+/- and HER2- . The expression of ki-67 is low, identifying these tumours as slow growing. These patients show the best prognosis with fewer instances of relapse. The luminal B subtype is also ER+, PR+/- and HER2-. However, the level of Ki-67 is high. This leads to tumours being of a higher grade with faster growth. The prognosis of these patients is poorer compared to luminal A. The HER2 positive subtype contains the expression of only the HER2 receptor. This tumour tends to be fast growing and aggressive <sup>179</sup>. For all three of these subtypes, patients can benefit from hormonal therapy and chemotherapy. The final subtype is TNBC. These tumours express no hormone receptors and therefore hormonal therapy cannot be used. This subtype accounts for 20% of all tumours and is the most aggressive form of breast cancer. Compared to other forms of breast cancer TNBC has the highest immune infiltrate. Due to this high level of immune infiltrate the work for this project was focused on this subtype of breast cancer <sup>179</sup>. (Figure 1.10)



**Figure 1.10. Subtypes of breast cancer.** The percentage of occurrence, markers for identification and prognosis of the different breast cancer subtypes.

## 1.9. Triple negative breast cancer

As mentioned above, TNBC is defined as the subtype of breast cancer where no hormone receptors can be detected immunohistochemically. Generally, TNBC tumours are larger in size with a higher grade and lymph node metastasis at diagnosis. The five-year survival rate in TNBC without recurrence is 79.6% compared to 84.6% for other breast cancers <sup>180</sup>. The chances of recurrence are also much higher in TNBC during the first 5 years. The rate of recurrence was seen to decline after 8 years <sup>181</sup>.

### 1.9.1 Subtypes of TNBC

To further improve therapy options for TNBC patients, attempts have been made to find subgroups within TNBC. Differences in treatment responses and survival have been shown for the different subtypes.

In 2011, from the analysis of 587 TNBC datasets, Lehmann et al. were able to identify 6 subtypes within TNBC. These included: basal-like 1 and 2 (BL1, BL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem like (MSL) and a luminal androgen receptor (LAR) <sup>182</sup>. The BL1 subtype has an enrichment of genes within the cell cycle and cell division pathways. Ki-67 is highly expressed, supporting the identification of a high proliferative phenotype. The BL2 subtype is enriched for the growth factor receptor genes *EGFR* and *MET*. The IM phenotype is associated with immune cell processes. The M subtype has high levels of cell

motility and cell differentiation pathway genes <sup>183</sup>. The MSL has similar properties but also has a lower rate of proliferation. The LAR group has an overexpression of the androgen receptor. Both the M and LAR subtypes showed the worst prognosis whilst the MSL subtype showed the best prognosis. Later in 2016 these subtypes were refined to just 4 subtypes: BL1, BL2, M and LAR. The identification of these subtypes was focused on responses to chemotherapy by the analysis of over 300 patients with neoadjuvant chemotherapy. From pathological analysis, and gene expression of both the tumour and stromal areas, they found that the level of tumour cells in the IM and MSL tissues were quite low with a high presence of immune cells and stromal cells. Therefore these subtypes were removed from the analysis. The reanalysis of the previous gene datasets showed the BL1 had the best prognosis and the BL2 and LAR group had the lowest response <sup>184</sup>. Similar subtypes were also identified by Burstein et al, where the basal like groups were divided based on the immune cell function <sup>185</sup>. **(Figure 1.11)**

Subtypes have also been classified by immunohistochemical staining of 13 antigens in 142 tumours. The antigens ranged from cytokeratins to markers such as p16, p53 and EGFR. The basal A cluster had high expression of p16 and an average to high expression of basal cytokeratins <sup>186</sup>. Ki67 was highly expressed. The basal B cluster had high levels of p53 with an average expression of Bcl2, CD117, WT1 and p16. The basoluminal cluster had high EGFR expression. The expression of basal and luminal cytokeratins varied. The luminal cluster had high levels of luminal cytokeratins. Basal A cluster showed the highest survival whilst the luminal group showed the worst overall survival <sup>186</sup>. **(Figure 1.11)**

Overall, the variations in the different subtypes emphasises the level of heterogeneity that is present in TNBC and how this is important for patient survival.

Lehman et al (2016)			
<b>Basal like 1</b> <ul style="list-style-type: none"> <li>- Highly proliferative (high Ki67)</li> <li>- Genes associated with cell cycle and cell division</li> <li>- Good prognosis</li> </ul>	<b>Basal like 2</b> <ul style="list-style-type: none"> <li>- Genes associated with growth factor receptors (EGFR)</li> <li>- Poor prognosis</li> </ul>	<b>Mesenchymal</b> <ul style="list-style-type: none"> <li>- Genes associated with cell motility and differentiation</li> <li>- Prone to resistance</li> </ul>	<b>LAR</b> <ul style="list-style-type: none"> <li>- High expression of androgen receptor</li> <li>- Luminal cytokeratins</li> <li>- Poor prognosis</li> </ul>
Burstein et al (2016)			
<b>Basal like immune suppressed</b> <ul style="list-style-type: none"> <li>- Downregulation of immune cell pathways</li> <li>- Poor disease free survival</li> </ul>	<b>Basal like immune activated</b> <ul style="list-style-type: none"> <li>- Upregulation of immune cell function</li> <li>- High STAT genes</li> <li>- Good prognosis</li> </ul>	<b>Mesenchymal</b> <ul style="list-style-type: none"> <li>- Cell cycle and DNA damage pathway genes</li> <li>- Claudin-low or mesenchymal stem-like</li> </ul>	<b>LAR</b> <ul style="list-style-type: none"> <li>- High expression of androgen receptor, prolactin and ErbB4</li> <li>- Oestrogen related genes</li> </ul>
Elsawaf et al (2013)			
<b>Basal A</b> <ul style="list-style-type: none"> <li>- Cytoplasmic p16</li> <li>- Basal cytokeratins</li> <li>- High proliferation</li> <li>- High grade</li> </ul>	<b>Basal B</b> <ul style="list-style-type: none"> <li>- High p53</li> <li>- Some expression of BCL2, CD117, WT1, p16</li> <li>- High grade</li> </ul>	<b>Basoluminal</b> <ul style="list-style-type: none"> <li>- High EGFR</li> <li>- Mixed expression of basal and luminal cytokeratins</li> <li>- Low survival</li> </ul>	<b>Luminal</b> <ul style="list-style-type: none"> <li>- Luminal cytokeratin CK7, CK18</li> <li>- Low survival</li> </ul>

**Figure 1.11. TNBC subtypes.** Subtypes identified from different papers using gene analysis and IHC analysis. <sup>184-186</sup>

## 1.9.2 Treatment

Patients with TNBC do not benefit from receptor targeting therapies such as trastuzumab and are currently limited to the options of chemotherapy, radiotherapy, and surgery.

Chemotherapy is the first line form of treatment for patients with both early and advanced stage TNBC. NICE guidelines recommend neoadjuvant chemotherapy with a platinum and anthracycline. Clinical trials have shown that neoadjuvant chemotherapy can increase pathological complete response by 10-15%. The expression of mutations can affect chemotherapy response rates with a significantly better response seen in *BRCA1/2* mutated tumours treated with carboplatin compared to docetaxel <sup>187</sup>. Combination chemotherapy is provided to patients with metastatic TNBC. However limited benefit is observed in survival.

Radiotherapy is offered to patients as an adjuvant therapy and has been found to improve survival of both older and younger patients <sup>188,189</sup>. The effects of

radiotherapy following mastectomies were dependent on subtype, with non-basal TNBC patients having a significant benefit in the reduction of recurrence whilst no effect was seen in basal-TNBC patients <sup>190,191</sup>

More targeted approaches have been sought after in TNBC. PARP inhibitors which target DNA-damage repair pathways have been approved in patients with *BRCA1/2* mutations. Olaparib and talazoparib showed a benefit in progression free survival when compared to chemotherapy. However no overall survival benefit was seen from the trials <sup>192,193</sup>. Targeting androgen receptors has also been trialled to try to combat LAR subtype tumours. Clinical trial results have not been promising, with some trials showing disease progression. Further clinical trials are being carried out with combination therapies <sup>194</sup>.

### 1.9.3 Immunotherapy

Immunotherapy has been discussed for the treatment of TNBC due to the high levels of immune infiltrate. From the clinical trials available on clinicaltrials.gov it is visible that 87 clinical trials are either recruiting, ongoing or completed for immunotherapy on TNBC. This is either as a monotherapy or a combination therapy.

The KEYNOTE 012 trial was a phase 1b trial that looked at the safety and activity of pembrolizumab in advanced PD-L1+ TNBC. The objective response rate from 32 patients was 18.5% <sup>195</sup>. The KEYNOTE 086 trial was a phase 2 study that looked at patients with metastatic TNBC in both PD-L1+ and PD-L1- groups. The use of pembrolizumab as a first line therapy in PD-L1+ patients showed a better response to those receiving the drug as second line therapy. No difference was seen between pretreated patients with both PD-L1+ and PD-L1- TNBC <sup>196</sup>. In the phase 3 trial KEYNOTE-119 there was no significant improvement with pembrolizumab when compared to chemotherapy <sup>197</sup>. Further clinical trials using other PD-L1 inhibitors avelumab <sup>198</sup> and atezolizumab <sup>199</sup> have been completed. These trials have also shown similar results with better responses as first-line therapy but a modest overall response.

Chemotherapy can lead to an increase in immunogenicity in tumours. This can be by increasing the number of antigens for immune activation as well as reducing immunosuppressive cells <sup>200,201</sup>. Therefore clinical trials have looked at the combination of chemotherapy and checkpoint inhibitors. The KEYNOTE-355 trial



compared chemotherapy plus pembrolizumab with chemotherapy plus placebo. The chemotherapy used in the trial was dependent on the physician and included paclitaxel, nab-paclitaxel, gemcitabine and carboplatin. An increase was observed in the progression free survival of the 'intention to treat' group from 5.6 months to 7.5 months <sup>202</sup>. The Impassion130 phase 3 trial combined nab-paclitaxel and atezolizumab. The addition of the checkpoint inhibitor increased the overall survival in the PD-L1+ patients from 18 to 25 months <sup>203</sup>. In 2019 the combination of nab-paclitaxel and atezolizumab was approved by the FDA for the use as a first-line therapy in late-stage TNBC. However, a further trial (Impassion131) which looked at the combination of paclitaxel with atezolizumab compared to a placebo found no benefit to patients irrespective of the PD-L1 status. The overall survival was also seen to reduce <sup>204</sup>. Differences in responses could possibly be explained by the responses to difference chemotherapies, gene status of patients and even immune cell infiltration. Due to no clear reason for the variations in results the combination treatment was withdrawn from clinics.

Positive results have been seen in early stage TNBC. The KEYNOTE-522 phase 3 trial showed that the addition of pembrolizumab with chemotherapy increased the pathological complete response from 51.2% to 64.8%. An improvement was also observed in PD-L1- patients <sup>205</sup>. In 2021, the FDA approved the use of pembrolizumab with chemotherapy as a neoadjuvant treatment in early stage high-risk TNBC. The combination of PD-1 and PD-L1 inhibitors with other therapies such as radiotherapies and PARP inhibitors are also being trialled <sup>206</sup>.

The use of immunotherapy as a treatment option in TNBC is promising, however the contrasting results seen with trials and different forms of chemotherapy show that a better understanding is required of the immunogenicity of cancers.

## 1.10. Hypothesis and Aims

As discussed in this chapter, matrix can act as a limiting factor towards T cell infiltration in the tumour. Through the analysis of matrix proteins associated with immunosuppression I identified VCAN expression to have an association with T cell location. Here, I test the hypothesis that VCAN inhibits T cell migration and therefore drives the production of an immune excluded phenotype. Therefore, targeting the key functional components of VCAN will restore T cell infiltration into the tumour epithelium, turning tumours into an inflamed phenotype.

The aim of the thesis was to explore the expression of VCAN in the different tumour immune phenotypes (desert, excluded and inflamed) and investigate the effect of the structural components of VCAN towards T cell infiltration.

This aim was achieved by the following objectives:

1. Identify the different tumour immune phenotypes in TNBC tissues and the association with VCAN localisation and expression. (Chapter 3)
2. Explore the association of VCAN post translational modifications of CS with tumour immune phenotype. (Chapter 4)
3. Optimise the isolation of VCAN from cell lines for use in functional assays. (Chapter 5)
4. Design models to test the effect of VCAN and its structural components in T cell invasion associated with tumour immune phenotype. (Chapter 6)

## Chapter 2: Materials and Methods

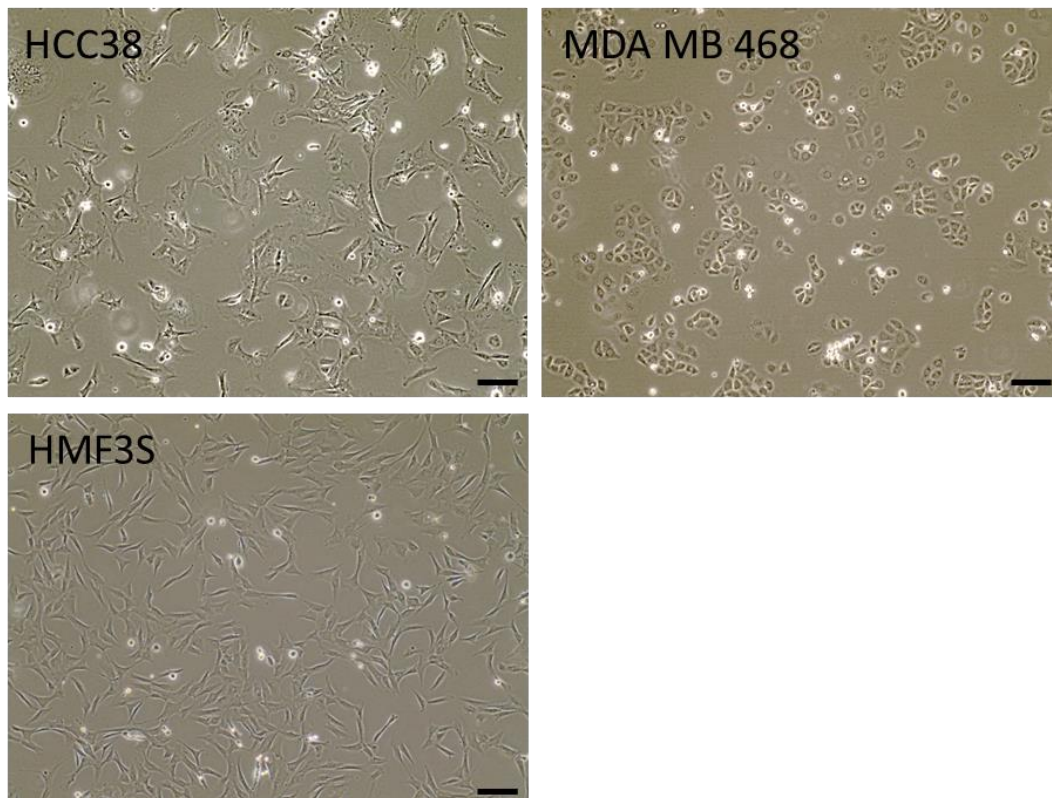
## 2.1. Cell lines and culture

TNBC cell lines HCC38 and MDAMB468 (Figure 2.1) were kindly gifted from Prof. John Marshall. The immortalised human mammary fibroblast cell line HMF3S (Figure 2.1) was purchased from Applied Biological Materials.

HCC38 cells were isolated from a mammary gland of a primary ductal carcinoma. Cells were grown in (Roswell Park Memorial Institute) RPMI medium (Gibco, 21875034) supplemented with 1% Glutamine (Sigma, G7513), 1% Penicillin/Streptomycin (P/S, Gibco, 15140122) and 10% Foetal bovine serum (FBS, Gibco, 10500064). MDAMB468 cells were isolated from the pleural effusion of an adenocarcinoma. Cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, 11965092) media supplemented with 1% Glutamine, 1% P/S and 10% FBS. HMF3S cells were immortalised by combined transduction with retroviruses carrying human telomerase (hTERT) catalytic subunit and temperature sensitive mutant of SV40 large T antigen. Cells were grown in DMEM/F12 medium (Gibco, 31331093) supplemented with 1% Glutamine, 1% P/S and 10% FBS.

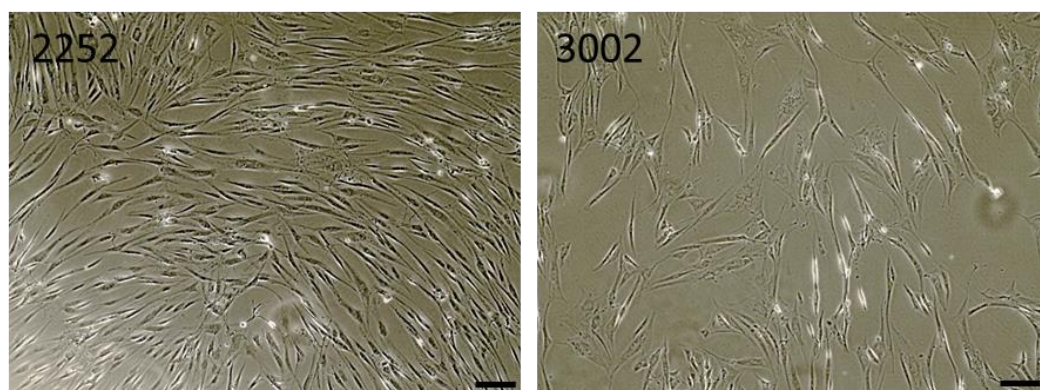
STR sequencing conducted by ATCC was used to authenticate the cells.

Mycoplasma testing was completed regularly with the MycoAlert PLUS mycoplasma detection kit (Lonza, LT07-710) with consistent negative results.



**Figure 2.1. TNBC cell lines HCC38 and MDAMB468 and immortalised mammary fibroblast HMF3S.** Magnification - 10x. Scale bar - 100µm. Images of cell lines taken with EVOS microscope. Magnification – 10x

Primary fibroblasts were obtained from the Breast Cancer Now tissue bank. Fibroblasts were obtained from healthy breast tissue (Figure 2.2), tumour tissue and surrounding tumour areas. The surround of the tumours is defined as over 5cm away from the tumour cells. Cells were grown in DMEM/F12 medium supplemented with 1% Glutamine, 1% P/S and 10% FBS.



**Figure 2.2. Fibroblasts from healthy patients.** Images of healthy (normal) fibroblasts taken with EVOS microscope. Magnification – 10x Scale bar - 100µm

Cells were split once a week with cells reaching a maximum 95% confluency. Cells were frozen down in freezing media made up of 10% Dimethylsulfoxide (DMSO) in FBS. Vials of cells were stored in liquid nitrogen.

## 2.2. Immunohistochemistry

### 2.2.1 Single stain

Tissues or collagen gels were fixed with 10% formalin (v/v) and embedded in paraffin by the pathology department. 4 or 5µm sections were sliced and mounted on charged slides. Slides were deparaffinised by leaving the slides at 60°C for 20mins to melt the paraffin and then submerging in Xylene for 3mins twice. The tissue was then rehydrated through a descending ethanol series of 100%, 90% (v/v), 70% (v/v) and 50% (v/v) ethanol for 2mins each. Slides were then placed in deionised water (dH<sub>2</sub>O). Antigen retrieval was completed using the antigen retriever (Aptum biologics). Depending on the stain either a citrate based buffer pH6 (Vector Laboratories, H-3300) or an EDTA (Ethylenediaminetetraacetic acid) based buffer pH9 (Abcam, ab93684) was used. Table 2-1 indicates the optimal form of retrieval per antibody. Slides were placed in the retriever for 25mins. Following this, slides were washed twice with phosphate buffered saline (PBS). Endogenous peroxidase was blocked using 3% H<sub>2</sub>O<sub>2</sub> (v/v) (Fisher Scientific, 10687022) diluted in PBS for 5mins or 0.3% H<sub>2</sub>O<sub>2</sub> diluted in methanol for 30mins. Slides were then washed with PBS-Tween 0.5% (v/v) (PBST) and then PBS for 3mins each. Slides were blocked with 2.5% goat serum (Vector Laboratories, S-1012-50) for 20mins. Primary antibodies were diluted in antibody diluent at the concentrations indicated in Table 2-1. Slides were incubated at either 1hr room temperature (RT) or overnight at 4°C. Slides were washed with PBST and then PBS for 3mins each. Impress HRP (horseradish peroxidase) secondary (Vector Laboratories, MP-7452, MP-7451) was used to detect the primary antibody, the solution was dropped on the slides and left for 25mins at RT. For the staining of biglycan, a biotinylated secondary antibody (Vector Laboratories, PI-9500) was used at 1:200 dilution for 30mins at RT and then treated with ABC solution (Vector Laboratories, PK-4000) for 20mins at RT. The ABC solution was used for hyaluronan staining as the hyaluronan binding protein was biotinylated. 3,3'-Diaminobenzidine (DAB) solution (Agilent, K3468) was made by adding 1 drop of reagent to 1ml of diluting solution. The DAB solution was added to the slides until a brown stain was observed. This time was kept consistent when using the same antibody for each batch of staining. Once the chromogen was

detected, slides were submerged in dH<sub>2</sub>O. Counterstaining was completed with 100% Gills I haematoxylin (Sigma, GHS1128-4L) for 30secs. Slides were then washed in tap water. Slides were dehydrated through an ascending ethanol series – 50% (v/v), 70% (v/v), 90% (v/v) and 100% for 2mins each. Slides were cleared in Xylene before mounting using DPX (Sigma, 06522) or vectamount (Vector Laboratories, H-5000).

### 2.2.2 Dual stain

For dual staining of slides, the procedure above was carried out up to the DAB staining with some modifications. Endogenous peroxidase and phosphatase was blocked using a dual-block solution (DAKO, S2003) for 10mins. Following the DAB staining, slides were washed in TBS-Tween 0.5% (v/v) (TBST). If the primary antibody was derived from the same species, a Fab fragment block step was included for 1hr at RT with a 1:100 dilution of the fragment (Jackson ImmunoResearch, 111-007-003, 115-007-003) to block any remaining sites that could be detected by the secondary antibody. The second primary antibody was added to the slides following the dilution in Table 2-1. Slides were incubated overnight at 4°C. Slides were washed with TBST and TBS for 3mins each. Impress AP kit (Vector Laboratories, MP-5401, MP-5402) or HRP kit was used to detect the primary. The reagent was dropped on the slide and left for 25mins at RT. Slides were then washed with TBST and TBS for 3mins each. To detect slides stained with the AP kit, Vector Red (Vector Laboratories, SK-5100) was used. To detect slides with the HRP kit, Vector VIP (Vector Laboratories, SK-4600) was used. Vector Red solution was made by diluting the reagents provided in the diluent. The stain was added to slides for 15mins at RT. Vector VIP was diluted by adding 3 drops of reagents in 5ml PBS. The stain was added to the slides for 14mins at RT. Slides were then washed in dH<sub>2</sub>O. Counterstaining was completed with 100% Gills I haematoxylin for 30secs. Slides were then washed in tap water. Slides were dehydrated through an ascending ethanol series. Slides stained with Vector Red were left in each ethanol dilution for 2mins each. Slides stained with Vector VIP were dipped in each dilution 7 times as the chromogen is sensitive to ethanol. Slides were cleared in Xylene before mounting using DPX or vectamount.

**Table 2-1. Immunohistochemistry antibodies.** IHC parameters for the different antibodies. PanCK (Pancytokeratin), COMP (Cartilage oligomeric matrix protein), CTSB (Cathepsin B), COL11A1 (Collagen Type XI  $\alpha$ 1), BGN (Biglycan). DCN (Decorin)

Marker	Antigen retrieval	Secondary	Dilution	RT or 4°C	Company	Code
<b><math>\alpha</math>-SMA</b>	pH6	Mouse	1:2000	RT	Sigma	A5228
<b>BGN</b>	pH6	Goat	1:500	RT	R&D	AF2667
<b>CD8</b>	pH9	Mouse	1:500	4°C	DAKO	M7103
<b>CD44</b>	pH6	Rabbit	1:1000	4°C	Sigma	HPA005785
<b>CD45</b>	pH9	Rabbit	1:100	4°C	CellSignalling	13917
<b>CD68</b>	pH9	Mouse	1:12000	4°C	Thermofisher	14-0688-82
<b>COL11A1</b>	pH6	Rabbit	1:100	RT	Sigma	HPA052246
<b>COMP</b>	pH6	Rat	1:75	RT	Abcam	AB11056
<b>CS</b>	pH6	Mouse	1:600	RT	Abcam	AB11570
<b>CTSB</b>	pH6	Rabbit	1:200	RT	Abcam	AB125067
<b>DCN</b>	pH6	Rabbit	1:500	RT	Proteintech	14667-1-AP
<b>FAP</b>	pH6	Rabbit	1:250	4°C	Abcam	AB207178
<b>FN1</b>	-	Rabbit	1:500	RT	Sigma	F3648
<b>Hyaluronan</b>	pH6	N/A	1:100	RT	Merck	385911
<b>PanCK</b>	pH6	Rabbit	1:1000	RT	DAKO	Z0622
<b>VCAN</b>	pH6	Rabbit	1:500	RT	Sigma	HPA004726
<b>VCAN (DPEEAE)</b>	pH6	Rabbit	1:400	RT	Abcam	AB19345

### 2.3. Masson's trichrome

Slides were deparaffinised and rehydrated as per the protocol for IHC staining. The slides were left in Bouin's solution (Sigma, HT10132) overnight at RT. Slides were washed under running tap water until all of the Bouin's solution was removed.

Weigert's haematoxylin solution (Sigma HT1079-1SET) was made up with an equal volume of part A and B. Tissues were drawn around with a hydrophobic (PAP) pen and the haematoxylin was added to the slides and incubated for 5mins. The haematoxylin was washed out under tap water for 2mins. Scarlet-Acid Fuchsin



(Sigma, HT151) was added to the slides and incubated for 15mins. Slides were rinsed in dH<sub>2</sub>O. Phosphotungstic acid (Sigma, HT152) and phosphomolybdic acid (Sigma, HT153) were diluted in water at a 1:1:2 ratio. The solution was added to slides and incubated for 10-15mins to remove the red staining from the collagen. Slides were rinsed in dH<sub>2</sub>O. Aniline blue (Sigma, HT154) was then added to the slides and incubated for 30mins. Slides were rinsed in dH<sub>2</sub>O. 1% acetic acid diluted in dH<sub>2</sub>O (Fisher Scientific, 10021123) was added to slides for 3mins. Slides were rinsed in dH<sub>2</sub>O then quickly rehydrated through 90% (v/v) and 100% ethanol. Slides were cleared in Xylene for 4mins then mounted with DPX.

## 2.4. RNAscope

### 2.4.1 Single

Single probe RNAscope was completed following the manufacturer's protocol (Advanced Cell Diagnostics (ACD), 322310). Within a week of the sectioning of Human TNBC FFPE tissues the protocol was carried out. Tissues were deparaffinised by being heated for 1hr at 60°C and then submerged in Xylene twice for 5mins. The slides were then submerged in 100% ethanol for 1min twice and then left to air-dry for 5mins. Tissues were outlined with a PAP pen. Hydrogen peroxide solution (ACD, 322330) was added to each slide and incubated for 10mins. Slides were washed in dH<sub>2</sub>O twice. Antigen retrieval was completed with 1X Target retrieval solution (ACD, 322000) for 15mins at 100°C. Slides were rinsed in dH<sub>2</sub>O by dipping 5 times and repeating in fresh dH<sub>2</sub>O. Slides were then dipped in 100% ethanol 5 times and left to air dry overnight. Protease plus solution (ACD, 322330) was added to each slide and slides placed in HybEZ Oven (ACD, PN 321710) set to 40°C for 30mins. Slides were washed in dH<sub>2</sub>O. The probe (ACD, 430071) was then added to the slides and incubated at 40°C for 2hrs. The slides were submerged in 1X wash buffer (ACD, 310091) for 2mins at RT with agitation. AMP1 was added to slides and incubated for 30mins at 40°C. Slides were washed in 1X wash buffer twice in between AMP incubations. AMP2 was added to slides with a 15min incubation at 40°C. AMP3 was added with a 30min incubation at 40°C. AMP4 was added and incubated for 15mins at 40°C. AMP5 was added with a 30min incubation at RT. AMP6 was added and incubated for 15mins at RT. The slides were washed and the DAB substrate was prepared by mixing equal volumes of solution A and B. The DAB solution was pipetted onto slides and incubated for 10mins at RT. The reaction was stopped by submerging slides in dH<sub>2</sub>O. 50% haematoxylin was made

by diluting Gills I haematoxylin in dH<sub>2</sub>O. The counterstain was completed for 2mins at RT. Slides were washed in dH<sub>2</sub>O and then submerged in 0.02% Ammonia to turn haematoxylin blue. Slides were washed in dH<sub>2</sub>O and dehydrated in 70% (v/v), 90% (v/v) and 100% ethanol for 2mins. The slides were cleared in Xylene for 5mins. Slides were left to dry at RT and mounted with DPX.

## 2.4.2 Duplex

For the duplex kit to detect 2 markers the manufacturer's protocol (ACD, 322430) was followed. The protocol follows the same process as the single marker detection up to the point of drying the slides. Slides were dried for 5mins following antigen retrieval. The protease plus reagent was then added to the slides and incubated at 40°C for 30mins. The probe solution was made with a 1:50 ratio of channel 2(ACD, 430071-C2) to channel 1(ACD, 490251). The slides were washed in dH<sub>2</sub>O and probes added to slides and incubated for 2hrs at 40°C. Slides were washed in 1X wash buffer and then placed in 5X SSC buffer overnight. AMP1-6 were the same as above. The first channel was then detected with a red signal. The detection solution was made by diluting Red B to Red A in a 1:60 ratio. The red solution was added to the slides and incubated at RT for 10mins. The slides were washed in 1X wash buffer twice and also between AMP incubations. AMP 7 was added to slides and incubated for 15mins at 40°C. AMP 8 was added and incubated for 30mins at 40°C. AMP9 was incubated for 30mins at RT and AMP10 for 15mins at RT. The green signal was then detected. The green solution was made with Green B to Green A at a 1:60 ratio. The solution was added to slides and incubated for 10mins at RT. Slides were rinsed quickly in dH<sub>2</sub>O. The slides were then dipped in 50% Gills I haematoxylin for 30secs and quickly washed in tap water for 30secs. The slides were then dried for 15-30mins at 60°C and mounted with DPX.

## 2.5. Image analysis

### 2.5.1 Stain deconvolution

FIJI (GNU General Public License) was used to deconvolute the chromagen stains from the IHC images. The plugin color deconvolution was used with the H-DAB option selected to extract DAB stains. For the Vector Red stains the H-AEC option was selected.

### 2.5.2 Co-localisation analysis

Tissue microarray (TMA) cores were selected based on the cores being intact for all 5 stains. Images were taken of the cores in all the images and they were aligned on Fiji. The DAB stain was deconvoluted and the stains were set to a threshold. The threshold was kept the same for images with the same marker. The co-localisation threshold plugin was used to form red-blue images. Output images were formed with blue pixels for stain 1, red pixels for stain 2 and white pixels for areas of co-localisation. The colour of the images was changed using the dichromacy setting with deuteranope as the selection to make the colours clearer. The color counter plugin was used to count the number of pixels for each colour.

### 2.5.3 Image overlay

To overlay stains from multiple markers, the images were aligned and then deconvoluted. A threshold was applied to the stains and a colour of either blue, red, yellow or green was applied to the positive areas. Using the overlay image option, images were overlapped at 100% with the removal of null areas.

### 2.5.4 Positive cell counting

QuPath<sup>207</sup> was used to count the number of positively stained cells in annotations (regions of interest). The colour detection vectors were changed to match the colour of the stained cells. Annotations were made for the area that was to be analysed. The positive cell count tool was optimised for each cell type and stain, based on the cell size and stain threshold of multiple images. The analysis was run as a batch for each stain.

### 2.5.5 Stain area detection

A thresholder was created for pixel classification on QuPath. The colour detection vector was changed to match the stain to be detected. The resolution was set to full and the channel was dependent on the stain used – either DAB, Vector Red or VIP. The prefilter was set to Gaussian and the threshold was dependent on the stain. Pixels above the threshold were labelled as positive and below were labelled as negative. The threshold was applied to all annotations to provide a percentage for

the number of positive pixels. The same threshold was then applied to the whole batch. A stain score was calculated by determining the level of matrix marker at a low, medium and high threshold. The equation for the stain score was  $\text{Score} = (\% \text{ low intensity} \times 1) + (\% \text{ medium intensity} \times 2) + (\% \text{ high intensity} \times 3)$ .

#### 2.5.6 Area classification

For QuPath to identify cells which were within the tumour and cells which were stromal, areas needed to be selected for training. Using PanCK staining as a guide, annotations were made on QuPath using the brush tool where the tumour area was classed as tumour and on the stroma area and classed as stroma to help train the software. The “Train object classifier” was selected and all classed annotations were selected. Cells were detected and the classifier was applied. More areas were selected for training, if the classifications were not correct. The classifier was then saved. Regions of interest were selected, and positive cell count was completed for staining. The classifier was then applied to the regions of interest. Cells were then annotated as being positive for the marker and also which area they were in based on the classifier.

#### 2.5.7 Tissue phenotype analysis

Areas of 2mm by 2mm were drawn at multiple areas around the tissue. Each area was at the tumour-stroma border containing around 50% tumour epithelium. Within each area, pixel classification was used to determine the percentage of matrix markers whilst positive cell count was used to identify the amount of CD8+ and CD68+ cells present. Area classification was used to determine the number of CD8 and CD68 positive cells in the tumour and the stroma. PanCK was used to determine the tumour area.

#### 2.5.8 Immune exclusion analysis

Following the Kather et al<sup>44</sup> paper, areas were outlined around the tumour-stroma border. On QuPath using the brush tool, a 100µm margin was drawn on either side of the border to form the Inner Invasive and Outer Invasive areas. On either side of these areas the stroma and tumour core areas were drawn. The number of cells positive for CD8 and CD68 were counted in each area and the percentage of pixels

of each matrix marker. This was repeated across multiple areas in the tissue. The data were then analysed to compare the number of immune cells in each region to identify the tissue immune phenotype of desert, excluded and inflamed. This is explained in detail in Chapter 3.

## 2.6. Gel models

### 2.6.1 Collagen gel model

Collagen gels were made with 3mg/ml rat tail collagen (ThermoFisher, A1048301). The ratio of components are shown in Table 2-2 and were dependent on the density of the gel. For gels containing cells, the number of cells were pelleted and resuspended in the required media for the gels. The collagen was first mixed with the 10X DMEM (Sigma, D2429) and the solution was mixed to turn to a yellow solution. Then 1M NaOH was added to form a pink solution, the solution was mixed and left in ice to prevent premature gelation. The cell suspension was then mixed into the gel to ensure that the cells were distributed evenly. 100µl of the gel solution was then pipetted into each well of a 96 well plate. The gels were left to set at 37°C for 1hr before taking the gels out with a spatula and placing them into a 24 well plate with the respective media for the cell line. Gels containing crude or enriched protein had the required amount of protein with PBS added prior to the addition of 1M NaOH. The volume for the cell suspension was adjusted accordingly.

**Table 2-2. Reagents for collagen gels of different stiffness.** Amount of reagent in µl.

	3mg/ml Collagen	10X DMEM	1M NaOH	Cells + media
0.05%	16.7	5	2	76.3
0.1%	33	5	2	60

### 2.6.2 Embedded gel model

To make the embedded gels, 0.1% (w/v) collagen gels were made as mentioned above with 300,000 tumour cells. The gels were treated with either only 50µg/ml ascorbic acid or with 10µg/ml TGFβ3 (Peprotech, 100-36E) and ascorbic acid in the media. The media was replenished after 3 days with the same treatments. At day 7 another gel was made with 1,000,000 HMF3S cells to a volume of 150µl per gel. 50µl of the gels were pipetted into a well and left to stand at RT for a few minutes. The tumour gel was then added into the centre of the well and the remaining 100µl gel pipetted into the well. The completed gel was then left to set at 37°C. The gels

were scooped and placed in wells containing either only 50µg/ml ascorbic acid or 10µg/ml TGFβ3 with ascorbic acid. The gels were cultured for 7 days. The gels were then transferred to a round bottom 96 well plate and 200,000 activated T cells were added to each well. The T cells were left to invade into the gel for 3 days. The gels were fixed with 4% formalin overnight at 37°C then washed with PBS and embedded in an agarose gel for paraffin embedding and sectioning.

## 2.7. RNA extraction

RNA was extracted from cell lines, frozen tissues and collagen gels. The extraction was completed using a Qiagen Mini kit (Qiagen, 74104) and manufacturer's protocols were followed.

Frozen tissues were initially sliced using a microtome by the pathology facility. The tissues were then further broken down using gentleMACs tubes (Miltenyi, 130-093-237) in 500µl RLT buffer solution. The RLT solution containing the digested tissue was then added to a Qias shredder to further homogenise the sample. The Qias shredder was spun at 8000g for 1min. Collagen gels were left to dissolve in 350µl RLT buffer in an eppendorf. The gels were pipetted up and down to break down the gel.

Cells were seeded in a 6 well plate at 100,000 cells per well and cultured for 7 days. Cells were seeded in a 24 well plate at a density of 40,000 cells for HCC38, 30,000 cells for MDAMB468 and 25,000 for HMF3S. This was dependent on the growth of the cells to ensure they were at a similar confluency at 7 days. For 6 well plates, 350µl RLT buffer was added to each well. For 24 well plates, 250µl RLT buffer was added to each well.

70% ethanol made in RNase Free water was added to each sample at an equal volume to the RLT buffer. The samples were then added to the mini columns and spun at 8000g for 21secs. This was repeated till all the sample was transferred through the column. The flow through was discarded and the 600µl RW1 was added to each well. The sample was spun at 8000g for 21secs and the flow through discarded. 500µl RPE buffer was added to each sample and the column was spun at 8000g for 21 secs. The flow through was discarded and this was repeated again. The column was then placed in a clean 2ml centrifuge tube to be spun again at 8000g for 1min to remove any buffer remnants. The RNA was then eluted from the column by adding 30µl RNase free water. The column was spun at 8000g for 1min.

The RNA concentration was determined using a nanodrop and quality of RNA was determined by ensuring the 260/280 ratio was around 2. The isolated RNA samples were stored at -80°C until required.

## 2.8. qRTPCR

### 2.8.1 cDNA transformation

To reverse transcribe the RNA to cDNA for qRTPCR (Quantitative reverse transcription polymerase chain reaction) analysis, the high-capacity cDNA reverse transcriptase kit (ThermoFisher, 4368814) was used. For the reverse transcription 250ng-1µg of RNA was used. The amount was dependent on the concentration of RNA. To each reaction, 2µl of 10X RT buffer, 0.8µl of 25X dNTP Mix, 2µl 10X RT Random Primers and 1µl Multiscribe Reverse Transcriptase was added. The amount of RNA required was added and then nuclease free water was added to total the volume to 20µl. The samples were then placed in a BioRad T100 ThermoCycler. The sample was heated to 25°C for 10mins, then 37°C for 120mins and 85°C for 5mins before being held at 4°C. The cDNA was diluted with nuclease free water to give a final concentration of 5ng/µl. The cDNA was stored at -20°C until required.

### 2.8.2 qRTPCR

Primers to detect genes of interest were selected from publications<sup>65,208-211</sup> and then cross-checked using the NCBi blast for specificity. The sequence for the primers are shown in Table 2-3. Primers were ordered from Invitrogen and diluted to a concentration of 100µM. A working stock solution containing both the forward and reverse primer was made to a final concentration of 10µM.

**Table 2-3. Primer sequences for each gene detected with qRTPCR.**

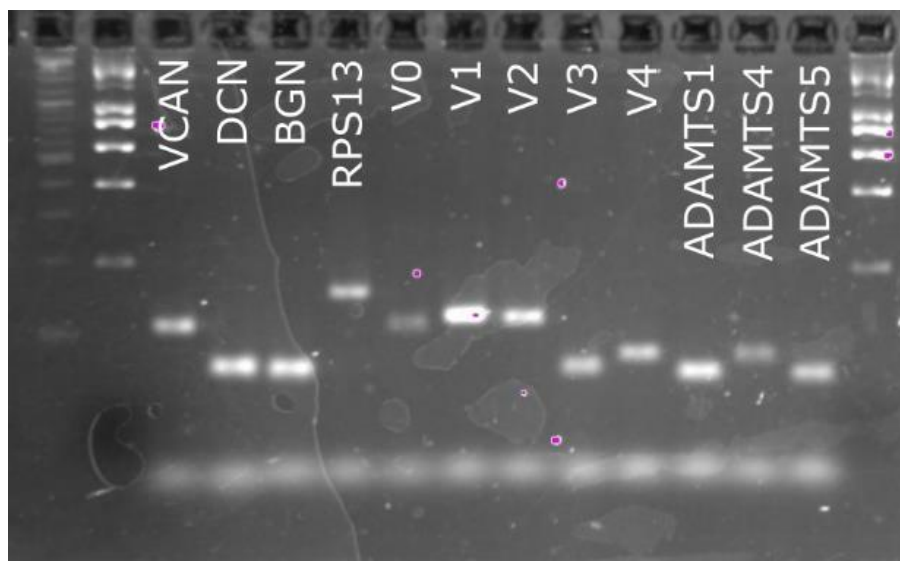
<b>Gene</b>	<b>Forward Primer</b>	<b>Reverse Primer</b>	<b>Amplicon bp</b>
<i>VCAN</i>	GTAACCCATGCGCTACAT AAAGT	GGCAAAGTAGGCATCGTT GAAA	110
<i>DCN</i>	GAAGTTCCTGATGACCGC GA	AAGATGGCATTGACAGCG GA	72
<i>BGN</i>	CACCAAAGTGGGTGTCAA CG	GATGCCGTTGTAGTAGGC CC	70
<i>RPS13</i>	TCGGCTTTACCCTATCGAC GCAG	ACGTACTTGTGCAACACC ATGTGA	153
<i>V0</i>	GCACAAAATTTACCCTGA CAT	TTAGATTCTGAATCTATTG GATGACCA	112
<i>V1</i>	CCCAGTGTGGAGGTGGTC TAC	CGCTCAAATCACTCATTCTG ACGTT	126
<i>V2</i>	CCCAGCAAGCACAAAATTT CAC	TAGGATAACAGGTGCCTC CGTT	122
<i>V3</i>	CCCTCCCCCTGATAGCAG AT	GGCACGGGGTTCATTTTG C	72
<i>V4</i>	CAGTACCACTGTTGAGGA AAAGAAAA	CGTTAAGGCACGGGTTCA TT	86
<i>ADAMTS1</i>	CAGCTTTCTTGCCATCAAA GCT	GGTGGACAAAGTGTAGTC ACCATTA	67
<i>ADAMTS4</i>	TCATCACTGACTTCCTGGA CAA	GAAAGTCACAGGCAGATG CA	83
<i>ADAMTS5</i>	AAATTCTGTGAAGAGACCT TTGGT	GCTGGTAAGGATGGAAGA CATTAA	66

For the qRTPCR reaction, PowerTrack SYBR Green Master Mix (ThermoFisher, A46109) was used. The sample was used in a 5µl reaction in a 384 well plate. The SYBR green solution was made with 2.5µl SYBR green and 0.1µl primer. The cDNA mix contained 1µl cDNA, 0.12µl sample buffer and 1.28µl nuclease free water. In each well of the plate 2.6µl of the SYBR green solution was added and then 2.4µl of the cDNA mix. The sample was run on the Quant Studio 7 (ThermoFisher). The first stage of the cycle (Hold stage) was at 95°C for 20secs. The second stage was the PCR stage where there were 40 cycles consisting of 95°C for 1sec and 60°C for 20secs. The final stage produced the melt curve where samples were at 95°C for 15secs and then 60°C for 1min and then 95°C for 15secs. The melt curves for each primer are shown in Appendix 1. The cycle threshold (Ct) was determined for each marker and normalised to the Ct value of RPS13 to give  $\Delta C_t$ . RQ was determined as  $2^{-\Delta C_t}$ .



### 2.8.3 Agarose gel electrophoresis for DNA

The amplicon from the qRT-PCR was checked by running the amplified DNA in an agarose gel. The agarose gel was made with 2% agarose in 1X TBE and 0.2ug/ml ethidium bromide. 15µl of the sample was pipetted into the wells of the gel and 5µl 1Kb ladder was added on either end of the samples. The gel was run for 2hrs at 80V. The gel was then visualised using the Chemidoc (Amersham). The bands are shown in Figure 2.3.



**Figure 2.3. Amplified products from qRT-PCR.** Agarose gel electrophoresis of amplified product following qRT-PCR reaction.

## 2.9. VCAN enrichment

The buffers used for the chromatography methods are listed in Table 2-4:

**Table 2-4. Buffer composition for chromatography.**

Buffer	Buffer composition
Wash buffer	0.15M NaCl, 0.05M Tris
Elution buffer 1	0.3M NaCl, 0.05M Tris
Elution buffer 2	0.5M NaCl, 0.05M Tris
Elution buffer 3	1M NaCl, 0.05M Tris
Elution buffer 4	1.5M NaCl, 0.05M Tris
Elution buffer 5	4M GuHCl, 0.05M Tris

## 2.9.1 Affinity chromatography

### 2.9.1.1 Conjugating hyaluronan to beads

10ml of Carboxylink beads (ThermoFisher, 20266) was taken into a 50ml falcon and washed with dH<sub>2</sub>O. 0.1M MES (2-(N-morpholino)ethanesulfonic acid) buffer was made and the pH was adjusted to 4.7 using HCl. 25mg of HA was dissolved in 25ml 0.1M MES buffer. The beads were allowed to settle before the water was removed. This was repeated twice to remove the buffer. The 1mg/ml HA solution was added to the beads. The pH of the solution was checked and adjusted to 4.7 if required. 0.1g of 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) was dissolved in 1ml MES buffer and then added to the bead solution. The beads were left on a roller for 3hrs at RT and then left to stand at RT overnight.

The HA solution was removed from the beads and 10ml 1% acetic acid was added to the beads. The beads were left on a roller for 10mins, and then left to settle before removing the acetic acid solution. 10ml of 1M NaCl was added and the beads were put on a roller for 10mins. After the beads had settled, the solution was removed and 10ml 0.5M formic acid was added to the beads. The beads were placed on the roller for 10mins and then left to settle. The solution was removed and the beads were washed with dH<sub>2</sub>O. The beads were then left in 0.5M Sodium acetate and 0.02% sodium azide solution. The conjugated beads were stored at 4°C.

### 2.9.1.2 Affinity chromatography with HA-beads.

To pack the column, the solution was removed from the beads and 20ml wash buffer was added to form a slurry. The slurry was poured into a 30ml gravity flow column (BioRad, 7321010) with the stopper closed. Once the beads settled the stopper was opened to allow the solution through. 10ml wash buffer was then added to the column and allowed to flow through. Protein free media was poured into the column and the flow through was collected. 10ml wash buffer was then added to the column followed by 10ml of each elution buffer. The wash and elutions were all collected in test tubes and placed on ice during the elutions. After the final elution, the beads were washed with wash buffer and then 0.5M Na-acetate with 0.5% sodium azide was added to the beads. The column was stored at 4°C.

The FT and elutions were tested for VCAN and the positive samples were stored at -20°C in falcon tubes.

### 2.9.2 Ion exchange

DEAE methacrylate beads were used as a weak anion exchange medium. 10ml beads were pipetted into a 50ml falcon tube and washed with dH<sub>2</sub>O to remove the buffer. They were then washed in wash buffer twice. Protein free media from the cell culture was spun down to pellet the cell debris. 100ml media was poured into a 150ml conical flask. The washed beads were added to the flask using the media to remove all the beads from the falcon. The beads were left on a shaker at 4°C for 2hrs. The bead slurry was poured into a 30ml empty gravity flow column. The stopper on the column was opened and the flow through was collected. Once the beads had settled, 10ml wash buffer was added to the column and collected. The beads were then eluted with 10ml of each elution buffer. The beads were washed with wash buffer and then stored in wash buffer at 4°C.

The FT and elutions were tested for VCAN and the positive samples were stored at -20°C in falcon tubes.

For use in functional assays, VCAN positive samples were sterile filtered. 100kDa centrifugal filters (ThermoFisher, 88524) were sterilised with 70% ethanol. Samples were concentrated and dialysed against sterile water using the centrifugal filters. Concentrated and dialysed protein samples were stored at -20°C short term and -80°C long term. A small aliquot of each sample was taken to quantify protein level.

### 2.10. Dot blot

PVDF membrane was activated in 100% methanol. The membrane was then washed in dH<sub>2</sub>O. A filter paper was dampened with water and placed on a thin sponge. The activated PVDF was placed on the filter paper. 10-20µl of each sample was pipetted onto the membrane. The membrane was left to stand till the sample had gone through the membrane and the protein was bound. Protein was detected on the membrane by staining with ponceau. The ponceau was washed with dH<sub>2</sub>O and the membrane was blocked with 5% skimmed milk (Sigma, 70166) in TBST for 15mins. The primary antibody was then diluted in the 5% milk and left on the membrane for 1hr at RT. The dilutions for each antibody are shown in Table 2-5. The membrane was washed with TBST twice for 3mins each. The secondary antibody was diluted in 5% milk at 1:5000. The membrane was incubated with the secondary for 30mins at RT. The membrane was washed with TBST three times for 3mins each. The membrane was then developed using ECL solution (Merck,

WBLUC0100) on a chemidoc (Amersham ImageQuant 600). For detection, the chemidoc was used at incremental settings with an image every 15secs.

**Table 2-5. Antibody dilutions for immunoblotting.**

Antibody	Dilution	Secondary	Company	Code
VCAN	1:500	Mouse	DSHB	12C5
BGN	1:1000	Goat	Novus	AF2667
DCN	1:1000	Mouse	Proteintech	14667-1-AP

## 2.11. Protein concentration using BCA

Protein standards were made using bovine serum albumin (BSA). Standards of 0.05, 0.1, 0.2, 0.5, 0.8, 1 and 2mg/ml were made. 10µl of each standard was added to a well of a 96 well plate. 2µl of samples were added to each well with 8µl of dH<sub>2</sub>O. Each standard and sample was pipetted as a duplicate. BCA solution was made by diluting the copper (II) sulphate solution in bicinchoninic acid solution at a 1:50 ratio. 200µl of the BCA reagent was added to each well. The plate was covered in foil and left at 37°C for 30mins. The absorbance of the plate was then read on a plate reader at 562nm. A standard line was then formed using the standard samples and the levels in each sample were determined by interpolating the line and multiplying the value by the dilution factor of 5.

## 2.12. Gel electrophoresis

Samples were run a 3-8% Tris-acetate gel unless otherwise stated. The samples were loaded into the gel and run at 120V for 1.5hrs with 1X Tris-Acetate buffer. For coomassie staining, gels were washed with dH<sub>2</sub>O and then coomassie solution (0.1% Coomassie G250 (Sigma, 1154440025), 25% methanol, 5% acetic acid) was added to the gel and left to stain for 30mins at RT on a plate shaker. The gel was then washed with dH<sub>2</sub>O twice and left in dH<sub>2</sub>O overnight to remove the background coomassie staining. Gel was imaged on the chemidoc.

## 2.13. PBMC isolation from leukocyte cones

Leukocyte blood cones were obtained from the St Bartholomew hospital. Under the hood, the cone was emptied into a 50ml Falcon. The remaining blood was collected

by rinsing the cone with cell buffer (PBS + 2mM EDTA). The blood was diluted 1:10 with cell buffer. 15ml Ficoll-Paque (VWR, 17-1440-03) was added to a new 50ml falcon tube. Slowly the diluted blood was added to the Ficoll-Paque, with caution to not disturb the Ficoll surface. The sample was centrifuged at 1400g for 20mins at RT without brakes. A PBMC (peripheral blood mononuclear cells) layer was formed between the Ficoll and the plasma. The PBMC layer was collected and resuspended in cell buffer. The PBMCs were centrifuged at 300g for 10mins at 20°C. The supernatant was removed to discard any platelets. The wash was repeated twice. The cells were resuspended in 50ml cell buffer and cells were counted at a 1:10 dilution.

### 2.13.1 T cell isolation from PBMC

T cells were isolated from the PBMCs using the EasySep Human T cell Isolation Kit (Stemcell, 17951). The PBMCs were pelleted and resuspended in cell buffer at  $5 \times 10^7$  cells per ml. 1ml of cells were added to a 5ml round bottom falcon tube. 50 $\mu$ l of isolation cocktail was added to the sample. The sample was mixed by pipetting up and down then left to incubate for 5mins at RT. The RapidSpheres were vortexed for 30 seconds and 40 $\mu$ l was added to the sample. Isolation buffer (cell buffer + 0.5% FBS) was added to the beads to top up volume to 2.5ml. The tube was placed in the EasySep magnet and left to incubate for 3mins at RT. With the tube still in the magnet, the solution was poured into a new falcon tube. The solution collected contained the isolated T cells. Isolated T cells were cryo-preserved in freezing media (10% DMSO in FBS).

### 2.14. T cell activation

T cells were activated using TransAct nanobeads (Miltenyi Biotec, 130-111-160). TransAct beads activate T cells with CD3 and CD28. Isolated T cells were resuspended in DMEM/F12 media with 10% FBS, 1% Glutamine and 1% P/S at  $1 \times 10^6$  cells per ml. 200 $\mu$ l of the T cell solution was added to a well of a 96 well plate to use as a control for the activation. 10 $\mu$ l of Transact per  $1 \times 10^6$  cells was added to the T cell solution. The T cells were plated into the 96 well plate at 200 $\mu$ l per well. T cells were left to activate for 4 days before use. Once activated, T cells were lifted from wells by pipetting up and down. Cells were pelleted and resuspended in media at desired cell density.

## 2.15. Flow cytometry

Flow cytometry was used to confirm the isolation and activation of T cells. Cells were spun at 2000rpm for 2mins. Supernatant was removed and cells were washed with PBS. The cells were pelleted at 2000rpm for 2mins. Fixable viability dye (FVD) was diluted in PBS at 1:1000. 100µl was added to the cells. The cells were mixed and left in the dark at RT for 20mins. Fluorescence activating cell sorting (FACs) buffer was made with 2% BSA, 2mM EDTA in PBS. Antibody mastermix was made with all the antibodies as well as solutions which contained all antibodies by one. Solutions were made for 50µl per sample. Cells were pelleted and resuspended in PBS. The cells were split equally into wells of a V-bottom 96 well plate. The plate was spun at 2000rpm for 2mins and the PBS was removed. 50µl of antibody mastermix or FMO solution was added to the cells. The plate was covered in foil and incubated at 4°C for 20mins. The cells were pelleted and 150µl of FACs buffer was added to each well. This was repeated and cells were resuspended in 200µl FACs buffer. The samples were fixed with 2% formaldehyde for 10mins at 4°C. The cells were pelleted, and cells were washed with FACs buffer. This was repeated and then cells were suspended in 200µl FACs buffer and transferred to FACs tubes. The tubes were stored at 4°C in the dark before analysis.

Staining panel for T cell isolation and activation: FVD (Aqua, 1:1000), CD45 (PE/Cy7, 1:100), PD1 (PerCP, 1:100), CD8 (AF700, 1:100), CD25 (BV421, 1:250), CD4 (BV605, 1:250), Tim3 (PE, 1:100)

Samples were analysed using a 4-laser Fortessa flow cytometer (BD). 5,000 live events were collected for each sample. Data was analysed using FlowJo V10. Flow cytometry data showing T-cell isolation and activation shown in Appendix 2.

## 2.16. Transwell assay

5 Units CSase (Sigma, C3667) was reconstituted in 0.01% BSA. 5X CSase buffer was made with 200mM TrisHCl, 200mM Sodium acetate pH8. Protein was CSase treated with 1µl CSase and buffer added to make 1X. ADAMTS4 (Biotechne, 4307-AD-020) was used at 0.5µg per 10µg of protein. Proteins were treated with CSase or ADAMTS4 overnight at 37°C. Collagen gel solutions were made up to form a 0.05% collagen gel with the desired protein added with PBS in replacement of the cell suspension. NaOH was added to the gels just before they were added to the wells. The upper chamber of the Clearview 96-well Plate for Chemotaxis (Sartorius,

4582) was coated with 15µl collagen gel containing the protein solution. Gels were allowed to set at 37°C for 20mins. SDF1 was diluted to 100ng/ml in DMEM/F12 media with 10% FBS, 1% glutamine and 1% P/S. 200µl of the SDF1 solution was added to each well of the bottom chamber. 40µl serum free media (SFM) (DMEM/F12 with 1% Glutamine and 1% P/S) was added on top of the gels. Activated T cells were counted and resuspended in SFM to give 10,000 cells per 50µl of media. 50µl of cell suspension was added to each well and plate was added to the Incucyte (Sartorius). The Incucyte was run using the chemotaxis module and the top of the membrane was scanned every 1.5-2 hours for 2 days. The scans were then analysed using the Incucyte software to count the number of T cells that had invaded through the gel.

## 2.17. Gel invasion assay

100ng/ml SDF1 was diluted in PBS. 50µl of each solution was added to a 96 well plate. The plate was left at 37°C for 3hrs. The solution was removed from the plate by tipping the plate upside down on tissue paper. PBS was then added to the wells and was removed by tipping plate upside down on tissue paper. The plate was left to airdry in the hood. 0.05% collagen gel solutions were made containing the protein solution. 50µl of collagen gel solution was added to the coated wells of the 96 well plate. The gel was left to set for 1hr. Activated T cells were seeded onto the wells at a density of 20,000 cells per well. The plate was left in the incubator for 3 days to allow cells to invade. The media was removed from the plate and the gels were fixed with 4% formalin at 4°C overnight. DAPI (4',6-diamidino-2-phenylindole) was diluted 1:2000 in PBS and added to the gels. Gels were stained with DAPI for 20mins at RT in the dark. The gels were washed and imaged using the ECHO Revolve microscope. Z-stack images were taken of each well.

## 2.18. Mass spectrometry

20µg of protein was diluted with 1M HEPES buffer to a volume of 100µl. 2.5µl 1M DTT (Dithiothreitol) was added to the sample. Samples were vortexed and left in an incubator at 23°C for 1hr in the dark. 5µl 45mM Iodoacetamide was added. The sample was vortexed and incubated at 23°C for 1hr in the dark. 1500U of PNGase F was added to the sample. The sample was vortexed and incubated at 37°C for 2hrs in the dark. 10µg GluC (endoprotease GluC) was solubilised in 50µl mass

spectrometry grade water (MS water). 1µg GluC was added to the sample. The sample was vortexed and incubated at 37°C overnight. 30µl 5X CSase buffer was added with 1.5µl 10U/ml CSase. The sample was vortexed and incubated at 37°C for 2hrs. Samples were then desalted using C18 desalting spin tips (ThermoScientific, 84850). The C18 spin tips were primed with 200µl acetonitrile (ACN). The tip was placed into an eppendorf and spun at 1500g for 3mins at 4°C. The flow through was discarded. 200µl wash buffer (0.1% Trifluoroacetic acid (TFA) (w/v), 1% ACN (v/v) and 99% dH<sub>2</sub>O (v/v)) was added. The tip was spun at the same parameters and flow through discarded. This was repeated. The sample was then added to the tip and spun at 1500g for 3mins at 4°C. The flow through was discarded and the tip was placed into a clean eppendorf. 250µl elution buffer (70% ACN (v/v), 30% dH<sub>2</sub>O (v/v), 0.1% TFA (w/v)) was added to the tip and the protein was eluted by spinning the column at 1500g for 3mins at 4°C. This was repeated to elute all the protein in the tip. The eluted protein was lyophilised in a speed vac (ThermoFisher Savant SPD1010) on a manual run for 6hrs. Samples were then reconstituted in 20µl 0.1% TFA (w/v). Samples spun at 13000rpm at 4°C for 5mins. Samples placed in a vial and run in the mass spectrometer (ESI-TRAP). The analysis of the mass spectrometry data was completed using Mascot Distiller. The parameters are listed in Table 2-6.

**Table 2-6. Mass spectrometry analysis parameters.**

Database	SwissProt_2021_02
Taxonomy	Homo sapiens (human) (20,396 sequences)
Search	MS/MS Ion search
Enzyme	V8-DE
Fixed Modifications	Carbomidomethyl
Variable modifications	CS-0S, CS-1S, Gln->pyro-Glu (N-term Q),Oxidation (M),Phospho (ST),Phospho (Y)
Mass values	Monoisotopic
Protein mass	Unrestricted
Peptide mass tolerance	10 ppm
Fragment mass tolerance	25 mmu
Peptide charge	2+ and 3+
Max missed cleavages	2



## 2.19. CS Mass spectrometry

500ml 50mM ammonium bicarbonate buffer (pH 7.6) was made in MS water. The buffer was heated at 95°C for 1hr to kill off bacteria that could degrade the glycans. 5 units of CSase was reconstituted in 1ml ammonium bicarbonate buffer to a concentration of 5U/ml and stored at -20°C.

### 2.19.1 Tissue sections

TNBC FFPE blocks were sectioned to 4µm per slide. One slide was stained for PanCK to identify tumour regions. Sections for CS extraction were deparaffinised in Xylene for 5mins. This was repeated with fresh xylene for 5mins. Slides were then placed in 50% Xylene and 50% Ethanol for 5mins. Slides were rehydrated in an ethanol series of 100%, 95%, 70% and 50% ethanol for 10mins each. Slides were washed in PBS for 10mins, then MS water for 10mins twice. The slides were then washed in the ammonium bicarbonate buffer. Using the PanCK staining the tumour areas were outlined using a needle. A fine tip PAP pen was used to go around these areas. 500µl CSase was diluted in 4.5ml ammonium bicarbonate buffer to reach a concentration of 0.5U/ml. 0.5U/ml CSase was added to the tumour regions of the slides and the slides were placed in a humidified box and incubated at 37°C overnight. CSase solution containing the extracted CS disaccharides was collected in 1.5ml eppendorfs. MS water was added to the slides if they had dried out and the water was added to the eppendorf. The sample was centrifuged at 14,000g for 10mins at RT to pellet any tissue remnants. The sample was then transferred to another eppendorf and the sample was stored at 4°C. 0.5U/ml CSase was added to the stromal areas of the slide and the slides were incubated overnight at 37°C in a humidified box. The CSase was collected as completed with the tumour regions. The samples from both tumour and stroma areas were lyophilised in a speed vac for 2hrs. MS water was added to bring the volume of samples to 30µl. Samples were stored at 4°C before analysis.

### 2.19.2 From cell culture

HCC38, MDA MB468, and HMF3S cells were grown in protein free media (PFM) as mono-culture, co-culture (HMF3S + tumour cells) and in conditioned media (HMF3S with 250µl tumour conditioned PFM) in 24 well plates. 500µl media was collected from each well after 7 days. 300µl of PFM from each condition was taken for CS

extraction. 20µg of crude and enriched protein from 2 rounds of VCAN enrichment with IEX. Samples were fixed with 4% PFA. Each sample was first dialysed using 0.5ml 10kDa centrifugal spin columns (Merck, UFC501024). The columns were spun at 14,000g for 10mins at RT. MS water was added to the sample and the sample was spun again. This was repeated twice. 5U/ml CSase was diluted to 0.5U/ml CSase in ammonium bicarbonate buffer. CSase was added to each sample to reach a final volume of 400µl. Samples were placed in eppendorfs and incubated on a shaker at 37°C overnight. The samples were placed in a 0.5ml 3kDa centrifugal spin column (Merck, UFC500324). The samples were spun at 14,000g for 10mins at RT. The eluted sample was collected and placed into an eppendorf to be lyophilised in a speed vac. The volume of the sample was made to 30µl with MS water and stored at 4°C.

The samples were sent to the University of Washington for analysis by LC-MS/MS<sup>212-214</sup> by Dr Kim Alonge.

## 2.20. Cell bacterial transformation

### 2.20.1 Plasmid expansion

Ampicillin was added to autoclaved LB agar to a final concentration of 100µg/ml. Agar was poured into petri dishes and left to set for 1hr. 1µg of plasmid was added to 40µl of DH5α competent cells (ThermoFisher, EC0112). A control was made with PUC19 plasmid and no plasmid. The bacteria were mixed by pipetting up and down and then incubated on ice for 30mins. The bacteria was vortexed and heat shocked at 42°C for 90 seconds. Samples were left on ice for 5mins. 200µl of SOC medium was added to the bacteria and was incubated for 1hr at 37°C on a shaker at 130rpm. The bacteria were spread across the agar plate using a inoculation loop and incubated at 37°C overnight.

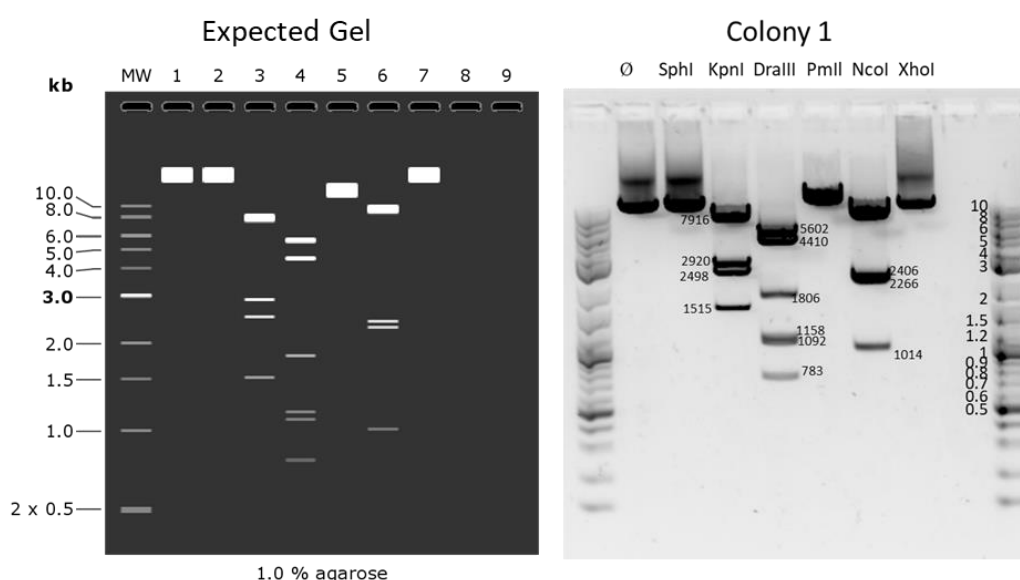
A single colony was selected from the agar plate and added to 50ml autoclaved LB broth with 100µg/ml ampicillin. The broth was left on a shaker at 37°C overnight. The broth was centrifuged at 6000g for 15mins at 4°C to pellet the bacteria. The DNA was extracted using the Maxi Prep kit (Qiagen, 12162) following manufacturer's protocol. The pellet was resuspended in 7ml of Buffer P1. Buffer P2 was added and sample was mixed. The samples were left at RT for 5mins. 7ml Buffer S3 was added to the sample. The sample was mixed and centrifuged for 1min at 4700rpm (rotor radius 8.6cm). The sample was left to stand for 5mins at RT.

The solution was then poured through the filter cartridge. 5ml of Buffer BB was added and the sample was vacuumed through the spin column. 0.5ml Buffer ETR was then vacuumed through the column followed by Buffer EB. The spin column was placed in a collection tube and centrifuged at 9,700rpm for 1min. The column was placed in a clean eppendorf and 400µl Buffer EB was added. The column was spun at 9,700rpm (rotor radius 8.6cm) for 1min to elute the DNA. The DNA concentration was measured using a nanodrop. The plasmid was stored at -20°C.

## 2.20.2 Enzyme restriction of plasmids

The V1 plasmid was purchased from Genscript (Appendix 3). The structure of the plasmid was confirmed using enzyme restriction. Enzymes: SphI (NEB, R0182S), KpnI (NEB, R3142S), DraIII (NEB, R3510S), PmlI (NEB, R0532S), NcoI (NEB, R0193S) and XhoI (NEB, R0146S) were used to digest the plasmid. A solution of 0.5µl enzyme, 0.5µg plasmid, 2.5µl Cutsmart buffer (NEB, B6004S) and 22µl RNase free water. Samples were incubated at 37°C with shaking.

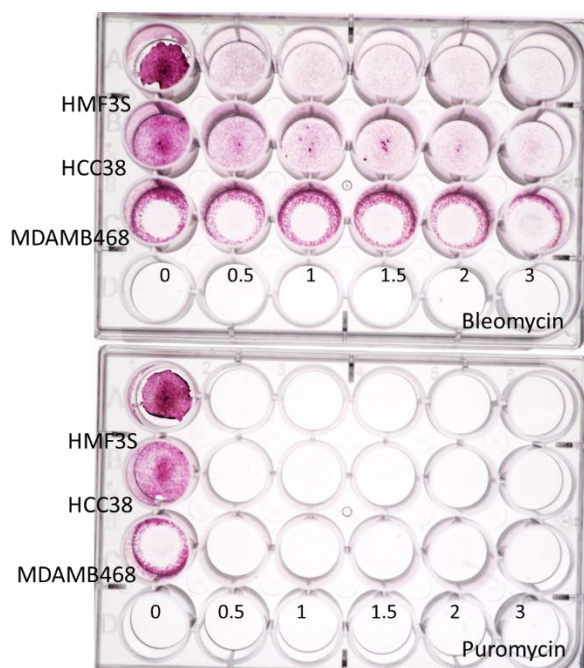
A 1% TAE (w/v) buffered agarose gel was made with gel red (Merck, SCT123). 5µl loading dye (NEB, B7025) was added to each sample. Samples with ladder (NEB, N3232S) were loaded onto the gel and run at 120V for 45mins. The gel was imaged under the chemidoc (Figure 2.4).



**Figure 2.4. Enzyme restriction of V1 plasmid.** Bands detected by agarose gel electrophoresis following enzyme digestion. Expected bands determined using Snapgene.

### 2.20.3 Antibiotic survival test

Cells were seeded onto a 24-well plate with 6-wells per cell line at 70% confluency and left to settle overnight. Antibiotic dilutions of Puromycin (Invivogen, ant-pr-1) and Bleomycin (ThermoScientific, J60727.MCR) were made at increasing concentrations and the wells were treated. The wells were treated each day for 3 days. The wells were then washed and stained with 0.25% crystal violet. The wells were washed with PBS and imaged (Figure 2.5).



**Figure 2.5. Drug dilution test of cells.** Crystal violet staining of wells following drug treatment. Drugs used at µg/ml.

### 2.20.4 Transient transfection

Cells were seeded onto a 24-well plate at 70% confluency and left to settle overnight. Lipofectamine 3000 (Invitrogen, L3000001) was used to transfect cell lines with the plasmids. Solution A was made with 25µl OptiMEM (Gibco, 31985062), 1µg DNA and 2µl P3000. Solution B was made with 25µl OptiMEM and 7.5µl lipofectamine. The solutions were mixed together and incubated for 15mins at RT. 50µl was then added to the well. The cells were incubated for 24hrs and then the media was changed. Cells were left to grow for 2 days before splitting into a T25 flask. Cells were then treated with antibiotic to select for the transfected cells. Cells were treated for 3 days and then grown in normal media to expand for use in assay.

## 2.21. Gene data analysis

RNAseq data was obtained from The Cancer Genome Atlas – Breast Invasive Carcinoma (TCGA BRCA) project. TNBC patients were identified from the analysis completed by Lehmann et al. In the study the TNBC patients within the study were sorted by TNBC subtype<sup>184</sup>. For the analysis 100 patients were selected across the 4 subtypes. The ID for the samples are shown in supplementary table 1. The gene data was extracted for matrix genes: *VCAN*, *FN1*, *COL11A1*, *COMP* and *CTSB*. The samples were analysed used CIBERSORTx to determine the immune cell subtypes. On GraphPad prism the CIBERSORTx data was correlated to the matrix genes. Matched normal and tumour datasets were also taken from TCGA BRCA and the matrix genes were compared. Dataset IDs shown in Appendix 5.

## 2.22. Statistical analysis

All the graphs and statistical tests were done on GraphPad Prism V9. All correlations were determined using Spearman Rank Correlation test. For comparing paired data amongst groups, RM One-way ANOVA was used. Unpaired data was calculated using One-way ANOVA or Two-way ANOVA with either Tukey's test or Welch correction depending on the group size and variations amongst samples. Data was statistically significant if the P value <0.05.

## Chapter 3: Characterising the tumour immune infiltrate and matrix expression in TNBC tissues.

### 3.1. Introduction

Cancer immunotherapies work for a limited number of individuals (approx. 10-20%), and therefore predictive markers that can help clinicians identify those most likely to respond would be advantageous. The tumour immune phenotype (TIP) is a classification in development that can be predictive of immunotherapy response<sup>41,49</sup>. TIP is a measure of both the number of immune cells in a tumour and where they are located in relation to the tumour epithelium. TIPs are characterised across three broad classes; desert, inflamed and excluded. Desert tumours have little to no immune cells, whereas, inflamed and excluded tumours both have high levels of immune cell infiltration but differ in the immune cell location with limited immune cells are in contact with the tumour epithelium in excluded tissues.

In this chapter, a set of 26 TNBC tissues have been spatially analysed for the localisation of CD8+ T cells and macrophages and the TIP of each tissues, has been determined. The TIP classification used here was adapted from the method outlined by Kather et al<sup>44</sup>. One of the key reasons immune cells are thought to be restricted to the stroma is the matrix composition. Following a previous analysis in our lab, five molecules were found to be upregulated at both gene and protein level in a signature that associated with immune suppression<sup>23</sup>. The expression of these molecules was explored at both gene and protein level to look at co-expression and co-localisation. VCAN and cathepsin B (CTSB) were found to be co-expressed. Further analysis comparing their expression to CD8+ T cell localisation identified VCAN as being associated with T cell localisation. The levels of VCAN expression and spatial location were compared between the different immune phenotypes to understand how VCAN expression could associate with immune cell trafficking.

#### 3.1.1. Chapter highlights at a glance:

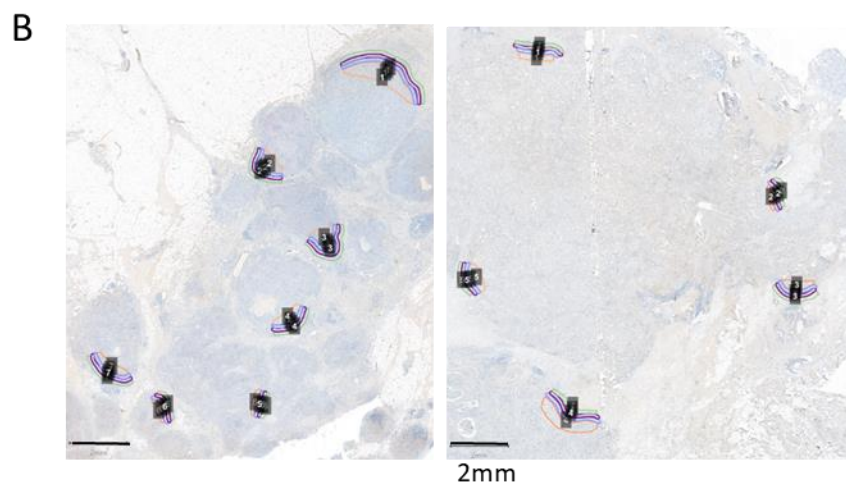
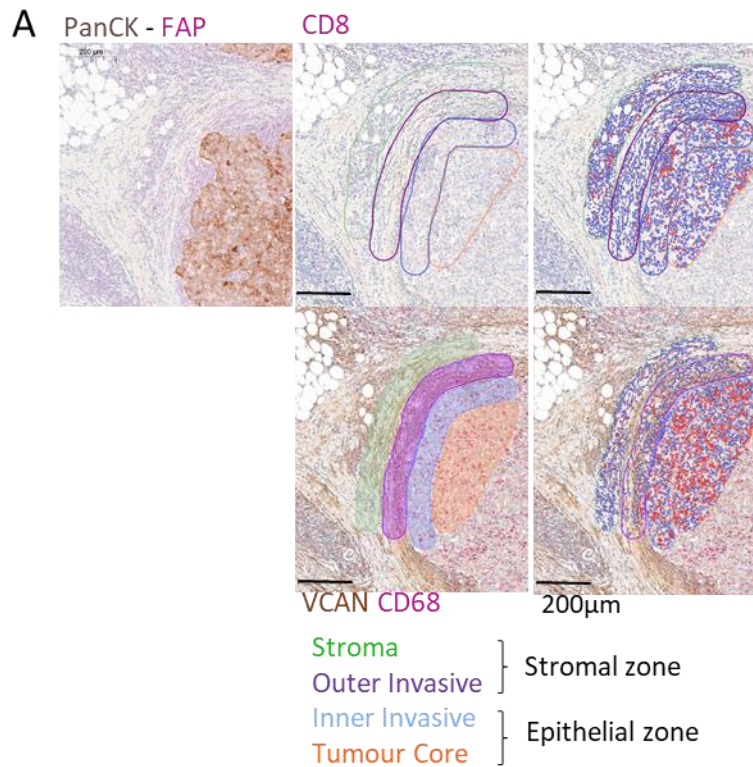
- Three TIPs observed in TNBC, with a majority being excluded
- VCAN expression is associated with CD8+ T cell localisation
- Immune exclusion is associated with VCAN expression around the tumour epithelium

### 3.2. Identifying TIPs in TNBC tissues

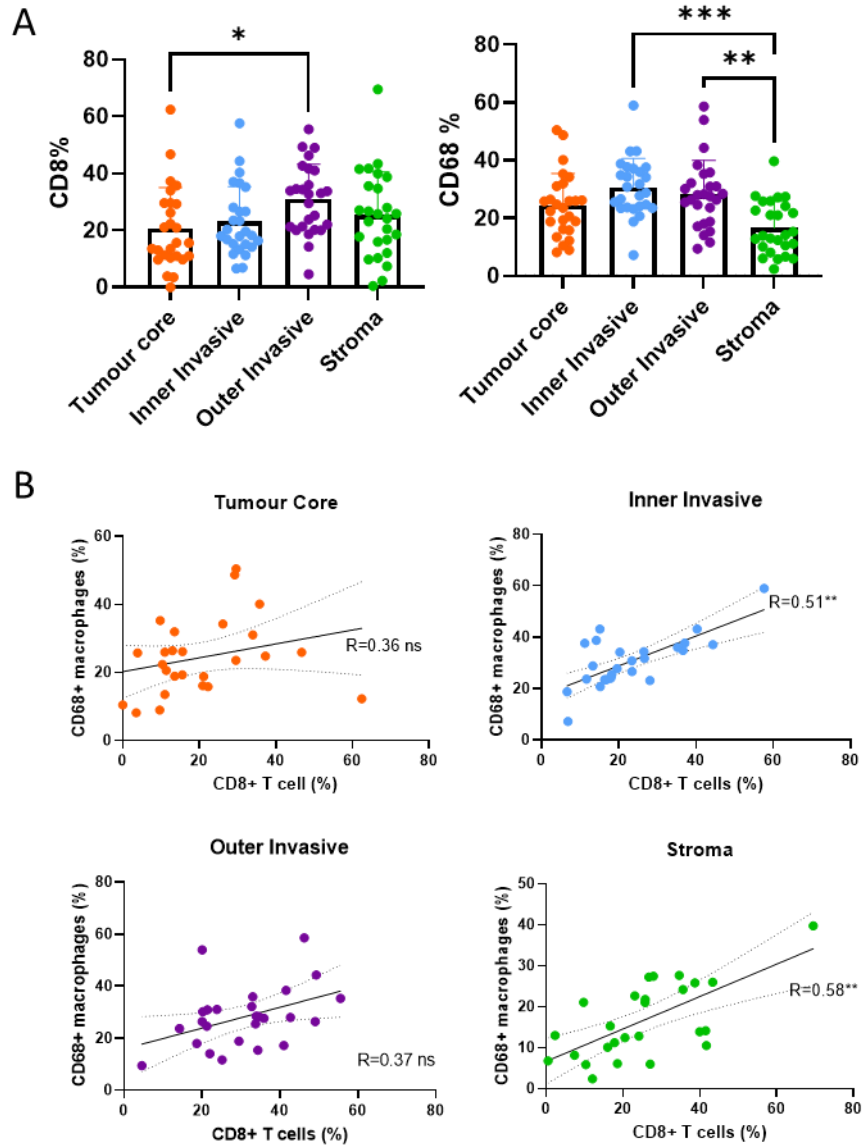
To characterise the TIP's across a library of 30 FFPE TNBC tissues, immunohistochemistry for CD8 (CD8+ T cells), CD68 (macrophages), PanCK, FAP and VCAN were used and analysed with the image analysis software QuPath <sup>207</sup>. From the staining only 26 tissues were suitable for analysis. The tissues were initially analysed following a method outlined by Kather et al <sup>44</sup>, which we found to be limited for the set of tissues used, and therefore we adapted this method, however both the original and adapted method are presented here. Starting with the original method <sup>44</sup> here, the areas around the tumour epithelium and stroma are separated into different regions and the number of immune cells within each region are counted. To define the regions, PanCK stain was used to confirm the tumour epithelium. We defined the tumour epithelium and stromal areas by PanCK and FAP staining respectively. To analyse the interactions between these areas, the inner and outer invasive areas were defined. The inner invasive area was established as a 100µm margin within the tumour epithelium and the outer invasive area as a 100µm margin within the stroma. Areas further into the tumour epithelium and stroma were also studied to form the tumour core and stroma. These regions were then grouped to form the epithelial zone (tumour core and inner invasive) and stromal zone (outer invasive and stroma) (Figure 3.1A). Within a tissue, multiple areas were analysed to be able to see the heterogeneity and the overall phenotype of the tissue. The number of areas sampled ranged from 2 to 7 and was dependent on the size and number of tumour islands present (Figure 3.1B).

The number of CD8+ T cells and macrophages were counted in each of these regions. CD8+ T cells were found to be significantly different between the tumour core and outer invasive area, whilst macrophages were found to be significantly lower in the stroma (Figure 3.2A). Positive correlations were observed between CD8+ T cells and CD68+ macrophages in the inner invasive and stromal regions (Figure 3.2B).



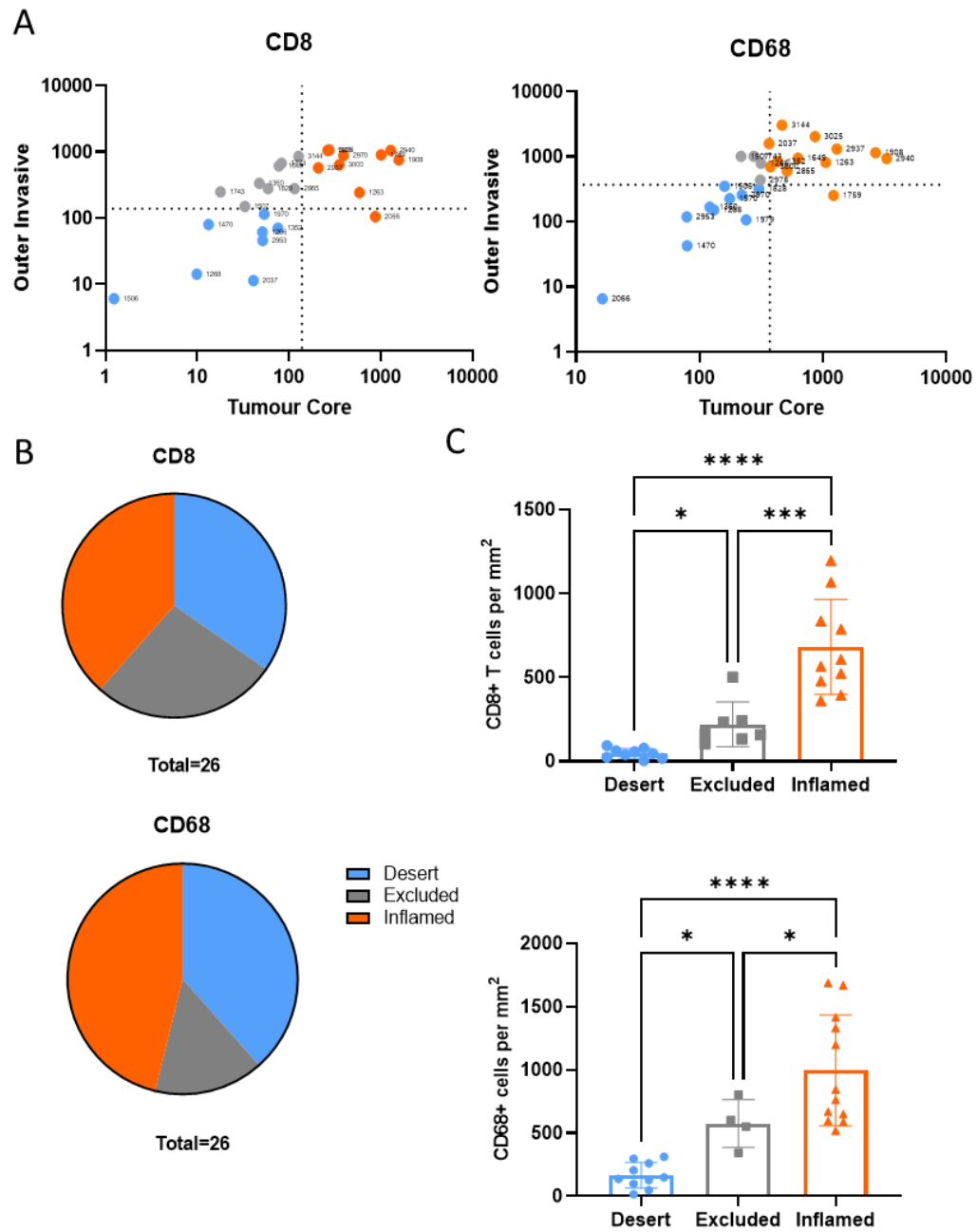


**Figure 3.1. Analysing the tumour epithelial border.** A) Regions around the border of the tumour epithelium were drawn using QuPath on 26 TNBC patient tissues. The outer and inner invasive margins were on either side of this border and were 100µm wide. The stroma and tumour core were then outlined. The stroma and outer invasive were grouped as the stromal zone. The tumour core and inner invasive were grouped as the epithelial zone. B) In a single tissue multiple areas were sampled. Areas ranged from 2-7 depending on the number and size of the tumour islands.



**Figure 3.2. Spread of immune cells across regions.** A) The percentage of CD8+ and CD68+ cells were calculated as the number of cells in each region against the total number of cells in all regions. The graphs show the spread of cells across the regions. Repeated measures one-way ANOVA. B) Correlations between the CD8+ T cells and CD68+ macrophages determined for each region. Spearman Rank Correlation. \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.005$ .

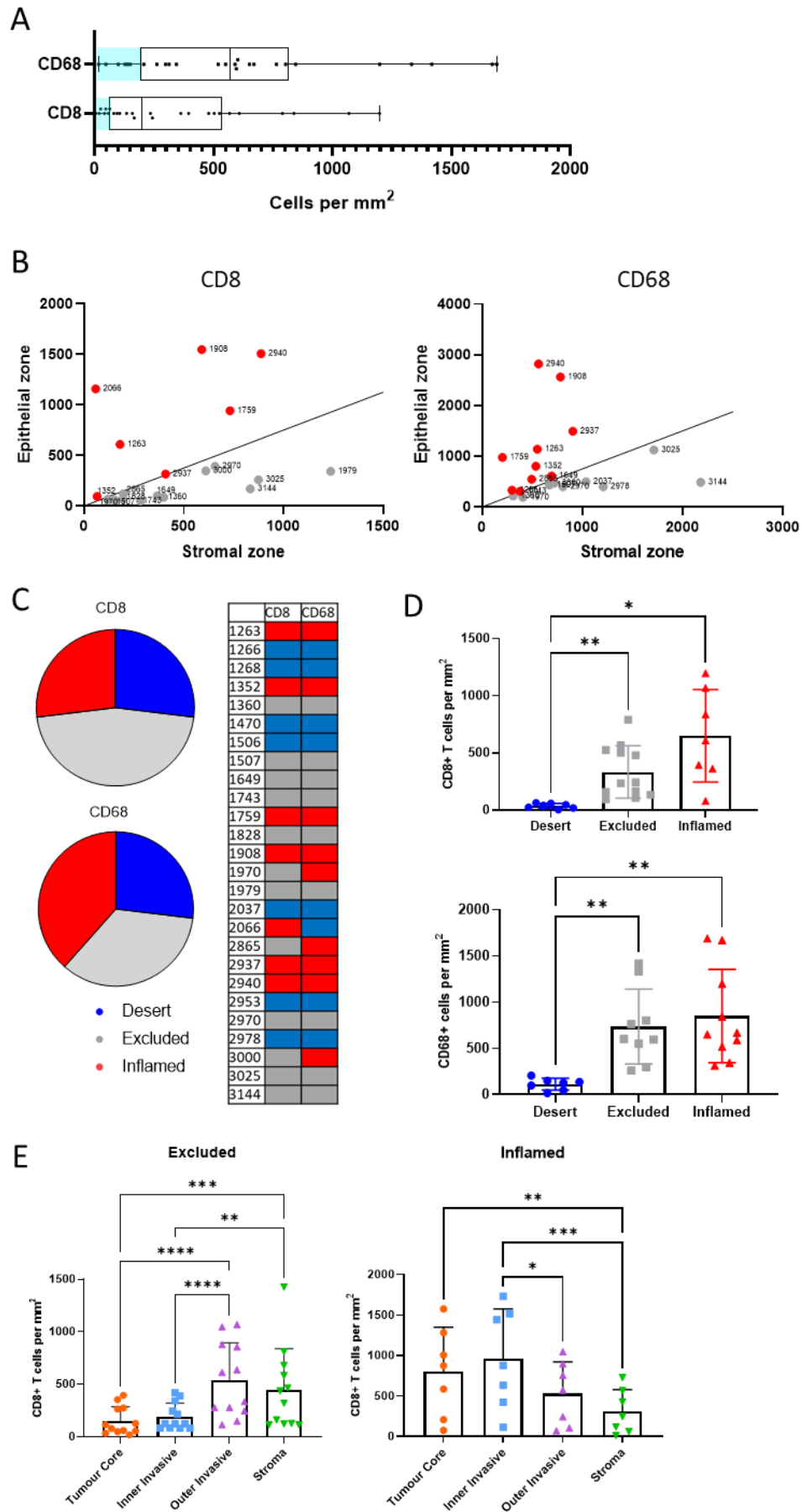
Using this data, the TIP was determined for each area by comparing the number of immune cells in each zone to the total immune cell count. This was repeated in multiple areas around the tissue and an average was made for each zone (Figure 3.1). The cut-off margins for the immune phenotypes was determined by the median of the total immune cell counts for all the tissues. When the number of CD8+/CD68+ cells in the epithelial area was greater than this median, the tissue was classed as inflamed (orange). If the cell count was below the cut-off in both zones then the tissue was classed as desert (blue). Tissues below the cut-off for the epithelial zone but above the cut-off for the stromal zone were classified as excluded (grey, Figure 3.3A). Overall, the analysis showed that a quarter of the tissues were excluded for CD8+ T cells and a majority of the tissues inflamed. For macrophages, an equal number of tissues were desert and inflamed and a small percentage had an excluded phenotype (Figure 3.3B). Moreover, comparison between the number of cells in the different phenotypes showed that there were significant differences between the total cell numbers in each phenotype (Figure 3.3C). This indicates that the tissues fell into the excluded category because of their low number of cells rather than cells being restricted to the stromal zone. Therefore, the cut-offs using this method were found to be misleading for TNBC tissues.



**Figure 3.3. Exclusion analysis using the Kather et al method.** A) Median of the number of immune cells in the tumour core and outer invasive taken as the margins. Cell count above this margin in the tumour core were defined as inflamed (orange). Tissues with a count below the margin in the tumour core but above in the outer invasive were excluded (grey) and those below in all regions were desert (blue). B) Pie chart showing the spread of the different phenotypes in the tissues. C) Immune cell counts in each phenotype. Mann-Whitney test to determine significant differences. Welch ANOVA. \* =  $P < 0.05$ , \*\*\* =  $P < 0.005$ , \*\*\*\* =  $P < 0.001$ .

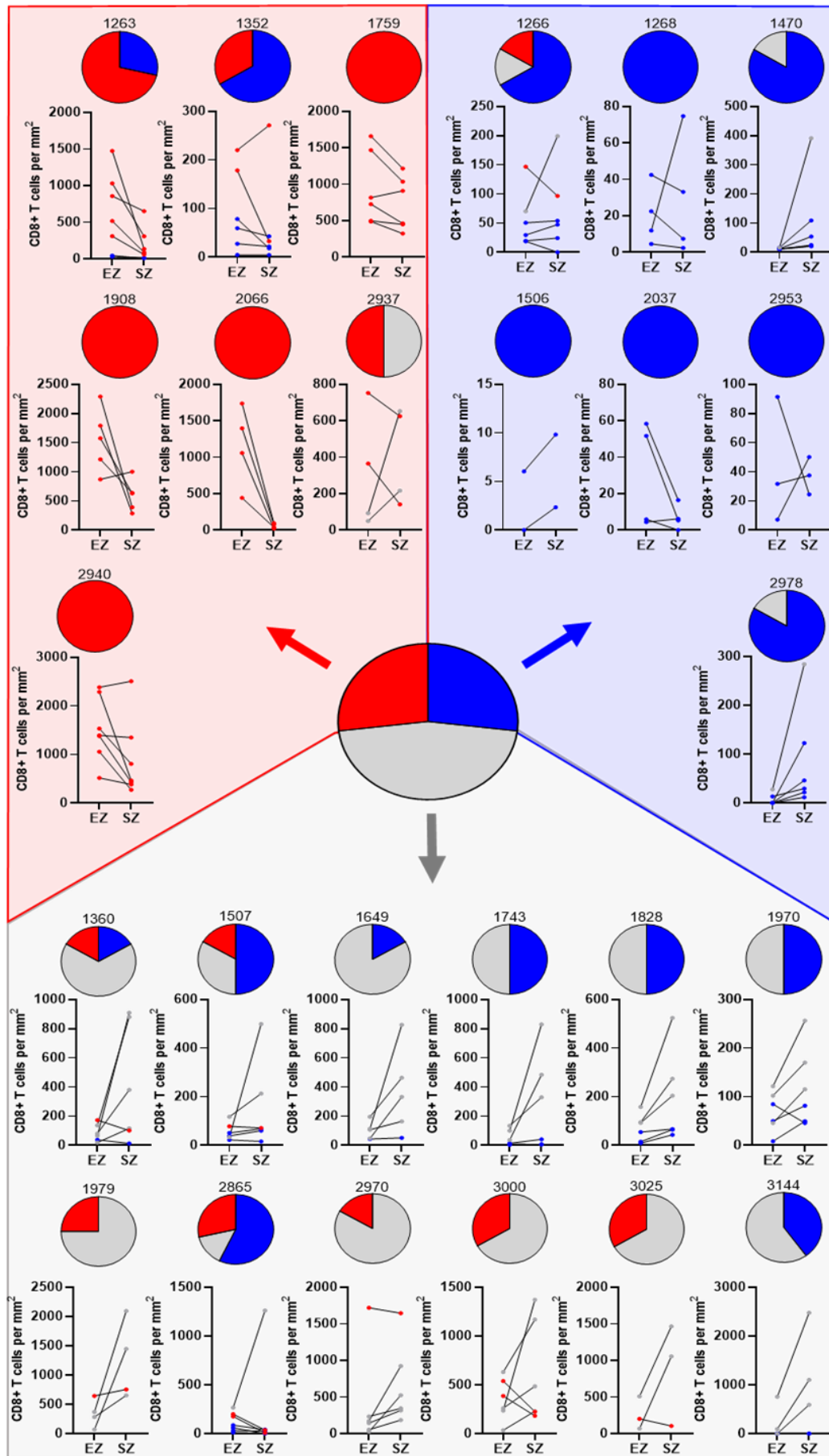
Excluded tissues are defined as tumours where the level of immune infiltration was comparable to inflamed tumours but the immune cells are restricted in the stroma. However, using the parameters from the Kather et al method did not seem to fulfil this definition and therefore the analysis method was adapted. First, desert tissues were identified by using the first quartile from the total cell count as the cut-off margin (Figure 3.4A). For the remaining tissues, the ratio of immune cells between the epithelial and stromal zone was calculated. A ratio greater than 0.75 classified tissues as inflamed (red) and below 0.75 as excluded (grey) (Figure 3.4B). Using a ratio to define the cut-off point allowed the tissues to be classified independently from the cohort. Within literature there has not been a consensus as to what this cut off should be. Within a study by Hammerl et al, this cut-off was 0.1<sup>49</sup>, whilst in a study by Li et al, 0.5 was taken as full infiltration<sup>215</sup> and ratio of 1 was used by Derks et al<sup>216</sup>.

Using the scoring method developed here (discussed in the previous paragraph), nearly 50% of tissues were found to be excluded and 25% were inflamed for CD8+ T cells. Interestingly, one tissue was found to be desert for macrophages despite being inflamed for CD8+ T cells (Figure 3.4C). A comparison of the total number of immune cells showed no significant difference between the overall cell counts in inflamed and excluded tissues indicating that the difference within excluded and inflamed phenotypes reside in the cell localisation (Figure 3.4D). Overall, the spread of cells across the four regions demonstrated that within excluded tissues there is a significantly higher number of cells in the outer invasive area compared to the tumour core whereas inflamed tissues showed the opposite, with more T cells within and close to the tumour core (Figure 3.4E).



**Figure 3.4. Exclusion analysis with adapted method.** A) Box plot showing the identification of desert tissues. B) Comparison of CD8 and CD68 in the epithelial zone and stromal zone. Line shows ratio of 0.75 where  $Y=0.75X$ . Tissues above the line were inflamed (red) and below were excluded (grey). C) Summary pie chart and table showing the spread of phenotypes across tissues. D) Number of CD8+ T cells and CD68+ macrophages in each phenotype. Welch ANOVA. E) Spread of immune cells across each region for each phenotype. RM-one way ANOVA. \* =  $P<0.05$ , \*\* =  $P<0.01$ , \*\*\* =  $P<0.005$ , \*\*\*\* =  $P<0.001$ .

The phenotypes determined from the analysis were the average phenotype of the tissue, however, we noticed that more than one phenotype could be present in a tissue. For example, using CD8+ T cells the TIP was calculated for the different areas in each tissue as an average across the three TIPs. The average TIP assigned to a tissue in general is also the TIP that has the highest frequency in that tissue. However, in a few tissues the overall TIP assigned is not the highest frequency TIP. This was seen in excluded tissues 1507 and 2865 as well as the inflamed tissue 1352. In the excluded tissues, one area within the tissues showed a large difference between the cells in the epithelial zone and the stromal zone causing the phenotype to be skewed towards excluded when the count was averaged. A similar case was also observed in 1352 where 2 areas with high cell counts lead to the TIP being assigned as inflamed (Figure 3.5). This data indicates the importance of looking at different areas within a tissue rather than analyse the tissue as a whole.





**Figure 3.5. Immune phenotypes within tissues.** Overall TIP of the tissues shown in the centre. The different tissues within these TIPs are shown and the phenotypes within multiple areas of these tissues. Pie chart shows the ratio of each phenotype in the tissues – Inflamed (Red), Excluded (Grey) and Desert (Blue). Graph shows the number of CD8+ T cells in the epithelial zone (EZ) and stromal zone (SZ) for each area.

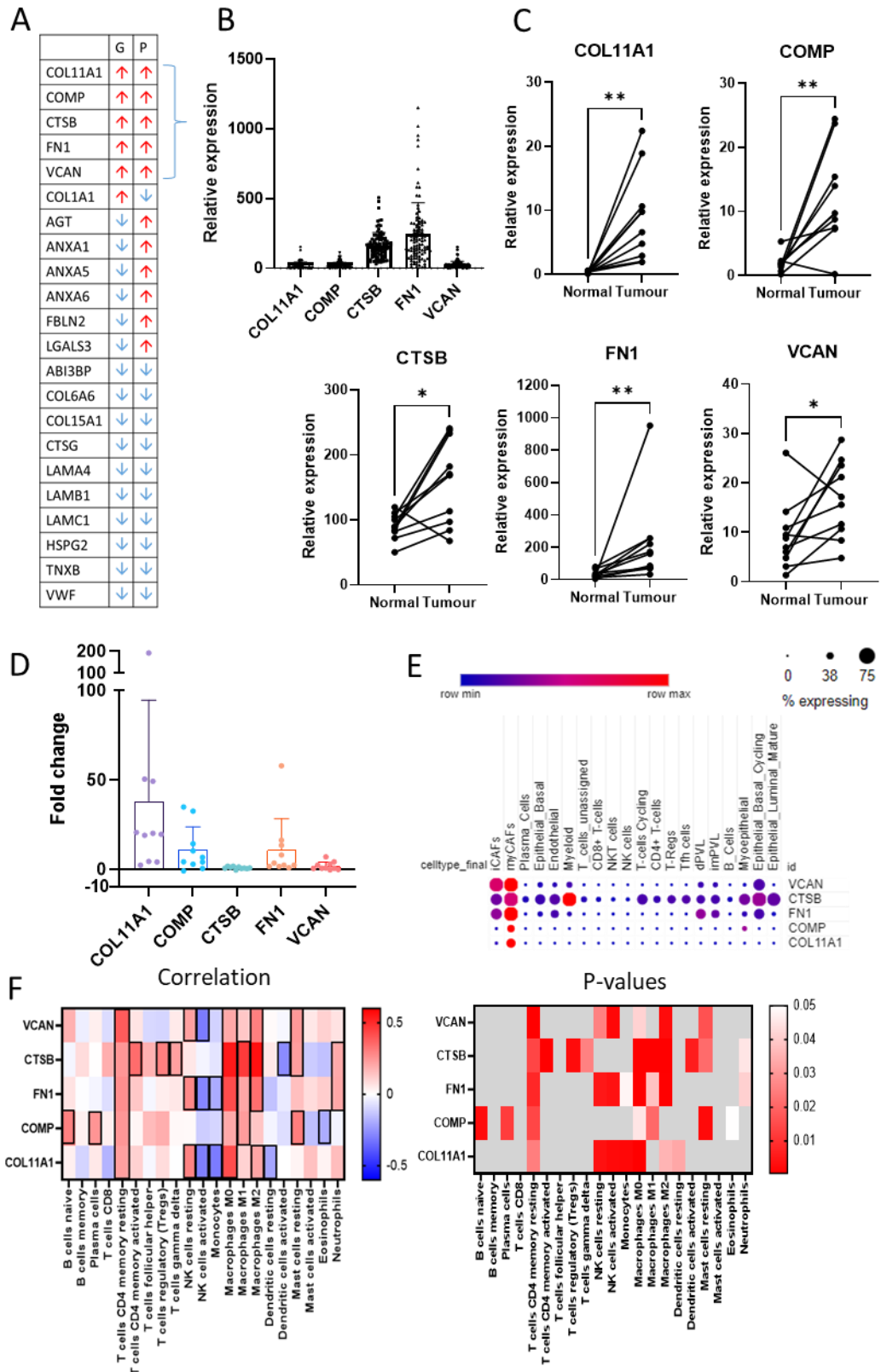
### 3.3. Expression of matrix index molecules in TNBC

Having assigned TIP values to each tissue, I next tested whether some particular matrix molecules associate with immune exclusion, focusing on the expression of the top 5 upregulated matrix molecules (FN1, COL11A1, VCAN, CTSB and COMP) from a signature previously identified in our group that associated with poor prognosis and immune suppression (Figure 3.6A) <sup>23</sup>. Initially, the expression of these molecules was explored at RNA level using TCGA datasets from 100 TNBC patients. RNA sequencing datasets were selected based on an analysis by Kalecky et al to identify TNBC patients from BRCA TCGA data sets and also the different subtypes of TNBC <sup>217</sup>. Out of the five matrix index genes, *CTSB* and *FN1* were the highest expressed (Figure 3.6B). Ten of these TNBC data sets also had matching data sets for normal and diseased tissue. Comparison between the normal and diseased tissues showed a significantly higher expression of these five genes in the tumour tissues (Figure 3.6C). From the overall fold change analysis, *COL11A1* had the highest increase in expression as well as the most variation. Despite *CTSB* being shown to have a higher expression (Figure 3.6B), the upregulation was the lowest (Figure 3.6D).

Using publically available single cell RNA sequencing datasets <sup>218</sup> and the single cell portal for analysis. The cells expressing these genes were analysed. Myofibroblasts (myCAFs) expressed all 5 genes while inflammatory (iCAFs) only expressed *VCAN*, *CTSB* and *FN1*. Basal epithelial cells also expressed *VCAN*, *CTSB* and *FN1* at low levels. The expression of *CTSB* was found in almost all cell types with the highest expression in myeloid cells while *COL11A1* was only expressed by myCAFs. Finally, *COMP* was expressed by a low number of myCAFs and myoepithelial cells (Figure 3.6E).

The relation between these genes expression and immune cell populations was explored using CIBERSORT<sup>219</sup>. CIBERSORT is an online tool and R package which can be used to deconvolute 22 immune cell populations from bulk RNA-sequencing data. The signatures for determining the immune cell populations were determined after the purification of immune cell subtypes and subsequent RNA profiling. The 100 patient samples from the TCGA were analysed using CIBERSORT. Spearman

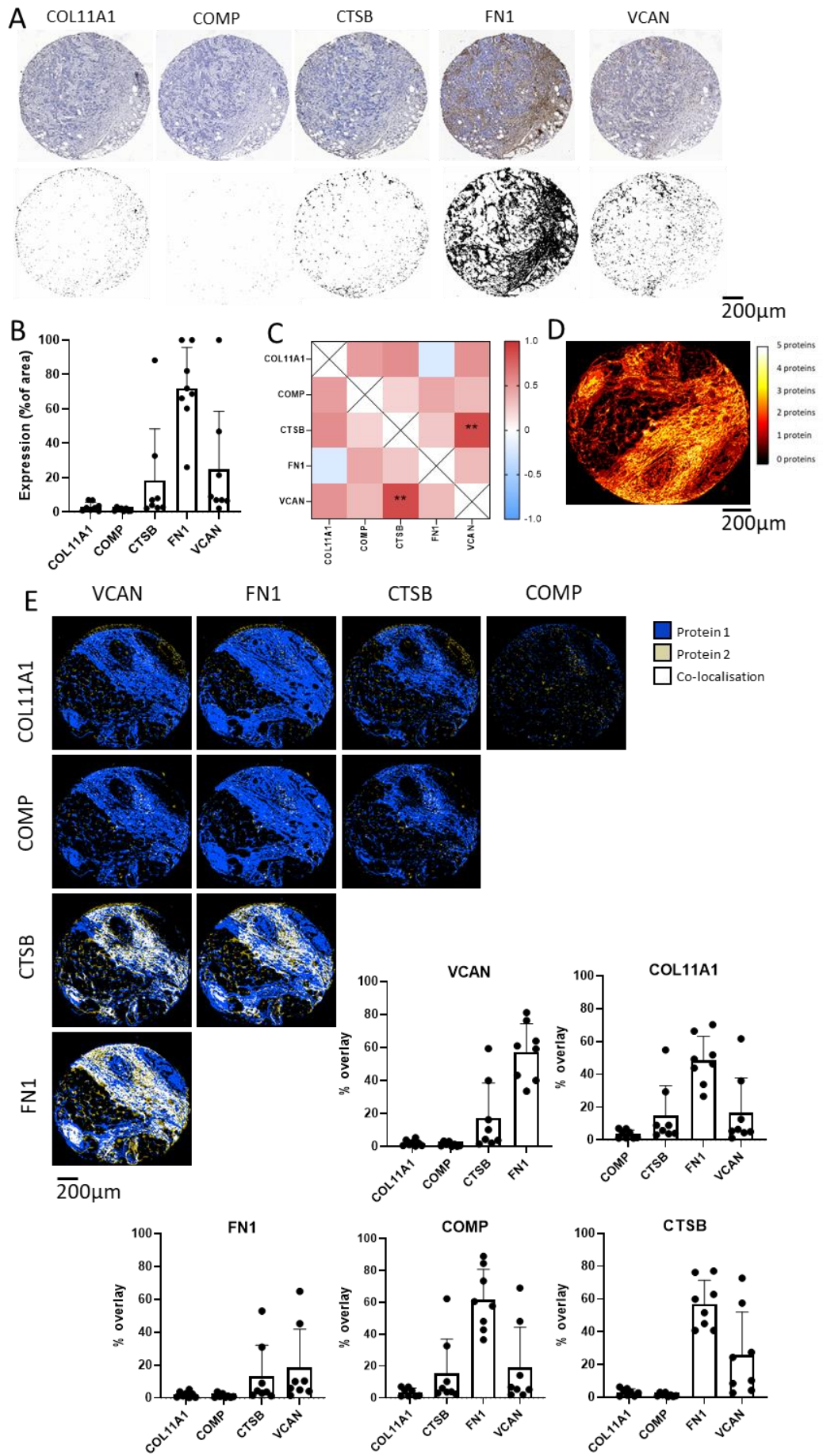
Rank correlation was completed to test how the expression of the five genes related to the different immune cell subtypes. CD4 memory T cells and M0 macrophages had a positive correlation with *VCAN*, *FN1*, *COL11A1*, *COMP* and *CTSB*. Overall, macrophages showed a positive correlation with the above mentioned genes, whilst monocytes negatively correlate, demonstrating that expression of these genes is possibly associated with monocyte differentiation to macrophages. *VCAN*, *FN1* and *COL11A1* expression favoured the resting state of NK cells in comparison to activated NK cells which showed a negative correlation. *COMP* was the only gene to show significant correlations with B cells, while *CTSB* significantly positively correlated with most of the T cell subsets (Figure 3.6F).



**Figure 3.6. Expression of overexpressed matrix index markers in TNBC.** A) The matrix index showing the upregulation or downregulation of genes and proteins from RNA sequencing and proteomic analysis. G=Gene, P=Protein. B) The relative gene expression of the 5 upregulated molecules from the RNA sequencing analysis of 100 TNBC patient tissues from TCGA datasets. C) Comparison of RNA expression in matched normal and tumour tissue from 10 TNBC patients from TCGA. Wilcoxon T test. \* =  $P < 0.05$ , \*\* =  $P < 0.01$ . D) Fold change determined by a ratio of the expression in normal and tumour tissue from the 10 patients. E) Single cell RNA sequencing data from publically available dataset of TNBC. Level of expression and number of cells expressed shown in dot plot. Graph formed using the single cell portal. F) Correlation of MI molecules with immune cell subsets. Immune cell subset levels determined using CIBERSORT with the TCGA dataset. P values also shown. Significant correlations indicated with a box.

The protein expression of these five markers was then investigated using IHC on human patient TNBC TMAs. Consecutive slides were stained for the five markers. Eight TMA cores were chosen for analysis as the remaining were not suitable for analysis on all 5 slides. To compare the level of staining, the positive DAB stain was deconvoluted using image J (Figure 3.7A). The percentage of protein was determined by the area of DAB staining from the deconvoluted images. FN1 was found to be highly expressed within the stroma of the tissues. VCAN and CTSB showed similar levels of expression (Figure 3.7B). Spearman Rank correlation showed that only VCAN and CTSB were significantly correlated (Figure 3.7C).

The overlaying of all the stains was then completed using the deconvoluted DAB stains. The overlay of all 5 proteins showed that there were regions of co-localisation of up to 4 proteins and some areas where only a couple proteins overlapped (Figure 3.7D). Co-localisation analysis was then completed on each pair of proteins. The first image was pseudo-coloured as blue and the second image as cream. The overlay of these led to white regions which were areas of co-localisation. The percentage area of the blue, cream and white areas were calculated and shown in Figure 3.7E. The graphs show how much of the protein staining overlaps with the other four proteins. Overall, FN1 showed the highest level of co-localisation with all proteins (around 60%), however its high expression across the whole core meant this co-localisation could not be considered specific. VCAN and CTSB showed similar levels of co-localisation with each protein. Despite this, the level of co-localisation between these proteins was not strikingly high (Figure 3.7E). One limitation of overlaying slides across multiple sections is that the pixels do not accurately overlay, This can lead to areas of co-localisation not being detected.

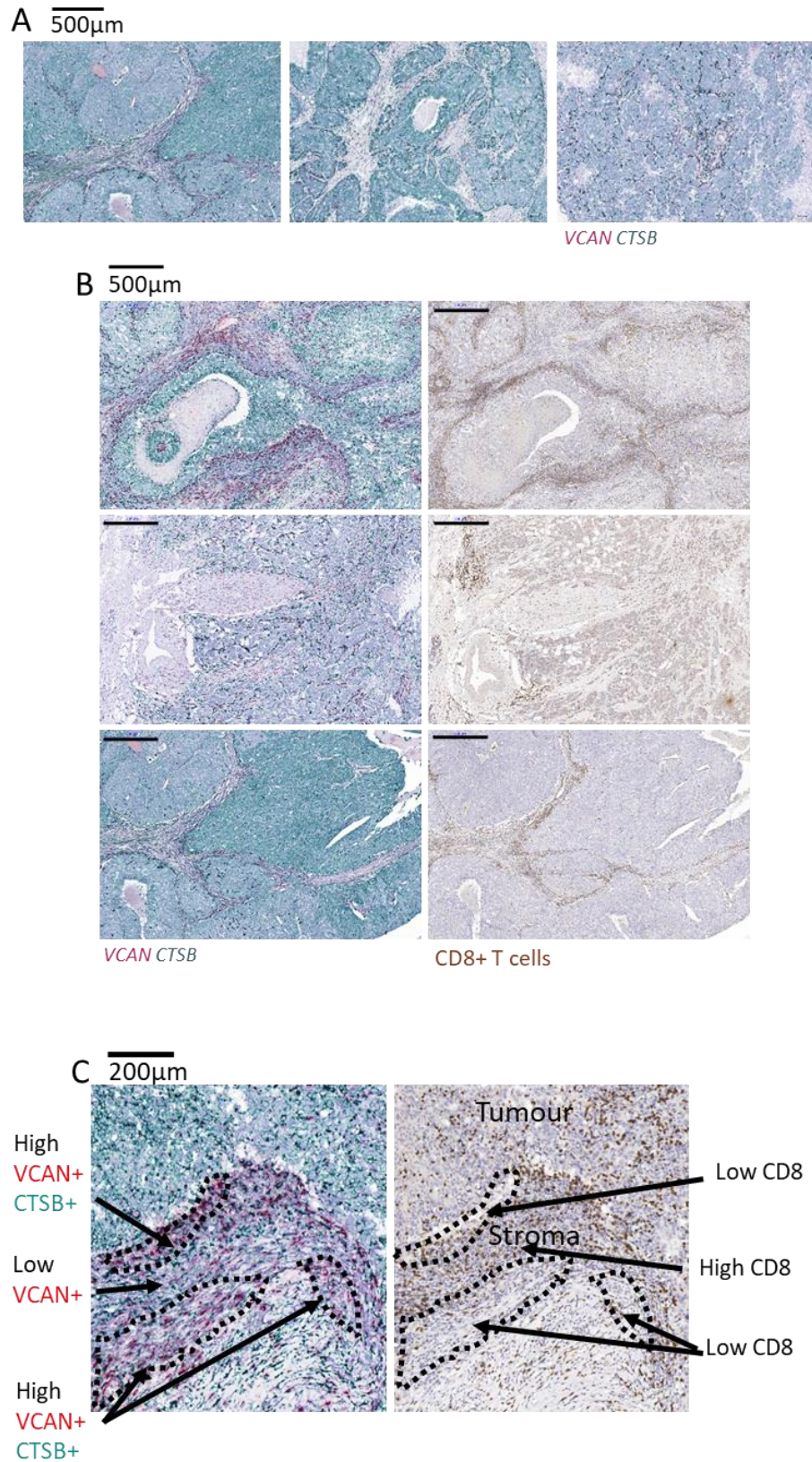


**Figure 3.7. Co-localisation analysis of matrix index proteins.** A) IHC completed on TNBC TMAs for the 5 proteins. Top images show stained TMA cores. Bottom images show DAB stains isolated using the colour deconvolution plugin on FIJI. B) Percentage of core stained for each marker from 7 cores on consecutive slides. Percentage determined using FIJI. C) Spearman rank correlation used to form a correlation matrix on the levels of each protein. D) Deconvoluted images were overlaid on FIJI to show areas of high levels of co-localisation. E) Pairs of stains overlaid to determine co-localisation between 2 proteins. Images overlaid on FIJI to produce an image with 3 colours – Blue for protein 1, Cream for protein 2 and white for overlay. Percentage overlay determined by number of white pixels (number of blue (or cream) pixels + number of white pixels).

Even though the level of co-localisation was not that high, VCAN and CTSS displayed similarities in terms of level of expression and visually looked to be present in similar areas. To look further into this, dual RNAscope was completed on FFPE sections from 10 TNBC patients. From the initial viewing of the stain it was clear that CTSS (green) was expressed by nearly all cell types in the TME (Figure 3.8A), recapitulating what was seen in the single cell RNA data (Figure 3.6E). VCAN however (red) was predominantly expressed in the stroma. Other than CAFs showing dual expression, no associations could be made between the RNA expression of VCAN and CTSS spatially in tissues.

To explore the immune landscape on the tissues, CD8 was stained on a consecutive slide. Comparisons of the CD8+ localisation and VCAN-CTSS RNAscope staining showed that CD8+ T cells localised more in areas where VCAN is expressed (Figure 3.8B). A deeper look into the localisation around the tumour epithelium border showed CD8 to be excluded from the tumour islands where VCAN expression was higher at the edge of the border (Figure 3.8C). Since T cells appeared to be unable to cross into areas where VCAN was expressed, I hypothesised that VCAN expression around the tumour epithelium border would be associated with CD8+ T cell exclusion, which was explored in the next section.





**Figure 3.8. Dual RNAscope of CTSC and VCAN.** 10 TNBC patient tissue sections were stained for VCAN and CTSC expression using dual RNAscope. A) Images of slides showing high CTSC (green) and VCAN (red) staining. B) CD8 DAB staining was completed on consecutive slides and compared to RNAscope staining. C) VCAN and CTSC expression was compared to CD8+ T cell expression around the edge of tumours.

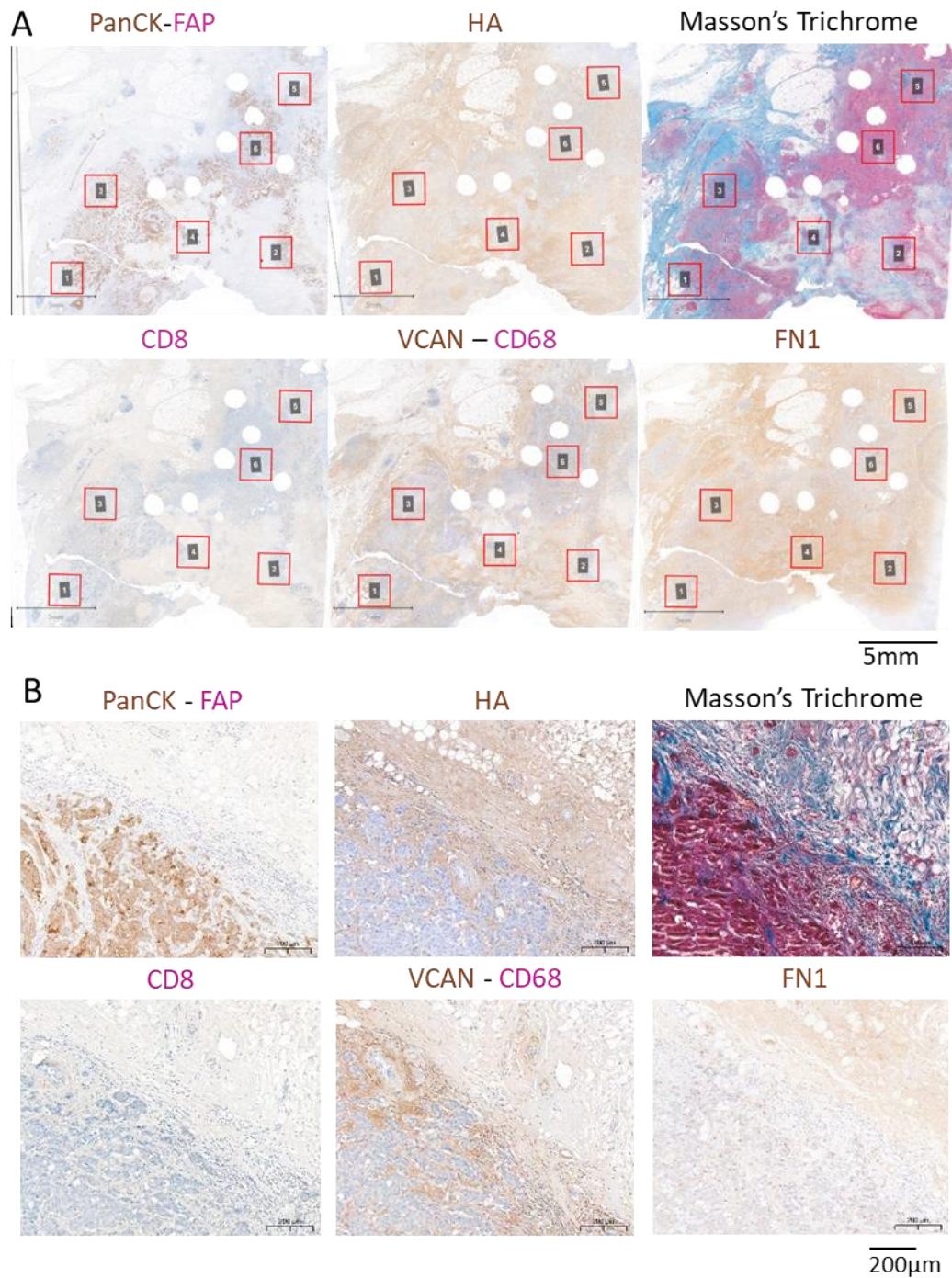
### 3.4. Expression of VCAN and its associated matrix molecules in comparison to CD8+ and CD68+ immune cells.

After having identified VCAN as a possible marker for immune exclusion, we wanted to study how the levels of VCAN, and its associated matrix molecules correlate with immune cell localisation. It has been shown that VCAN binds to three major matrix molecules named HA, FN1 and collagen<sup>92</sup>. These markers are also known to be highly expressed within the ECM, therefore it was important to look at whether any associations seen with VCAN are independent of these other molecules.

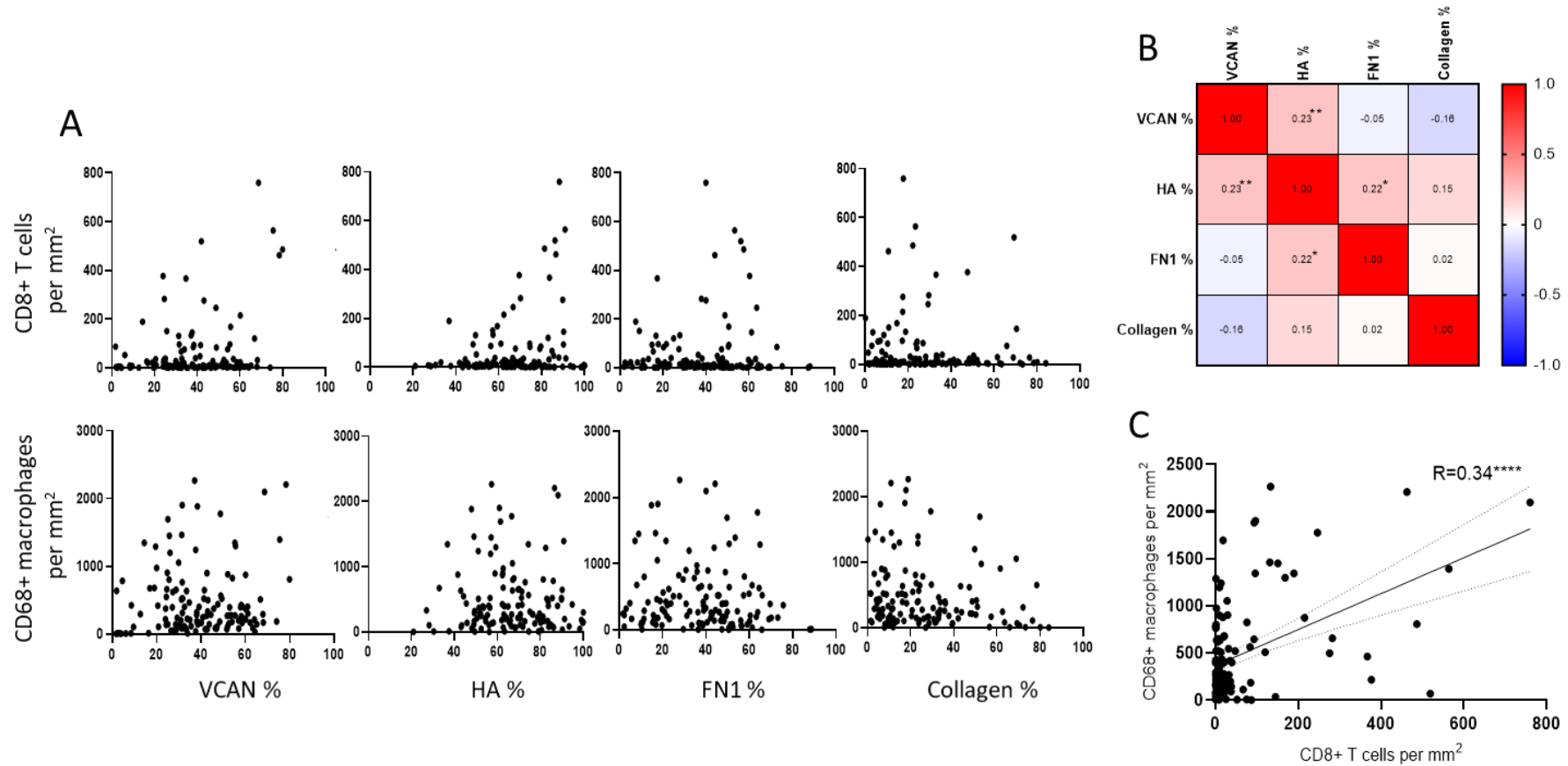
Using the library of 26 TNBC tissues from human patients, I analysed HA, PanCK, FAP, FN1, VCAN, CD8 and CD68 using IHC. Massons trichrome was used to detect collagen. The abundance of matrix markers and the cell counts were analysed on QuPath. Due to the size of the tissue and the variations in the level of adipocytes around the edge of the tissue, it was difficult to decide the margins of the tissue for whole tissue analysis. For this reason, 2mm-by-2mm areas around the tissue were randomly sampled to be able to detect tissue heterogeneity and patterns between markers. The number of areas sampled within each tissue ranged between 2 and 7 and on the basis of the number of tumour islands. Areas were selected from the PanCK-FAP stained sections to limit bias towards the expression of ECM and immune cell markers. The same areas were then analysed on other sections. To be able to accurately place the areas, images were aligned on QuPath and regions of interest copied on every image (Figure 3.9A). Because the main focus of the analysis was the immune infiltration into the tumour, regions of interest focused on areas which contained both tumour epithelium and stroma (Figure 3.9B).

The abundance of each matrix marker was then quantified as percentage area and the number of CD8+ and CD68+ immune cells were quantified as cells per mm<sup>2</sup>. To understand whether the matrix levels associated with macrophage or CD8+ T cell recruitment to the tissue, the cell count was correlated to the percentage of matrix. No correlation was observed between any matrix and immune cell markers (Figure 3.10A). Comparisons of the matrix markers showed positive correlations (VCAN and HA; HA and FN1 (Figure 3.10B)). CD8+ T cells and CD68+ macrophages were also positively correlated. (Figure 3.10C).





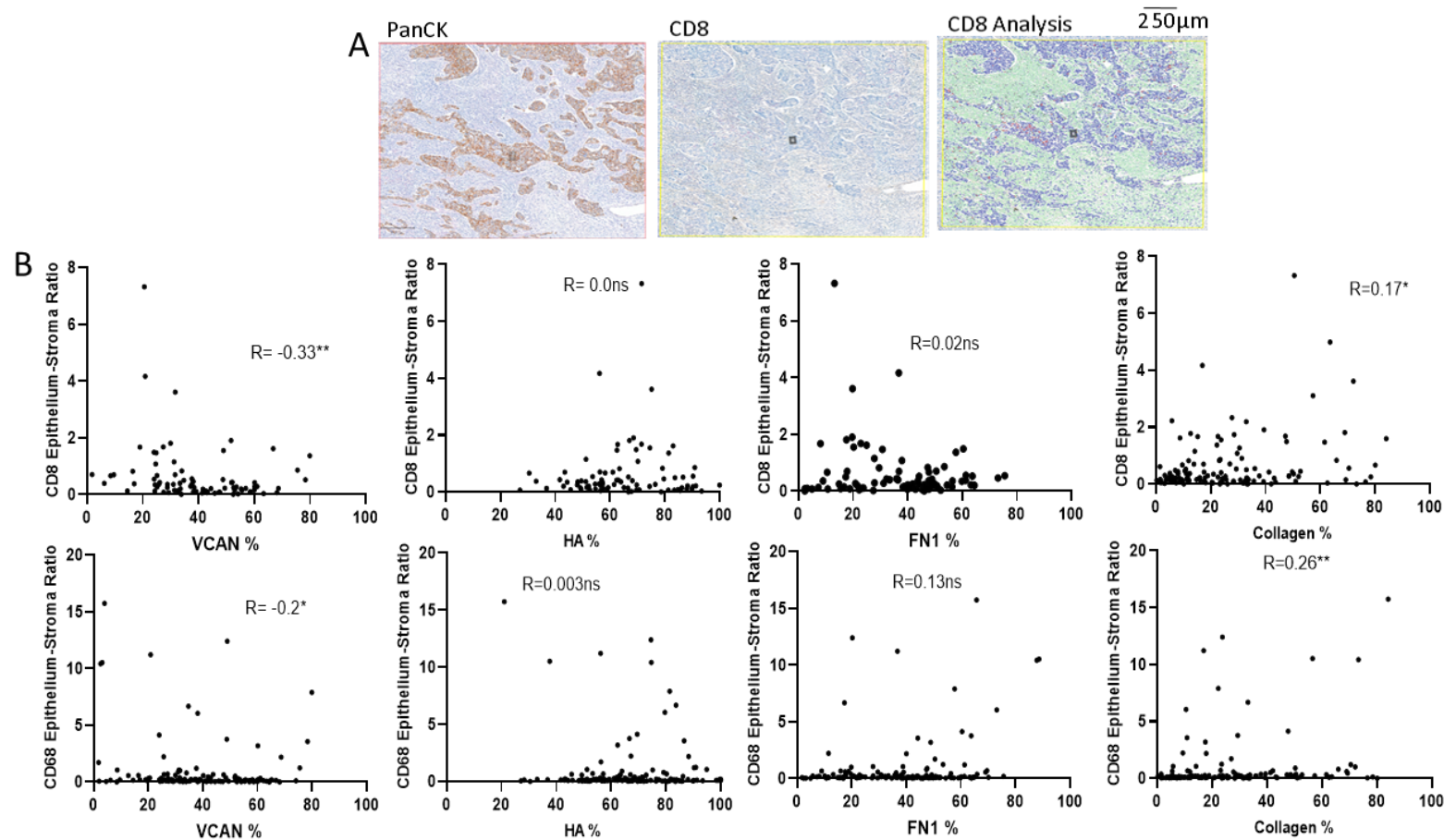
**Figure 3.9. Phenotyping tissues for matrix and immune cells.** A) Single or dual IHC for PanCK-FAP, HA, CD8, VCAN-CD68 and Massons trichrome carried out on consecutive sections of 30 TNBC patient tissues. 4mm<sup>2</sup> areas sampled around tissue at the tumour invasive border for each stain. B) Single area of tissue stained for the different markers.



**Figure 3.10. Comparison of matrix markers and CD8+ T cells and CD68+ macrophages.** A) 26 TNBC tissues stained for immune cell markers CD8 and CD68, and matrix markers VCAN, HA, FN1 and collagen (massons trichrome). QuPath used to determine number of cells and percentage area of matrix marker expression. Number of cells in each area compared to matrix marker. B) Spearman Rank correlation of the different matrix markers. C) Spearman rank correlation of levels of CD8+ T cells and CD68+ macrophages in each area. \* =  $P < 0.05$ , \*\* =  $P < 0.01$ , \*\*\*\* =  $P < 0.001$ .

I then studied the level of immune infiltration from the stroma to the tumour epithelium. To achieve this, QuPath was trained to identify which areas were tumour (using PanCK as a tumour marker) and which were stroma. This setup was then applied to the tissues stained for CD8 and CD68. This led to cells which were identified as being within the stroma labelled green and cells within the tumour labelled blue. Cells which were positive for CD8 and CD68 were stained dark green if they were present in the stroma regions and red if they were present in the tumour regions (Figure 3.11A).

To identify the presence of any correlation between the amount of immune cells in the different compartments and the different matrix molecules, we calculated a ratio between the positive immune (CD8<sup>+</sup> or CD68<sup>+</sup>) cells in the tumour epithelium and stroma and we compared this to the accumulation of each different matrix markers within the tissue. A high ratio indicated more cells are present in the epithelium compared to the stroma. A significant negative correlation was observed for VCAN and CD8<sup>+</sup> T cells ( $R=-0.33^{**}$ ) as well as for CD68<sup>+</sup> macrophages ( $R=-0.2^{*}$ ). No correlations were observed for HA and FN1. Moreover, a significant positive correlation was observed between collagen and the immune cells (CD8<sup>+</sup>  $R=0.17^{*}$ , CD68<sup>+</sup>  $R=0.26^{**}$ ). Although these correlations were not strong, I could hypothesise that the association between VCAN and CD8<sup>+</sup> T cell migration is independent of the expression of VCANs binding partners (Figure 3.11B).



**Figure 3.11. Tumour epithelium-stroma ratio of immune cells and the matrix.** A) QuPath was trained to identify cells as being in the tumour or in the stroma. PanCK IHC was used to guide the selection of areas. CD8 T cells and CD68+ macrophages could be identified as being either in the epithelium or stroma. B) The ratio of CD8+ T cells and CD68+ macrophages were compared in the epithelium and stroma to give a ratio. This was correlated to the matrix and two-tailed Spearman Rank Correlation was completed. \* =  $P < 0.05$ , \*\* =  $P < 0.01$ .

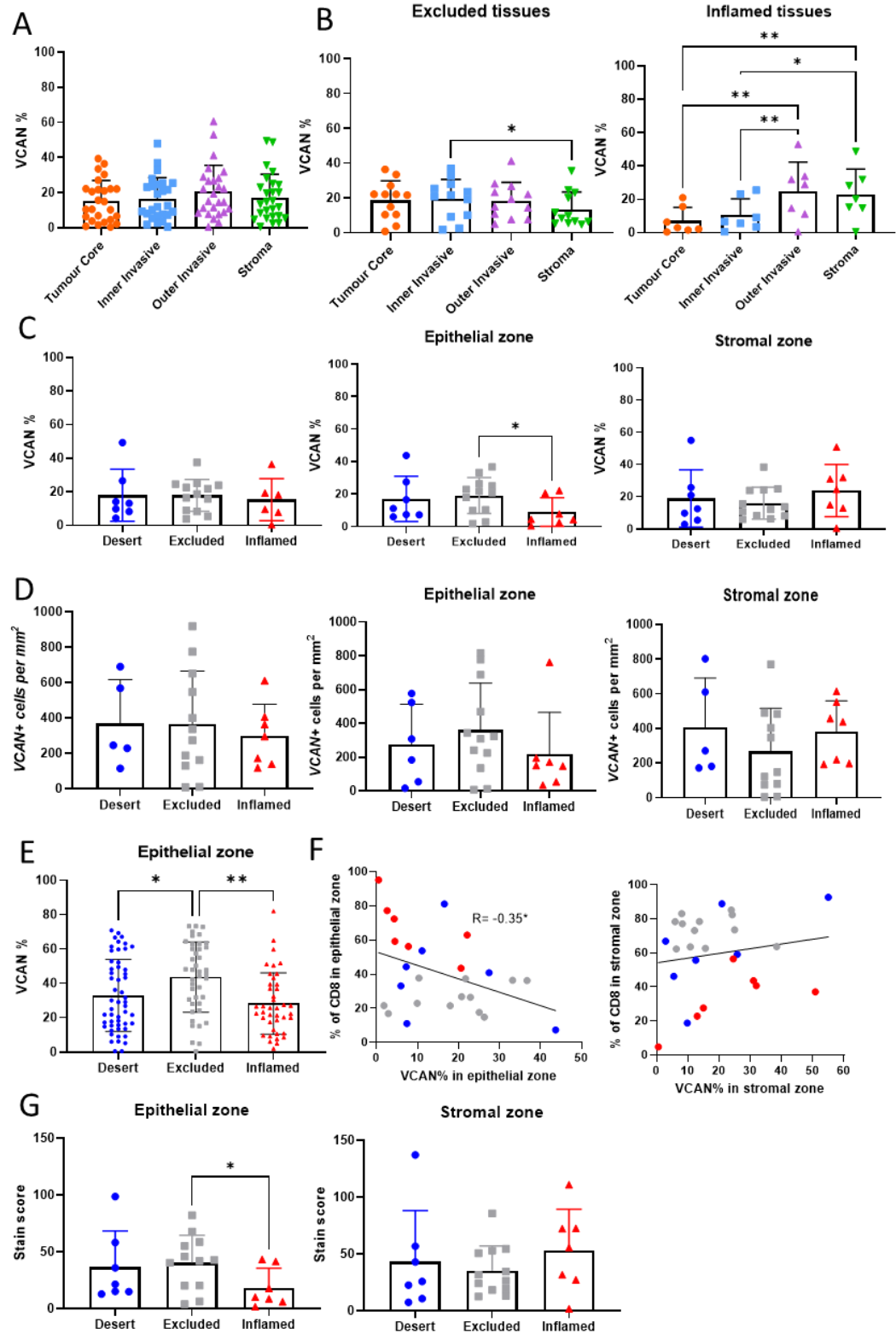
### 3.5. Immune exclusion associates with VCAN expression in the epithelial zone

Having identified VCAN as a protein of interest, the accumulation and expression was compared to the TIPs previously assigned. VCAN was quantified as a percentage in the same areas and regions as the immune phenotyping using the VCAN-CD68 dual stained sections (Figure 3.12A). The analysis was completed on the immune phenotype observed for CD8+ T cells. Overall, VCAN seemed to be equally spread in all the tissue regions analysed. When segregating tissues as excluded and inflamed, differences started to be observed. In excluded tissues, there was an even spread across regions whilst in the inflamed tissues there was a significantly reduced amount of VCAN in the epithelial zone compared to the stromal zone (Figure 3.12B). The overall VCAN present in a tissue section was determined as the average percentage for all 4 regions. No association was seen between the average percentage of VCAN and the immune phenotype (Figure 3.12C). When VCAN levels were averaged in the epithelial zone (tumour core + inner invasive) and the stromal zone (outer invasive + stroma), a significant increase in the level of VCAN in the epithelial zone within the excluded tissues was observed (Figure 3.12C). The same finding was also validated at RNA level using RNAscope analysis of the same tissue areas (Figure 3.12D). Analysis based on the phenotypes within the tissues also showed a similar pattern with the levels of VCAN in the epithelial zone of the excluded tissues being significantly higher (Figure 3.12E).

After comparing the percentage of CD8+ T cells in the epithelial and stromal zone to VCAN expression in the same areas, I observed a significant negative correlation ( $R=-0.4$ ) in the epithelial zone (Figure 3.12F), emphasising the idea that high abundance of VCAN in the epithelial zone is associated with immune exclusion.

When calculating the percentage area, the intensity of the staining isn't taken into account. Areas of high intensity staining can be attributed to high levels of protein localised in that area. To identify whether the differences in the percentage area of VCAN in the epithelial zone are due to the VCAN localising in specific areas, the areas of low, medium and high intensity were determined and a stain score was calculated. VCAN in the inflamed tissues was still significantly lower than excluded

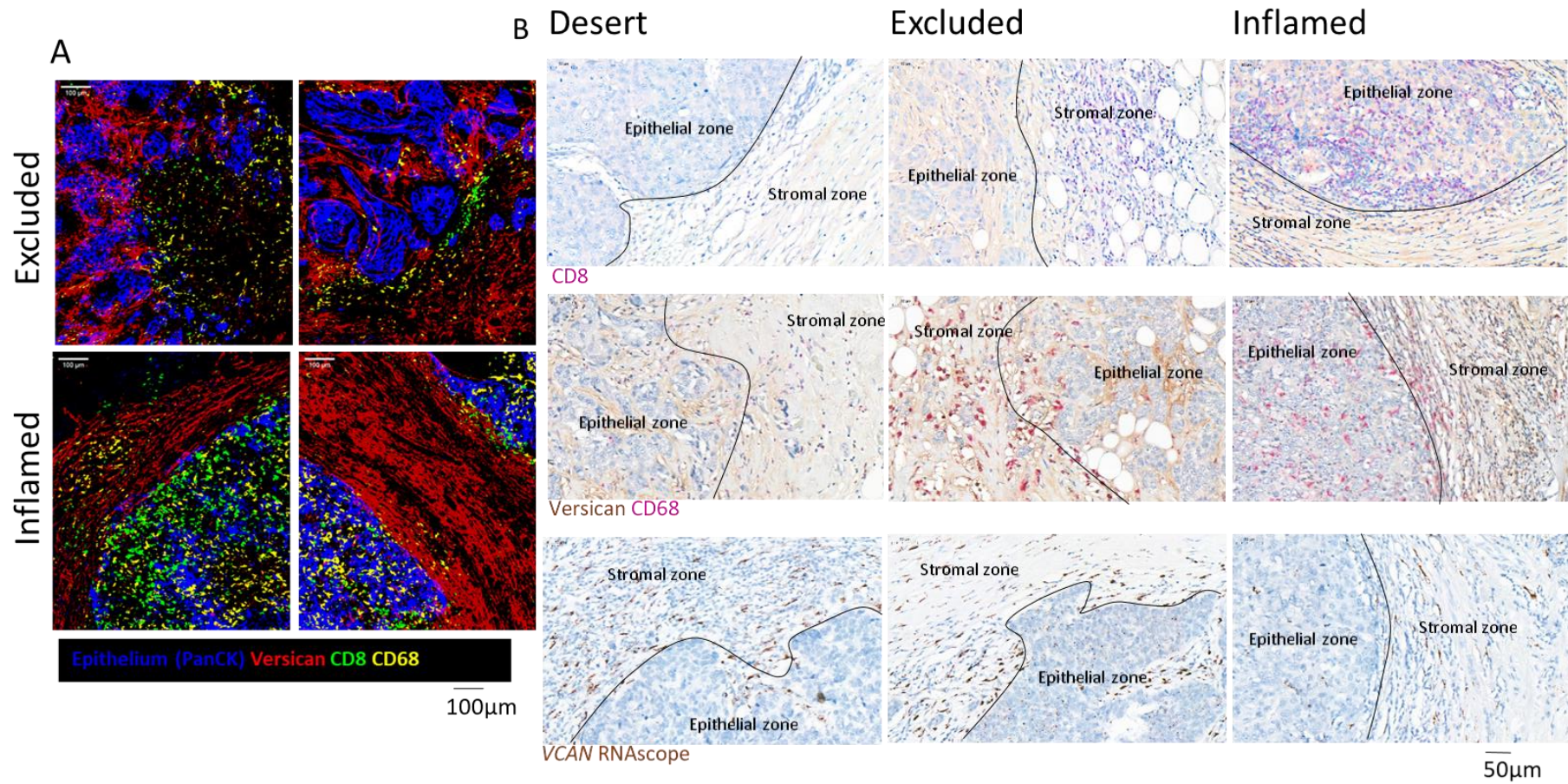
in the epithelial zone showing that the intensity was not a factor for this difference (Figure 3.12G).



**Figure 3.12. VCAN expression and immune phenotype.** A) Dual staining for VCAN-CD68 and CD8 analysed in same areas for the 4 different regions. VCAN expression was determined in the same regions. Spread of VCAN expression in different regions shown. B) Spread of VCAN expression in excluded and inflamed tissues. RM One-way ANOVA. C) Overall VCAN expression and expression in epithelial zone and stromal zone between phenotype. Welch ANOVA. D) RNAscope staining for VCAN analysed for VCAN expressing cells in the epithelial zone and stromal zone. Number of VCAN+ cells in each phenotype and within the epithelial and stromal zone showed. Welch ANOVA. E) Phenotype determined per area of the tissue. VCAN levels in epithelial zone within each area of the tissue. F) Percentage of T cells in the epithelial and stromal areas determined from the overall number of cells. Correlation between percentage of VCAN and CD8 in the epithelial and stromal zone. Spearman Rank Correlation. G) Percentage of VCAN at different intensities transformed into a score. Welch ANOVA. \* =  $P < 0.05$ , \*\* =  $P < 0.01$ .

Representative images of the different TIPs are shown in Figure 3.13. Stains for PanCK, VCAN, CD8 and CD68 were deconvoluted and overlaid using Image J. Analysis of the overlapped images demonstrated that VCAN (red) is present around the tumour cells (Pan-CK - blue) in the excluded tissues but in the stroma in inflamed tissues. A difference in immune cell infiltration can also be observed with greater numbers of CD8 (green) and CD68 (yellow) cells in the tumour areas of inflamed tissues. (Figure 3.13A) The differences in VCAN expression at gene and protein levels as well as immune cell infiltration into the epithelial zone between the different phenotypes are shown in Figure 3.13B. Overall from the analysis, it could be concluded that VCAN expression within the epithelial zone is associated with an excluded phenotype.

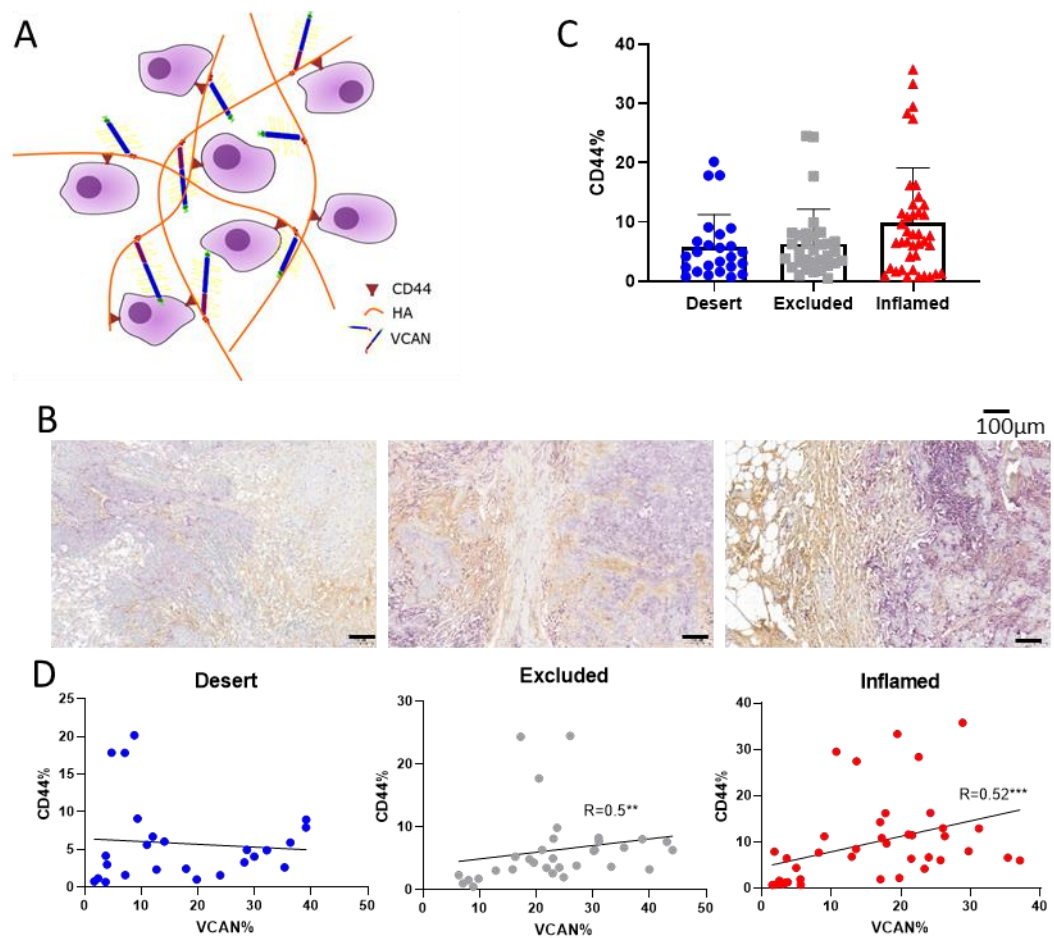




**Figure 3.13. Representative images of immune phenotypes and VCAN expression.** A) IHC staining for PanCK, VCAN, CD8 and CD68 deconvoluted on FIJI and overlapped. B) Representative images of the different phenotypes of VCAN, CD8 and CD68 IHC and VCAN RNAscope.

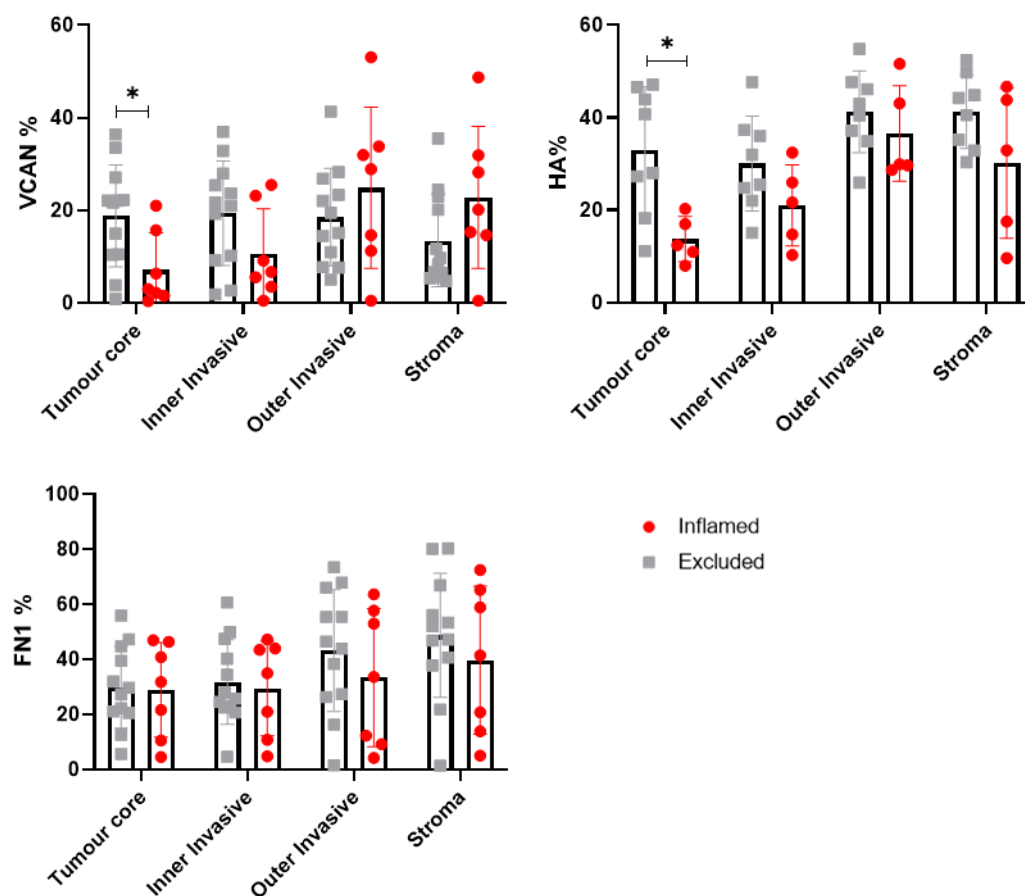


The localisation of VCAN can be affected by other proteins within the ECM. CD44 is a receptor expressed by multiple cell types including tumour cells and can bind VCAN directly through the CS chains or indirectly through its binding to HA<sup>220,221</sup> (Figure 3.14A). It was hypothesised that tumour cells in excluded tissues may have high levels of CD44 which will lead to greater levels of VCAN. Using a dual stain, VCAN (brown) and CD44 (purple) were stained on the same tissue sections (Figure 3.14B). The epithelial zone was quantified in the same areas as the TIP analysis. This allowed the CD44 expression to be compared to the phenotypes within the tissue. From the analysis, there was a greater level of CD44 in the inflamed areas however this wasn't significant (Figure 3.14C). It must also be noted that immune cells can also express CD44, which could be causing the increase in levels in inflamed tissues. When comparing CD44 to VCAN, a significant positive correlation was seen in both excluded and inflamed tissues (Figure 3.14D), indicating no clear link between the expression of CD44 to VCAN expression in excluded areas.



**Figure 3.14. CD44 expression and immune phenotype.** A) Diagram showing how VCAN and HA interact with CD44 on tumour cells. B) Dual IHC completed for VCAN (brown) and CD44 (purple). C) Quantification of CD44 staining in the epithelial zone. Kruskal Wallis test. D) Correlation between VCAN and CD44. Spearman rank correlation. \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$

The differences between regions in the immune phenotypes were also determined for VCAN associated matrix markers. The same areas were analysed on TNBC patient tissues stained for HA and FN1 (Figure 3.15). Similarly to VCAN, HA showed a significant difference between excluded and inflamed tissues in the tumour core. However, within the outer invasive and stroma regions, the levels of HA in the excluded tissue were higher than in the inflamed and the inverse for VCAN. . No difference was recorded between excluded and inflamed tissues in the tumour core and inner invasive regions for FN1 although a slight increase was seen in the stromal regions. From this data, I found that there may be some association between the binding of HA and VCAN towards the TIP.



**Figure 3.15. Comparisons of matrix markers in CD8 excluded and inflamed tissues.** Expression of VCAN, HA, and FN1 in excluded and inflamed tissues within the four regions across the tumour epithelial border. Multiple Mann-Whitney tests. \* = P<0.05

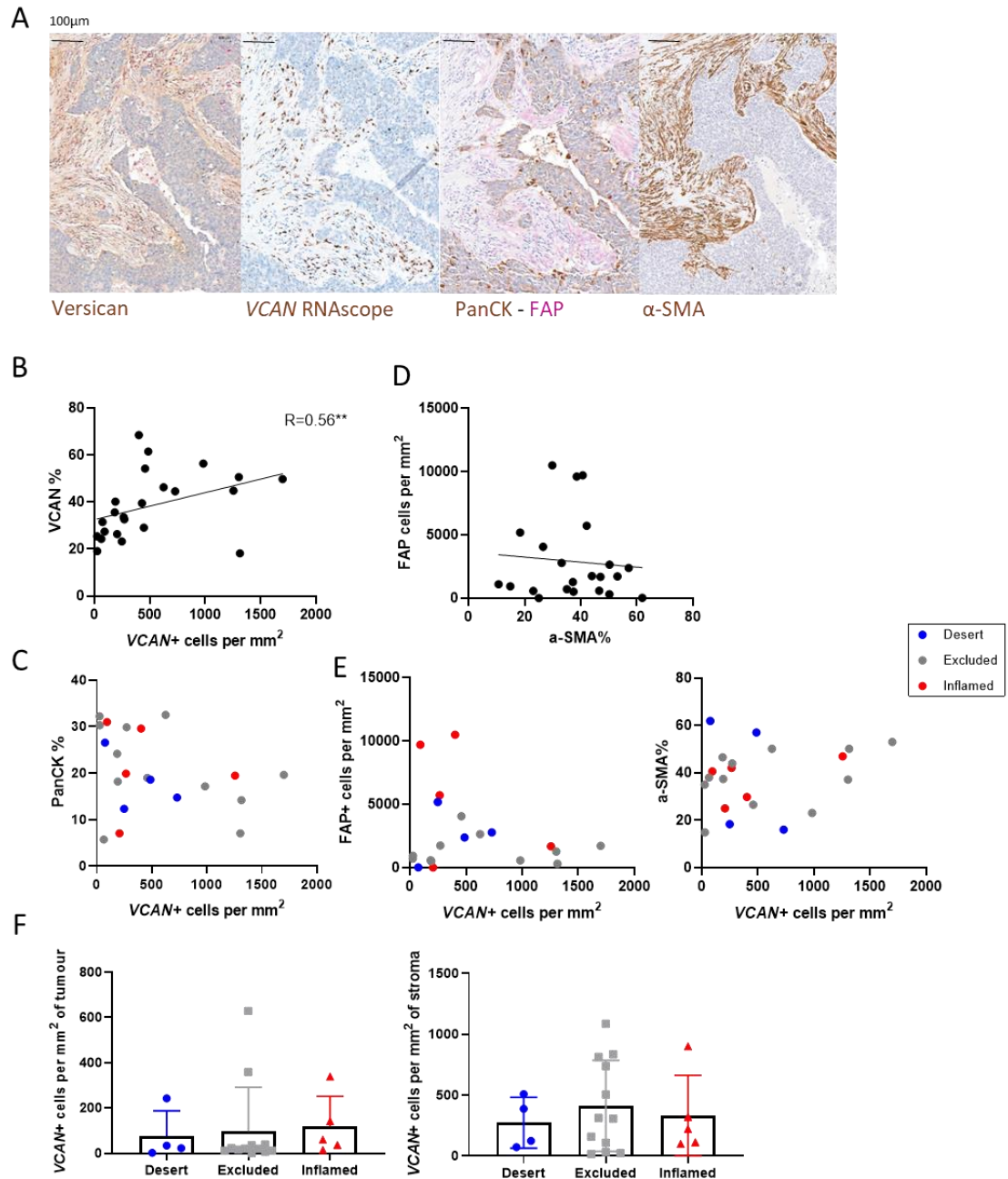
To deduce whether HA was associated with VCAN localisation and expression, the levels of HA were compared with VCAN. The trend of HA expression in the excluded tissues contrasted to VCAN where the levels increased in the stromal regions (Outer invasive and stroma), however the inflamed tissues were similar with HA significantly increasing in the outer invasive region. The correlation of HA to VCAN was then determined for the different regions. The epithelial zone – tumour core and inner invasive regions – did not show any correlations for both the excluded and inflamed tissues. This suggests that HA is not linked to the significant increase in VCAN seen in the epithelial zone of excluded tissues. Positive correlations were seen between VCAN and HA in the outer invasive region of both excluded and inflamed tissues and in the stroma of inflamed tissues (Figure 3.16). Overall from this it can be concluded that despite observing higher HA levels in the tumour core of excluded tissues, this was not associated with VCAN expression, and the association of VCAN and excluded phenotype is independent of HA.



### 3.6. Tumour cells and fibroblasts express VCAN in TNBC

VCAN can be expressed by multiple cells in the TME as seen in Figure 3.6E. This includes tumour cells and fibroblasts. The expression of VCAN was explored in the TNBC tissues and compared to cell markers for tumour cells (PanCK) and fibroblasts (FAP,  $\alpha$ SMA) (Figure 3.17A). Areas around the tumour epithelial border were sampled around the tissue for analysis. Comparison of VCAN+ cells from the RNAscope analysis with the accumulation of VCAN protein showed a positive correlation ( $R=0.56^{**}$ ) (Figure 3.17B) indicating a correlation between transcription and translation of VCAN. To look at how disease level would associate with VCAN expression, the percentage of tumour area using PanCK was compared to the VCAN expression. No correlation was observed showing that tumour area does not affect VCAN expression (Figure 3.17C).

Two commonly reported CAF markers associated with immune exclusion are FAP and  $\alpha$ -SMA<sup>222</sup>. No correlation was observed between these markers indicating they are expressed independently of each other (Figure 3.17D). To deduce whether the activation of fibroblasts to express these markers is associated with VCAN expression, the levels of these markers were compared to the number of VCAN+ cells from the RNAscope technique. The correlation of the markers and VCAN expression was calculated overall and per TIP. No correlation was observed between both markers and VCAN expression (Figure 3.17E). I then looked at whether the levels of expression by tumour cells and stromal cells were associated with immune phenotype. Areas of the tumour and stroma were determined and the number of positive cells in these regions were counted. No difference was observed between the number of VCAN+ tumour cells and VCAN+ stromal cells in each phenotype (Figure 3.17F). Overall from the analysis it was observed that there was no specific cell or cell phenotype that seemed to be associated with VCAN expression in excluded tissues.

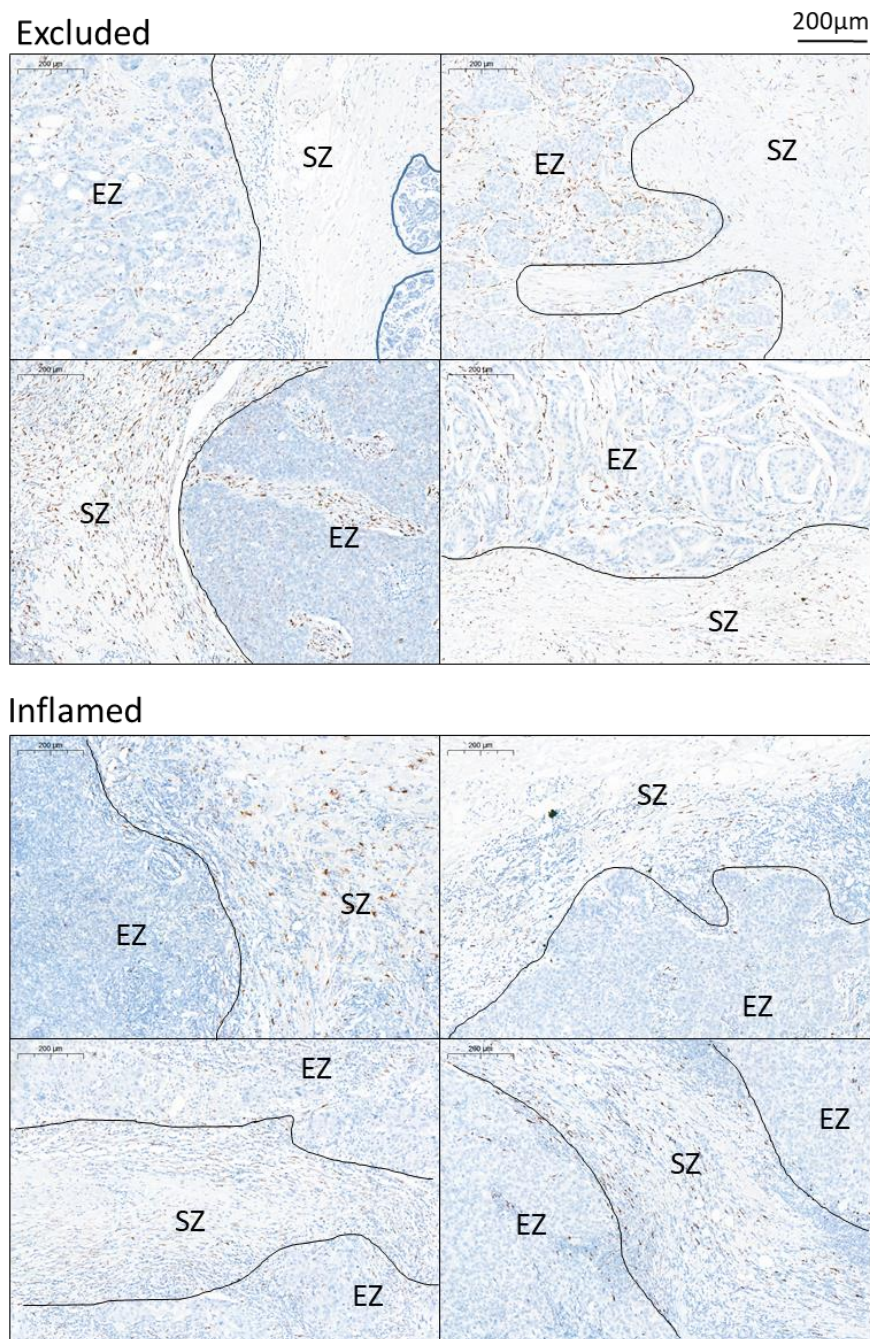


**Figure 3.17. Comparing VCAN expression in tumour cells and fibroblasts.** IHC stains for PanCK, FAP, α-SMA and VCAN were compared to VCAN RNAscope staining. Representative images from 18 tissues analysed. B) Number of VCAN expressing cells were compared to the area of VCAN protein. C) Number of VCAN expressing cells compared to PanCK expression. D) Comparison of levels of fibroblast markers FAP and α-SMA. E) Number of VCAN expressing cells compared to fibroblast markers α-SMA and FAP. E) Effect of stroma area compared to fibroblast marker. Spearman Rank correlation used for all comparisons. F) Comparing VCAN+ tumour cells and VCAN+ stromal cells in the different phenotypes.

Next, the location of the VCAN expressing CAFs were compared between the excluded and inflamed phenotypes. It was observed that within the excluded tissues, higher levels of VCAN was expressed by fibroblasts within the epithelial



zone compared to inflamed tissues, where VCAN expressing fibroblasts were scattered around the stromal zone (Figure 3.18). This led to the impression that VCAN expression in the excluded tissues was triggered by cell-cell communication between tumour cells and fibroblasts.



**Figure 3.18. RNAscope of excluded and inflamed tissues.** DAB (brown) staining shows the presence of VCAN RNA within the cells. Images representative for 12 excluded tissues and 7 inflamed tissues. SZ – stromal zone, EZ – epithelial zone.



Having found predominantly fibroblasts and tumour cells are expressing VCAN in the TNBC tumour microenvironment with little expression by other cells such as myeloid cells (Figure 3.6E), I wondered whether the type of VCAN isoform expressed may vary between cell types and how cell to cell interactions could affect isoform expression. The roles of each isoform differ as detailed in Chapter 1, therefore identifying which isoforms are highly expressed could indicate the function of VCAN within the tissue. V0 and V1 isoforms are associated with tumour proliferation and inflammation<sup>144,223</sup> whilst V2 promotes apoptosis<sup>147</sup>. The V3 isoform has been found to be inhibitory to the inflammatory effects of V0 and V1 due to the absence of CS chains<sup>223</sup>. The V4 isoform has only been detected in breast cancer, therefore I was interested in identifying whether it was expressed in TNBC tissues and cells<sup>65</sup>.

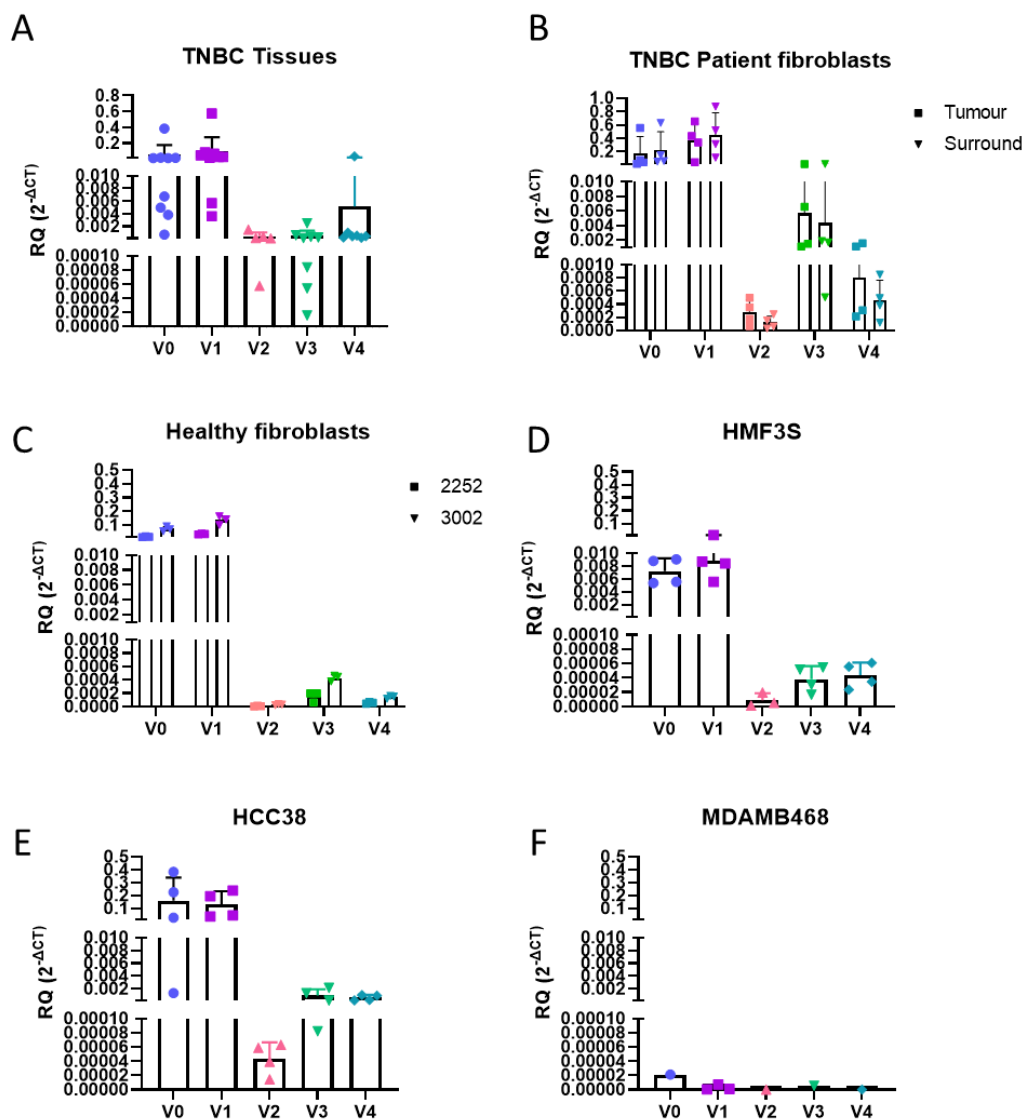
RNA was extracted from 9 frozen TNBC tissues. Tissues were studied to have an idea of which isoforms were most prominent within TNBC. V0 and V1 were the highest expressed isoforms and present within all tissues. The isoform V2 was not at a detectable level for four tissues whilst one tissue also had isoforms V3 and V4 at levels lower than the detectable range (Figure 3.19A). Comparable patterns of isoform expression were observed across the 9 tissues analysed.

Next I used qRT-PCR to test VCAN isoform expression within individual cell populations. Primary fibroblasts were isolated from 4x TNBC patients and 2x healthy donors by the breast cancer now tissue bank. The TNBC patient fibroblasts were isolated from tumours and their matched surround tissues where surround is defined as the areas greater than 5cm from the tumour edge. Analysis of the primary fibroblasts showed a similar pattern of isoform expression. However lower levels were seen with V2 and V4. Higher levels of V4 were seen in the fibroblasts from tumour areas compared to the surround however this was not significant (Figure 3.19B). For the healthy donor fibroblasts, the pattern of isoform expression was similar to both the tumour derived fibroblasts and whole tissue analysis, however expression levels were lower, indicating that there is an upregulation of VCAN expression in tumour tissue (Figure 3.19C). HMF3S, an immortalised human mammary fibroblast cell line was also tested and levels of V0 and V1 were also the highest with the levels of V2, 3 and 4 being quite low. These levels were comparable to the healthy fibroblasts (Figure 3.19D).

TNBC tumour cell lines also expressed VCAN in a similar isoform pattern to that seen in whole tissues and fibroblasts, however the levels of expression were quite

variable. For example, the HCC38 cell line expressed each isoform at a comparable level to the fibroblasts and whole tissue (Figure 3.19E), whereas MDA MB468 cell line expressed low levels of VCAN with the isoforms V2, 3 and 4 just within detectable limits (Figure 3.19F).

Overall, from this analysis, it can be concluded that the V0 and V1 isoforms are the highest expressed isoforms in TNBC and that the expression of isoforms are not specific to cell type. The increase in expression from healthy to tumour associated fibroblasts indicates the TME does increase VCAN expression within fibroblast populations.



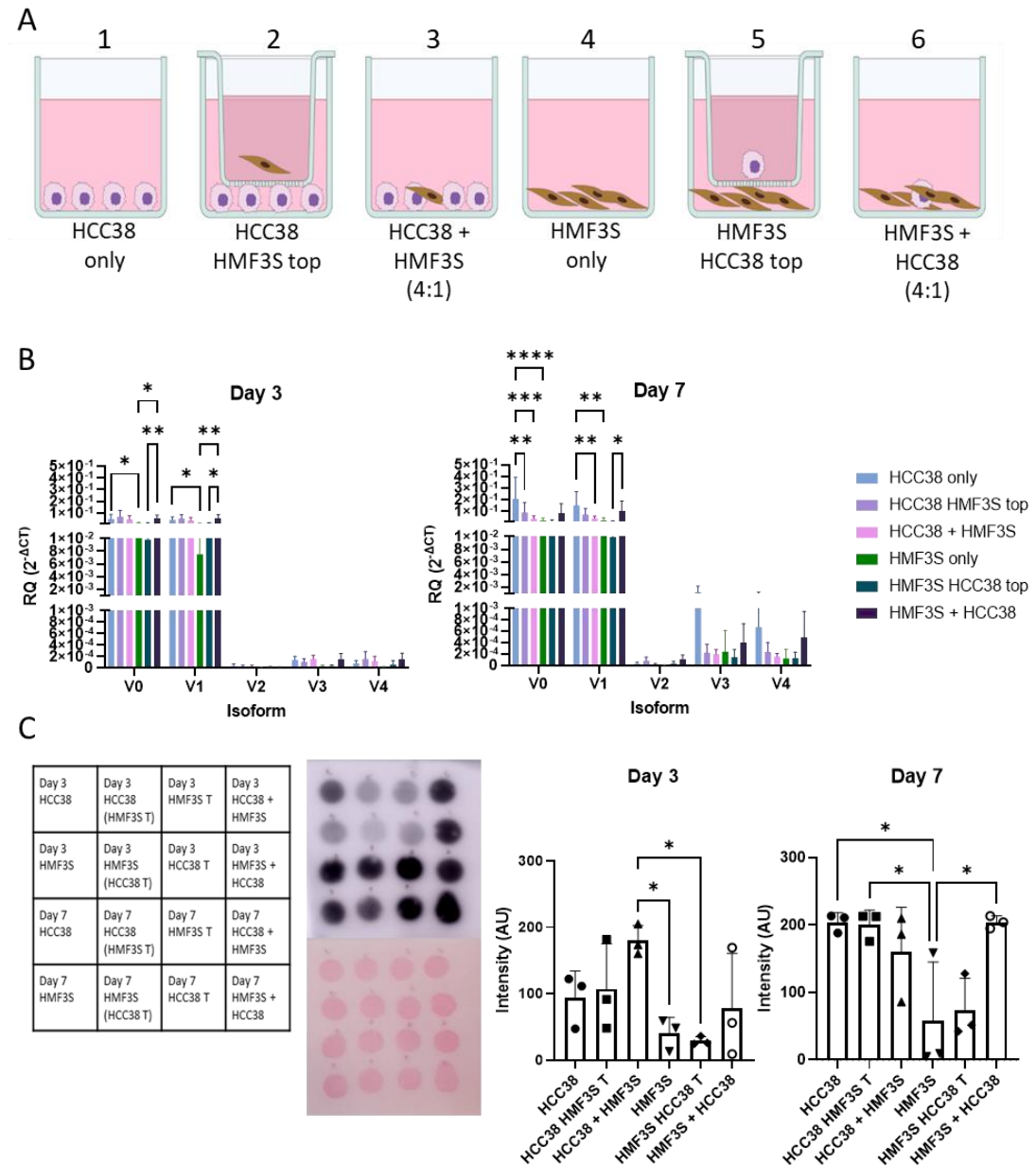
**Figure 3.19. Expression of VCAN isoforms in tissues and cell lines.** RNA was extracted from A) frozen TNBC tissues (N=9), B, C) primary fibroblasts from TNBC patients (N=4) and healthy donors (N=2), D) immortalised fibroblast cell line (HMF3S) and E, F) TNBC cell lines HCC38 and MDAMB468 (N=4). qRT-PCR was completed for each isoform and normalised to housekeeping gene RPS13.

To explore whether a tumour-fibroblast interaction may enhance VCAN expression and potentially production of certain isoforms, I investigated how cell to cell contact can affect VCAN expression. As discovered from the RNAscope analysis, excluded tissues showed fibroblasts in close proximity to tumour islands whilst in inflamed tissues the fibroblasts were away from the tumour cells. This suggested perhaps some tumour cells were influencing CAF location and VCAN expression, or vice-versa. To explore tumour cell –fibroblast cross talk in regulating VCAN expression a co-culture and transwell model was used. Six condition arms (C1-6, Figure 3.20A) were compared, mono-cultures, a transwell model where tumour and fibroblast cells are separated, and a co-culture where tumour and fibroblast cells could make contact. A protein free medium was used for the cell-culture with media collected and RNA extracted at days 3 and 7.

At day 3, the qRTPCR analysis showed that in all conditions V0 and V1 were the highest expressed. No significant differences were observed between HCC38 cells alone (C1) or with HMF3S in a transwell (C2) and co-culture model (C3). HCC38 cells alone (C1) had a significantly higher expression than HMF3S alone (C4). Co-culturing HMF3S with HCC38 cells (C6) significantly increased expression in comparison to HMF3S alone (C4) or in a transwell with HCC38 (C5). At day 7, there was an increase in the expression of V3 and V4 isoforms. For the V0 isoform, growing HCC38 alone (C1) showed a significantly higher expression compared to the co-culture with HMF3S (C3) and HMF3S cells alone (C4). For the V1 isoform HCC38 cells alone (C1) expressed significantly higher levels than in a co-culture (C3) and HMF3S cells alone (C4). HMF3S expression of V1 was increased through co-culture with HCC38 (C6). It was interesting to observe that HMF3S cells with HCC38 in a co-culture (C6) had increased levels of V0, V1 and V4 compared to HCC38 with HMF3S in a co-culture (C3) (Figure 3.20B). This was not significant but indicated the importance of cell proportions towards gene expression. Overall from this data, it could be seen that there is an increase in expression from day 3 to day 7 with all isoforms. The effect of the co-culture was clear with the HCC38 cells, where the addition of HMF3S seemed to reduce expression at day 7. As this was observed in both a transwell and co-culture model, it suggested that direct cell to cell contact was not important for this effect. With the HMF3S cells, comparisons of the co-culture and the transwell showed the importance of cell-cell contact. When comparing this to the TIPs where excluded tissues are represented by the co-culture and inflamed by the transwell model, it suggests that high expression of the

V0, V1 and V4 isoform are associated with an excluded phenotype with a higher level observed in the HMF3S co-cultures.

The impact on the protein level was determined by detecting the level of protein with dot blot analysis. The membrane was stained with ponceau to compare the overall protein levels and probed for VCAN. The G1 domain of VCAN was probed for to detect all isoforms. The highest level of VCAN was detected in the co-culture with HCC38 and HMF3S (C3) at day 3. At day 7, no impact was seen in the expression from HCC38 for any condition. A striking increase was observed in VCAN expression from HMF3S when co-cultured with HCC38 for 7 days (Figure 3.20C), this increase was similar to what was observed at RNA levels (Figure 3.20B). Protein expression from the HCC38 conditions (C1-3), did not seem similar to the RNA levels where the mono-culture of HCC38 had the highest expression. Overall I conclude from these data that VCAN expression is stimulated by tumour-fibroblast interaction, which may result from cell-cell contact rather than a secreted factor, although more work would need to be conducted to be conclusive.



**Figure 3.20. Effect of cell cross-talk on VCAN expression.** HCC38 cells and HMF3S cells were seeded in a 24 well plate at density of 40,000 cells. The cells were either left in a monoculture (condition 1, 4), transwell with alternate cells added (condition 2, 5) or co-culture with the alternate cell (condition 3,6). In both the transwell and co-culture 10,000 cells of the alternate cell was added. The media of the cells was changed to protein free media. At Day 3 and Day 7 the media was collected and RNA extracted from the bottom well. B) qRTPCR was completed to detect changes in isoform expression. C) A dot blot was completed with 15 $\mu$ l media and probed for anti-VCAN to detect changes in level of VCAN protein. The membrane was stained with ponceau to compare protein levels. Densitometry was completed to compare VCAN levels. N=3. ANOVA. \*= $P<0.05$ , \*\*= $P<0.01$ , \*\*\*= $P<0.005$ , \*\*\*\*= $P<0.001$ .

### 3.7. Discussion

TNBC is known to have the highest level of immune cell infiltration compared to other types of breast cancer, however as seen from the data presented here, this can vary between patients. Studying the complexity of different TIP phenotypes in TNBC could be useful to better understand and target TNBC. A strategy to assign TIPs to cancer tissues was first reported by Kather et al <sup>44</sup>, where multiple tumour types and immune cell markers were analysed. TNBC was not included within this analysis so we looked to assign these TIPs to our subset of TNBC tissues. The analytical technique shown in this paper was initially followed to classify the TIPs, however, from the results there was a significant difference between the number of CD8+ cells in the inflamed and excluded tissues, indicating that this phenotype may describe a poor overall infiltration, as opposed to a poor infiltration into the tumour epithelium. For this reason, the analysis technique was modified to first identify desert tumours and then used an immune cell based ratio of 0.75 to discriminate between excluded and inflamed tissues. By using a ratio to interpret the excluded and inflamed tissues, this method can be applied to different tissue libraries and results can be comparable. I have found that the excluded phenotype makes up the majority of tissues and also that more than one of the TIP phenotypes can be present in the tissues. A quarter of the patients showed an inflamed phenotype which has been shown to have a positive association with patient outcome <sup>224</sup> as well as immunotherapy response <sup>225</sup>. Clinical trial data has shown that overall response rate to pembrolizumab only in TNBC patients was 21.4% <sup>226</sup>, similar to the number of inflamed tissues identified in this study. The identification of TIPs in TNBC has since been completed by different labs using varying analysis techniques. Gruosso et al, also identified excluded tissues being the highest phenotype (60% of tissues) <sup>45</sup>, whilst Hammerl et al only classified 26% as excluded <sup>49</sup>. The ranges in the classification of phenotypes shows the importance of a conclusive method to identify TIPs. To confirm that the method used in this analysis was superior it would need to be compared against the other methods in patient tissues where data is available for immunotherapy responses.

Previous studies exploring TIPs have not considered the possibility of intratumoural heterogeneity. Through my analysis I found that the TIP's could vary across a single tissue. Intratumoural heterogeneity of infiltrating lymphocytes has been associated with variations in immunotherapy response <sup>227</sup>. This shows the importance of spatial analysis of tissues to have a clear understanding of its phenotype and potential response to therapy.

The matrix has been identified to play a key role in the migration of immune cells and immunotherapy response <sup>228</sup>. Having identified the phenotypes, we wanted to look into possible targets within the matrix that are associated with the different phenotypes and immune cell trafficking. In a study previously completed in the lab, a signature was identified with 22 molecules that associated with immune suppression and tumour progression. From this signature 5 molecules were upregulated at both gene and protein expression<sup>23</sup>. To look at the importance of these matrix molecules, I first looked at the expression of these genes in publically available TNBC datasets. The genes were correlated to immune cell subtypes using CIBERSORT to identify the cells from the RNAseq data. All genes were found to positively correlate with macrophage infiltration. TAMs have been found to adhere to FN1 and collagens through integrins in breast cancer <sup>229</sup>, while proteoglycans like VCAN can affect TAM migration through the attraction of chemokines <sup>230-232</sup>. I then looked at the protein expression of these molecules in TNBC TMAs. 5 consecutive sections of the TMA were previously stained for these markers. For the analysis only a small number of cores could be analysed as the cores were required to be in a condition for analysis for all 5 markers, which was not the case for a majority of the cores. Initial observations of the TNBC TMA cores showed that there were similarities between VCAN and CTSB expression. FN1 showed a high co-localisation with all molecules, however the expression was ubiquitous which led me to think that any relations with other proteins were not specific. VCAN and CTSB had similar levels of co-localisations with other proteins however they did not show high levels of co-localisation with each other. No previous studies have shown links between VCAN and CTSB, although TGF $\beta$  may be the driver of an indirect association by increasing VCAN production and by being activated by the latent form of CTSB <sup>126,233</sup>.

To look further into the expression of VCAN and CTSB, dual RNAscope was completed with the aim of identifying co-expressing cells. CTSB was expressed in all cells as seen in the scRNAseq data, while VCAN was mostly stromal. As VCAN has been previously associated with T cell infiltration <sup>234,235</sup>, the expression of this protein was studied in relation to CD8+ T cell localisation. The expression of VCAN by the cells surrounding the tumour was found to be associated with CD8+ T cells localisation, raising the hypothesis that VCAN might be implicated in the poor infiltration of the CD8 T cells in the tumour. For this reason, VCAN was chosen as protein of interest associated to immune exclusion in our subset.

The association between VCAN and CD8+ T cells was further explored by comparing the protein expression to the total number of CD8+ T cells and CD68+ macrophages and the ratio of these cells in the tumour and stroma. A negative correlation was observed, further confirming that VCAN is associated with CD8+ T cell exclusion. VCAN has multiple binding partners within the ECM <sup>92</sup> which can orchestrate the architecture of the tissue and affect the organisation of VCAN in the ECM. To study whether the association of VCAN and immune exclusion was independent to its ECM binding partners – HA, FN1 and collagen, these binding partners were studied for their correlation to the tumour stroma ratio of CD8+ T cells and CD68+ macrophages. HA and FN1 displayed no correlation while collagen had a positive correlation, confirming that the association seen with VCAN was independent to its binding partners.

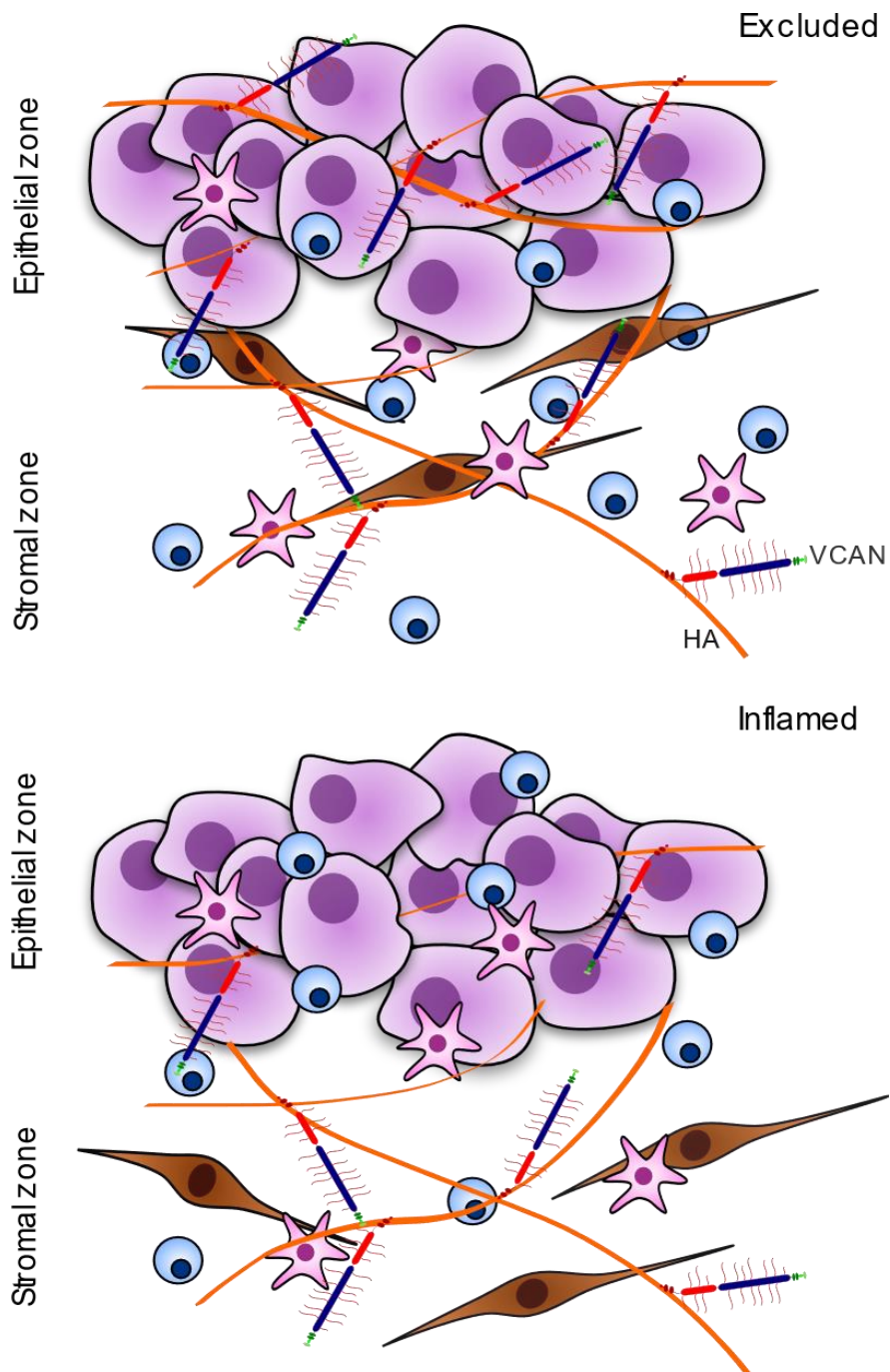
The levels of VCAN were then determined in the regions tested to see whether the expression of VCAN in specific areas was linked to an excluded phenotype. Excluded tissues were found to have higher levels of VCAN in the epithelial zone. High levels of matrix lining tumour islands have been shown to act as an obstacle for T cell migration <sup>236</sup>. Proteoglycans can be present within the glycocalyx of tumour cells. The glycocalyx around tumour cells has been questioned for its role in immune cell evasion <sup>237,238</sup>. VCAN is not a membrane bound proteoglycan so it is not directly part of a cell glycocalyx, however it can bind to other components of it. For example, CD44 was investigated as it can bind to VCAN directly through the CS chains or indirectly through HA <sup>91</sup>. No difference in CD44 was seen between excluded and inflamed tissues which may be due to CD44 also being a marker for T cell activation. A positive correlation was seen between VCAN and CD44 in both excluded and inflamed, suggesting VCAN binding to CD44. With no significant difference between CD44 levels in excluded and inflamed tissues it could not be confirmed that this was the cause for high VCAN levels in the epithelial zone. HA was also examined to see if higher levels in the epithelial zone affected VCAN levels, however no correlations were observed between VCAN and HA in the epithelial zone in both inflamed and excluded tissues. This showed that HA was not linked to the possible role of VCAN in immune exclusion.

The production of VCAN between the inflamed and excluded tissues were compared. CAFs could be seen to be the highest expressing cells of VCAN from the RNAscope images and the scRNAseq data. Previous studies have shown that CAFs expressing FAP and  $\alpha$ SMA are involved in CD8+ T cell exclusion <sup>222</sup>. These markers were stained to see whether their expression associated with VCAN



transcription via comparison with RNAscope. No correlation was observed between FAP and  $\alpha$ SMA with *VCAN* RNA expression, showing that CAF expression of these markers cannot be directly associated with *VCAN* levels. The location of the CAFs seemed to show greater importance with *VCAN*<sup>+</sup> CAFs observed around the tumour islands in excluded tissues but spread across the stroma in the inflamed. Tumour cells have been shown to affect the activation of fibroblasts to CAFs, leading to the production of matrix proteins <sup>7</sup>. Co-culture and transwell studies with tumour cell line HCC38 and immortalised mammary fibroblast HMF3S, showed that the cell to cell contact between the cells can affect the expression of *VCAN*, with HCC38 increasing the *VCAN* levels in HMF3S. Yeung et al showed that TGF $\beta$  expressed from tumour cells was able to induce *VCAN* expression in fibroblasts and the neutralisation of TGF $\beta$  with an antibody significantly reduced *VCAN* expression <sup>126</sup>. This pattern could be seen to be associated with the excluded phenotype where the expression is triggered in fibroblasts surrounding tumour cells. Contrastingly the fibroblast cell line HMF3S looked to reduce *VCAN* expression in the TNBC cell line HCC38. Fibroblast derived matrices have been shown to affect the mRNA levels of genes in cancer cells, this includes matrix genes such as *COL4A1* <sup>239</sup>. Further experiments to look at the HMF3S derived matrix would be required to see if this is the case.

To conclude, from this chapter, I was able to identify all three TIPs within TNBC tissues and associate these phenotypes to the matrix protein *VCAN*. Initial studies were completed to look at the expression of *VCAN* where it was observed that excluded tissues had greater levels in the epithelial zone (**Figure 3.21**). The possible cause behind this could be through the stimulation of fibroblasts. In the next chapter, the post-translational modification of *VCAN* with the addition of CS chains will be studied to understand how the structure of *VCAN* could affect CD8<sup>+</sup> T cell migration.



**Figure 3.21. Summary of excluded and inflamed tissue phenotypes.** Excluded tissues found to have high levels of VCAN in the epithelial zone with VCAN expressed by fibroblasts close to tumour cells. Inflamed tissues have VCAN expressed predominantly in the stroma with VCAN expressed by fibroblasts distant from tumour cells.

## Chapter 4: The expression and arrangement of chondroitin sulphate proteoglycans in the TME

## 4.1. Introduction

From the previous chapter it was identified that VCAN is associated with exclusion in tissues. Interestingly high expression of VCAN in excluded tumours was seen within the tumour parenchyma, whereas inflamed tissues showed low or appeared negative for VCAN in the same region. Tumour cell and fibroblast interactions were studied and were found to effect the expression of the V0 and V1 isoforms.

The structure of VCAN is important to its function within the TME. One of the key structural components of VCAN is the disaccharide CS. The V0 and V1 isoforms display the highest number of CS chains. CS is glycosylated to VCAN via the GAG domains as described in Chapter 1. Other proteins within the ECM can express CS. Using mass spectrometry, previous studies have identified multiple proteins express CS<sup>240,241</sup>. These proteins were identified to be known proteoglycans as well as glycoproteins.

CS has a negative charge and can bind to water and chemokines in ECM affecting the structure of the ECM and the migration of immune cells. The expression of CS within the ECM is not consistent amongst proteoglycans. CS is formed of GalNAc and GluC which can be sulphated at different sites to give rise to different isoforms<sup>242</sup>, further implicating the roles it has within the TME.

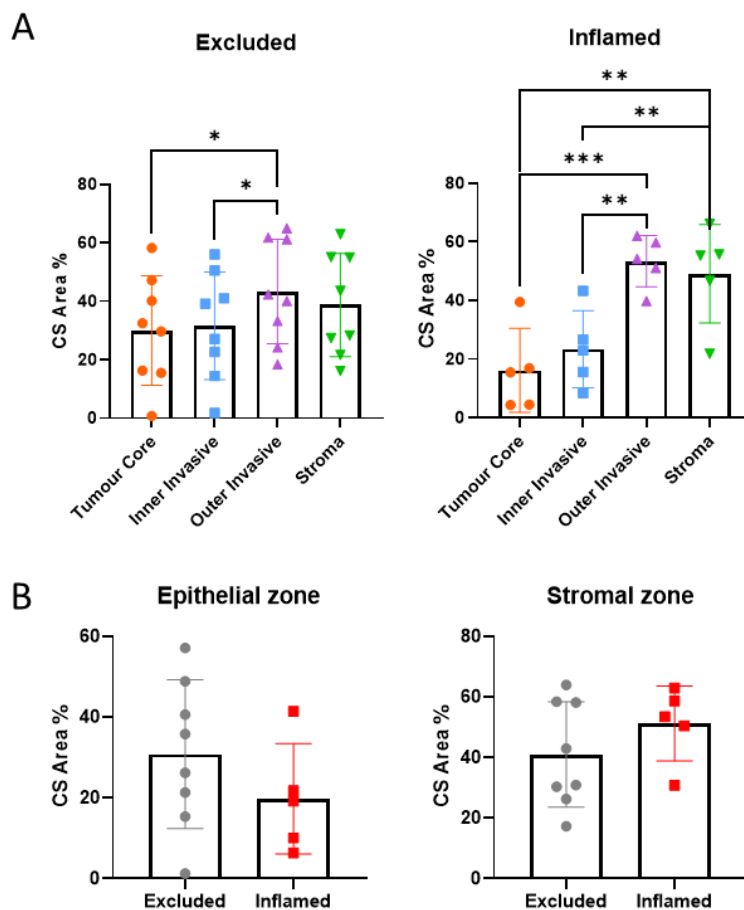
In this chapter, I looked at the levels of CS in the different regions around the tumour epithelium identified in Chapter 3 and how it varies in inflamed or excluded tissues. Highly expressed CS proteoglycans BGN and DCN, were analysed to understand if the associations observed were due to CS or to VCAN specifically. I then studied the different CS isomers in the tissue to identify whether a particular isomer was associated with an excluded phenotype.

### 4.1.1. Chapter highlights at a glance:

- VCAN correlates with CS in inflamed tissues
- Proteoglycans DCN and BGN do not associate with TIPs
- Levels of CS-A and CS-C isomers in the stroma differentiate inflamed and excluded tissues

## 4.2. Localisation of CS in TNBC

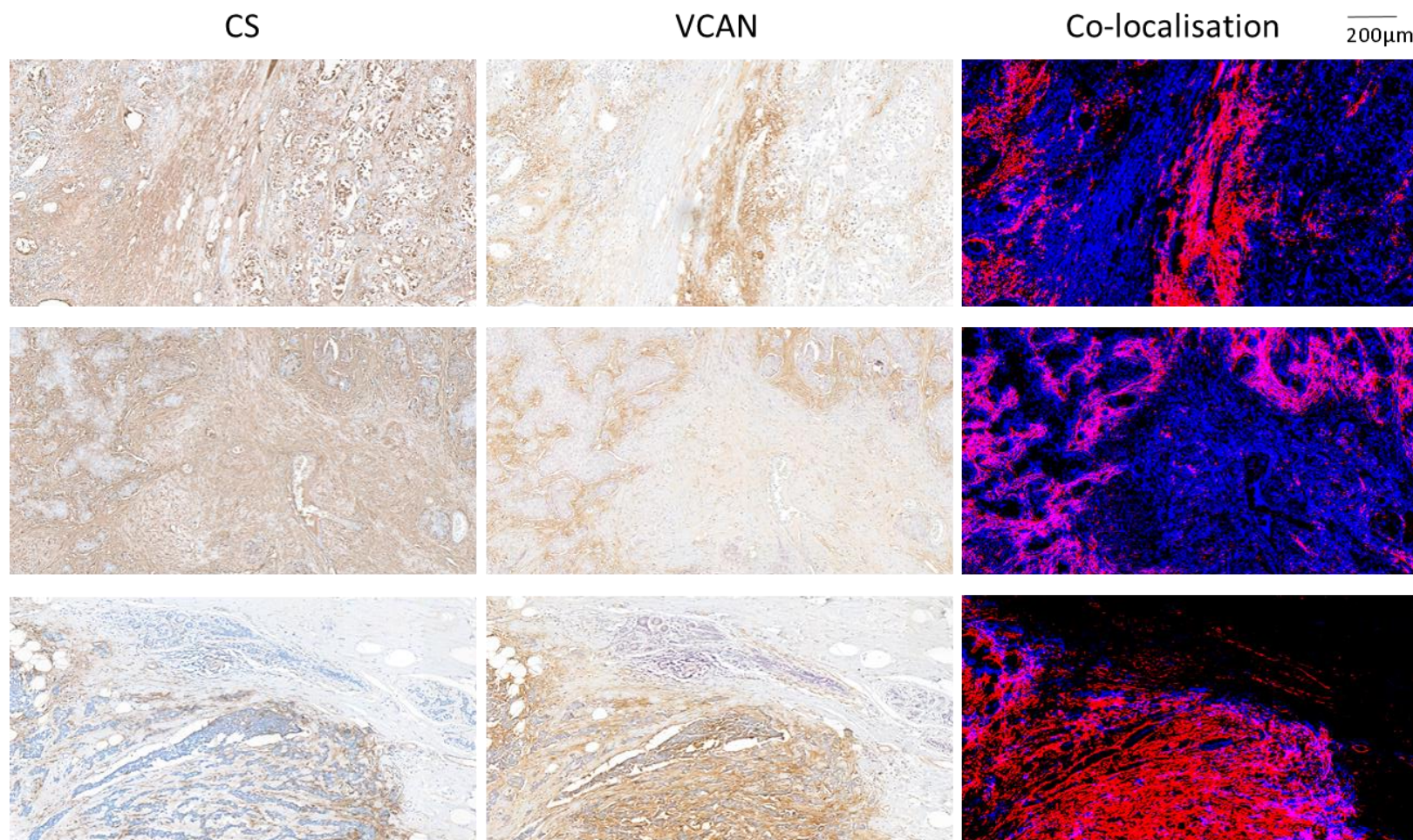
The levels of CS were explored in the four regions across the tumour epithelial border. In the excluded tissues CS was significantly higher in the outer invasive area compared to the tumour core and inner invasive. This was also the case with the inflamed tissues where there was also a significantly higher level in the stroma (Figure 4.1A). The trend in CS expression within the inflamed tissues were similar to the trend observed with VCAN (Figure 3.12B). With a significantly higher level of VCAN seen in the epithelial zone of excluded tissues compared to inflamed tissues (Figure 3.12C), I looked at the levels of CS in the different zones. From the analysis, there was a higher level of CS in the epithelial zone of excluded tissues but this was not significant. In the stromal zone there was a high level in all the inflamed tissues whilst there was more variability in the excluded tissues which almost appeared to represent two groups, a high and low CS group (Figure 4.1B).



**Figure 4.1. CS expression across the tumour epithelial border.** A) Level of CS across the different regions in 9 excluded tissues and 5 inflamed tissues determined from IHC staining. RM One-way ANOVA. \*= $P < 0.05$ , \*\*= $P < 0.01$ , \*\*\*= $P < 0.005$ . B) Comparison of CS levels in the epithelial zone and the stromal zone. T-test.

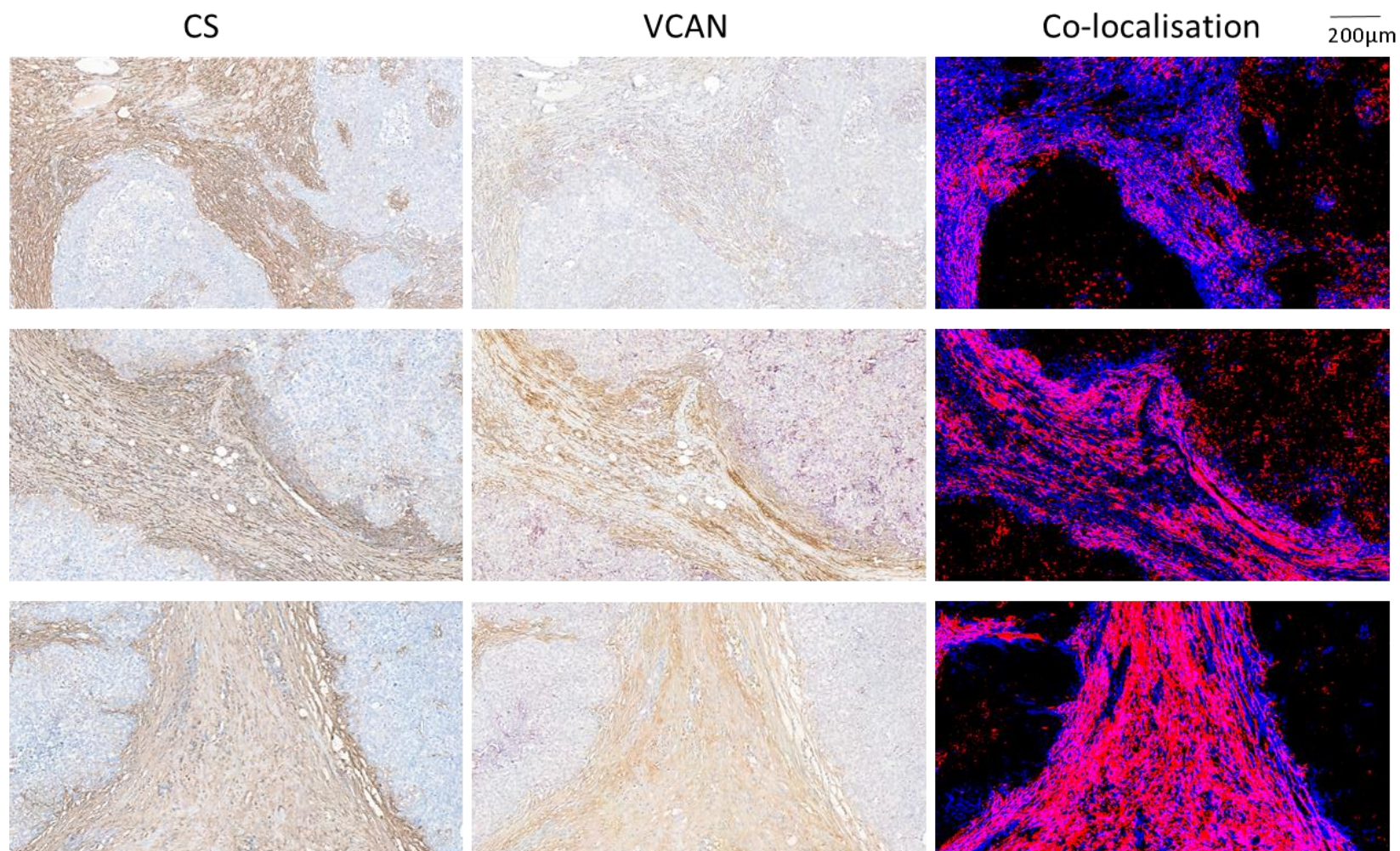
To deduce whether the levels of CS are linked to the association seen with VCAN and immune exclusion, the co-localisation and correlations between CS and VCAN were explored. Co-localisation analysis was completed by overlaying the staining for CS and VCAN. This was completed in different areas of excluded tissues and inflamed tissues. From the overlay it was interesting to observe that there were high levels of variation between the level of co-localisation. In Figure 4.2 it can be observed, that there are excluded tissues where there is little to no co-localisation and there are tissues where there is co-localisation (purple) but higher levels of CS (blue) or higher levels of VCAN (red). In the inflamed tissues a range was also observed however in most cases there was a greater amount of CS compared to VCAN (Figure 4.3). There was no correlation between VCAN and CS in the excluded tissues in any region (Figure 4.4). Significant correlations were observed in the inner invasive and stromal regions. Poor levels of correlation and variable co-localisation in the excluded tissues indicates that the expression of CS on VCAN may be irregular compared to a more consistent level in inflamed tissues. This could also be associated with other CS proteoglycans being present.





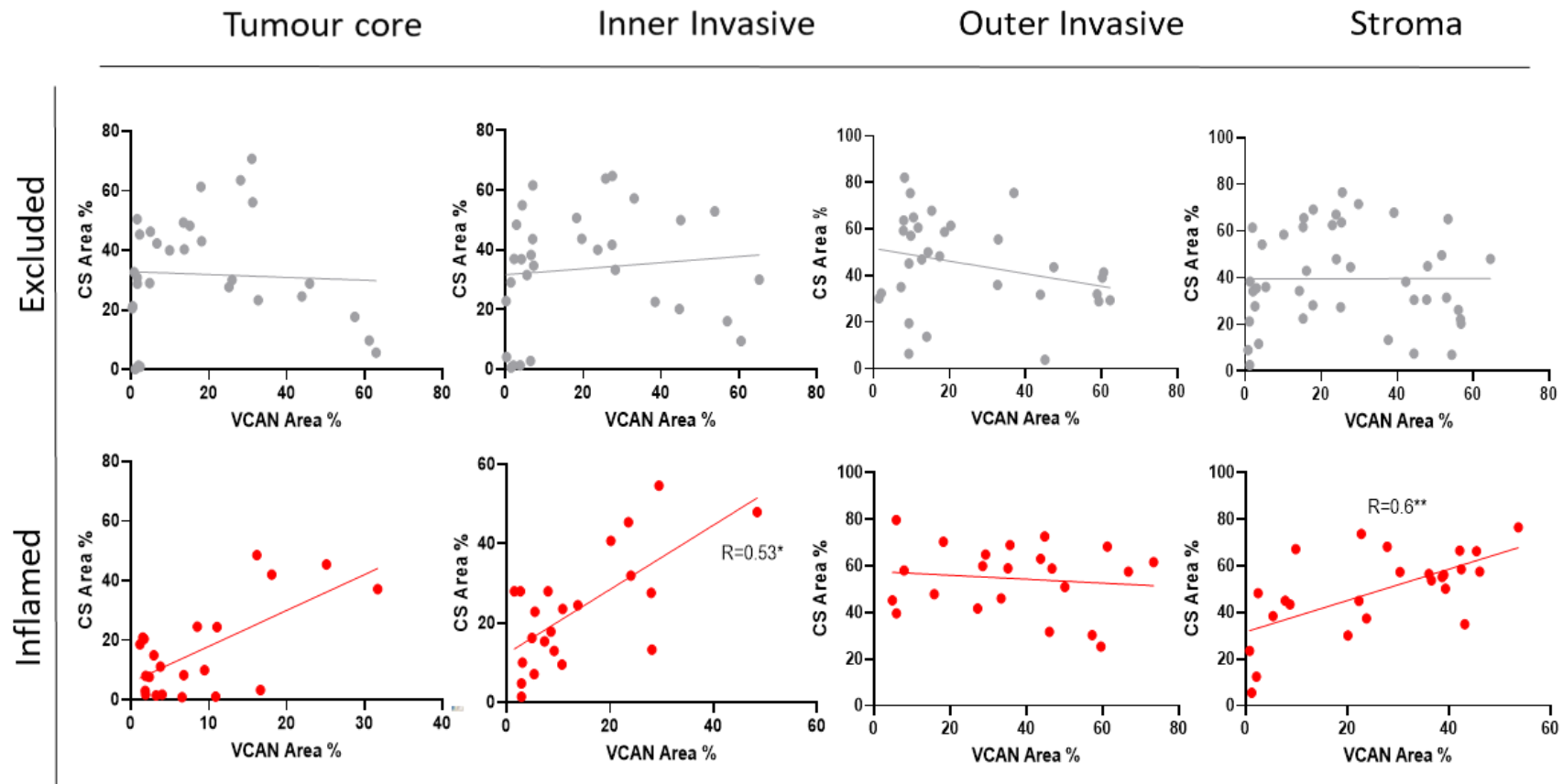
**Figure 4.2. Co-localisation of CS and VCAN in excluded tissues.** IHC was completed for CS and VCAN on consecutive sections. The same areas were taken and deconvoluted for the DAB stain. Stains were then overlaid to look at co-localisation – Blue (CS), Red (VCAN), Purple (co-localisation).





**Figure 4.3. Co-localisation of CS and VCAN in inflamed tissues.** IHC was completed for CS and VCAN on consecutive sections. The same areas were taken and deconvoluted for the DAB stain. Stains were then overlaid to look at co-localisation – Blue (CS), Red (VCAN), Purple (co-localisation).



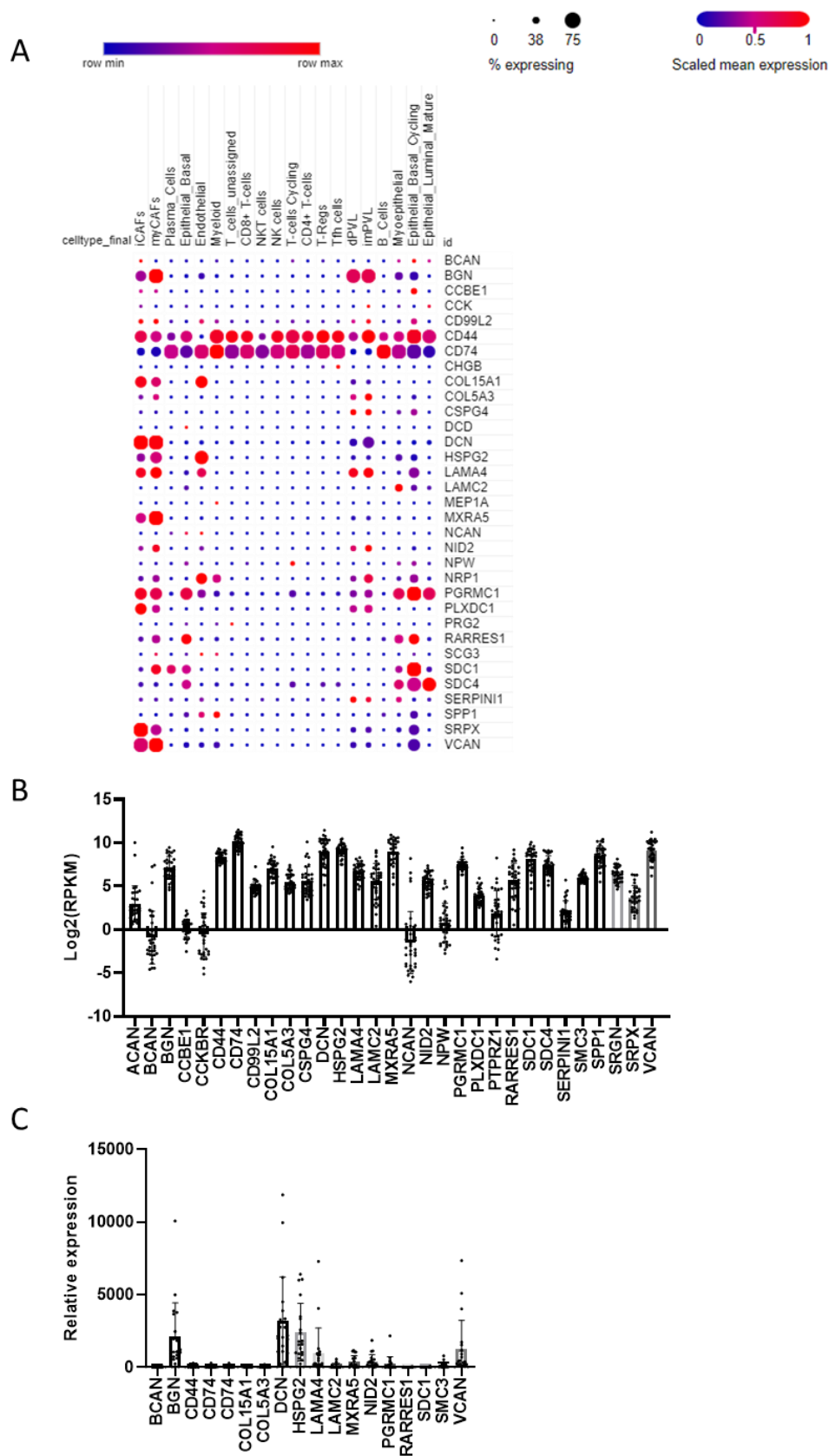


**Figure 4.4. Correlation between VCAN and CS in the different regions of excluded and inflamed tissues.** IHC was completed for CS and VCAN. Areas for each region were outlined for each stain. 9 excluded tissues and 5 inflamed tissues were selected and 5 area were analysed within each tissue. Each dot represents one area. Spearman Rank Correlation.  $*$ = $P < 0.05$ ,  $**$ = $P < 0.01$ .

### 4.3. CS proteoglycans in TNBC

To explore CSPGs that may also correlate with inflamed and excluded phenotypes I began by performing a literature search to identify CSPGs highly expressed in TNBC. CSPGs expression levels were first explored at a single cell level, using the single cell portal with data from Wu et al <sup>243</sup>. The levels of gene expression for each CSPG were detected in different cell types in TNBC. CD44 and CD74 were found to be expressed by nearly all cells. Aggrecan has the highest number of CS chains, however there was no *ACAN* expression seen in any cell type, leaving *VCAN* as the proteoglycan with the most number of CS chains in TNBC. Fibroblasts (iCAFs and myCAFs) expressed the most common proteoglycans – *BGN*, *DCN*, *LAMA4*, *MXRA5* and *VCAN*. Epithelial cells only expressed a selected group of genes. *HSPG2* was found to be mostly expressed by endothelial cells (Figure 4.5A).

The expression of these genes were then investigated in datasets from TNBC human tissues analysed by RNAseq and proteomics. A majority of the genes were expressed at similar levels between the different patients. No key proteoglycans could be observed from the data, with most genes being expressed at comparable levels (Figure 4.5B). From the proteomics analysis not all CSPGs could be analysed as the levels were too low. The highest expressed proteoglycans were *BGN*, *DCN*, *HSPG2* and *VCAN* (Figure 4.5C). As the interest was to look at proteins within the stroma and around the tumour epithelium, *HSPG2* was not seen as a protein of interest. *BGN*, *DCN* and *VCAN* were further explored.

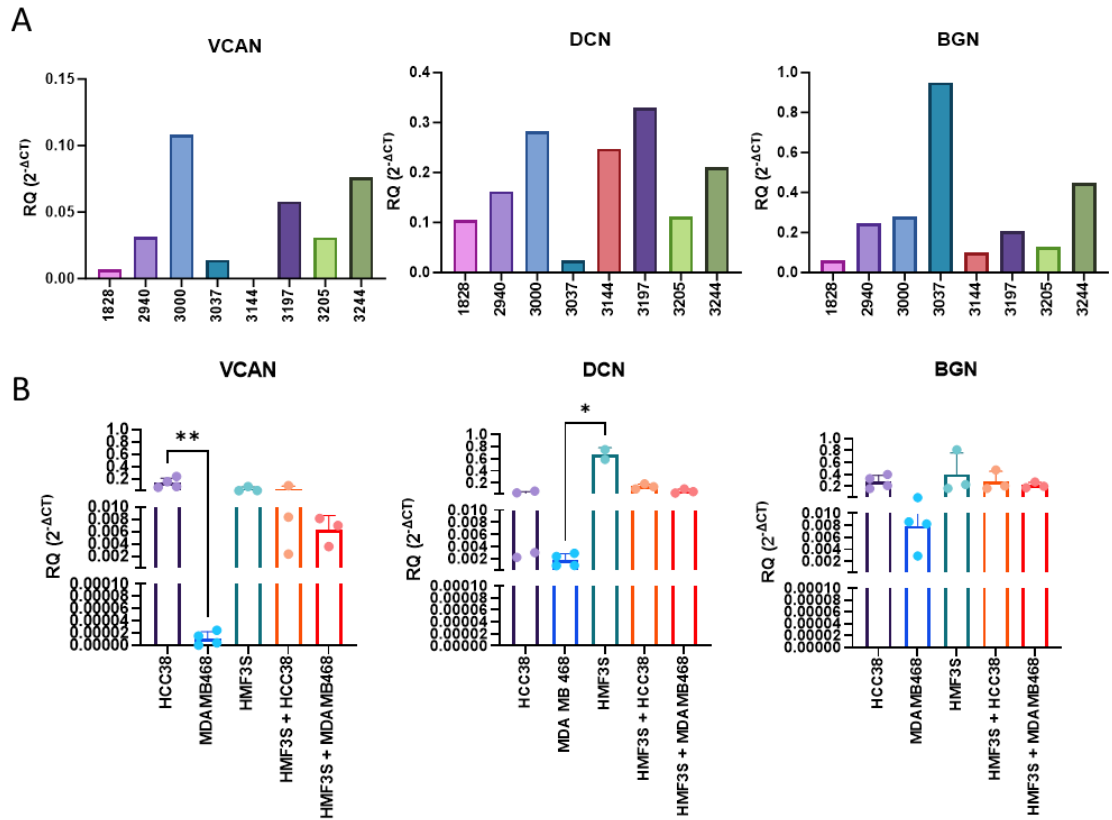


**Figure 4.5. Expression of CS proteoglycans in TNBC.** A) Single cell RNAseq data from online dataset of human TNBC tissues<sup>218</sup>. B) Bulk RNAseq of 43 human TNBC tissues. C) Proteomics of 20 human TNBC tissues.

To further investigate these three CSPGs and confirm what was observed from the RNAseq data, I analysed their gene expression level in tissues and cell lines of TNBC using qRT-PCR. qRT-PCR was used to identify the mRNA level of all isoforms of *VCAN*, *DCN* and *BGN*. Primer sequences were selected to detect all isoforms of each proteoglycan. I first studied the levels in frozen TNBC tissues, for comparison with cell lines. Differential expression was observed for each proteoglycan in the different tissues. No trend was found between the proteoglycans, indicating there is no connection between their expression (Figure 4.6A).

TNBC cell lines HCC38 and MDA MB468 and mammary fibroblast cell line HMF3S, were compared for their expression. The co-culture of the tumour cells with the fibroblast cell line was also tested to see whether cell-cell contact can impact the expression of the proteoglycans, as observed with *VCAN* isoform expression in the previous chapter (Figure 3.20). Comparing TNBC cell lines MDA MB468 and HCC38, MDA MB468 had significantly lower expression of *VCAN*, and also lower, but not significantly so, levels of *BGN* and *DCN*. HMF3S fibroblasts expressed the highest levels of *DCN* with comparable levels of *VCAN* and *BGN* with HCC38 (Figure 4.6B).

The co-culture of TNBC cells lines with HMF3S did not significantly change the level of proteoglycan expression. A lower level of *VCAN* was observed with the HMF3S-MDA MB468 co-cultures but this may be due to MDA MB468 cells expressing low levels of *VCAN* (Figure 4.6B). Overall, from this analysis, cell-cell contact between tumour cells and fibroblasts does not appear to inhibit or stimulate proteoglycan expression. This contrasts with what was observed in the previous chapter, indicating that the interactions may be isoform specific to *VCAN*.

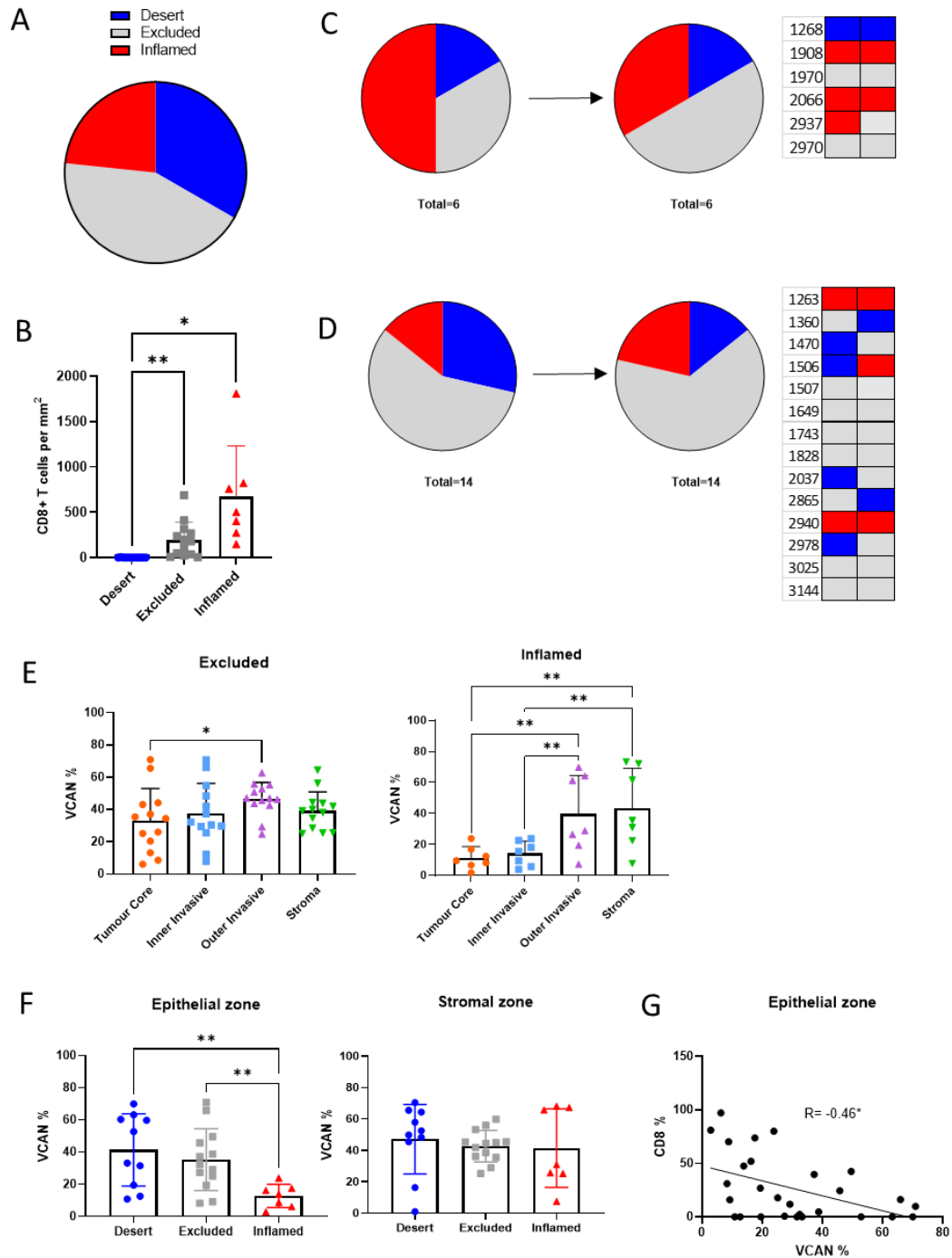


**Figure 4.6. CS proteoglycan expression in tissues and cells in mono-culture and co-culture,** RNA was extracted from A) 8 frozen TNBC tissues and B) TNBC cell lines in monoculture and co-culture with HMF3S. qRT-PCR was carried out for VCAN, DCN and BGN. The levels were normalised to RPS13 expression. N=4. Kruskal-Wallis test. \*= $P < 0.05$ , \*\*= $P < 0.01$

#### 4.4. Association of BGN and DCN with tissue immune phenotypes.

Having identified BGN and DCN also being highly expressed in tumours, the levels in the different TIPs were examined to see if they associated in a similar manner seen with VCAN.

Due to the unavailability of slides from the previous subset of TNBC tissues, a new subset of 30 TNBC tissue blocks were analysed for the TIP and compared to VCAN. 20 of the blocks in the new subset were from the same patient with tissue sections received from either the same block or a different block. Following TIP analysis, the percentage of inflamed tissues was similar to the level in the previous analysis, but there was an increase in desert tissues and a reduction in the percentage of excluded (Figure 4.7A). No significant difference was observed between the number of CD8+ T cells in the excluded and inflamed tissues as seen in previous analysis (Figure 4.7B). Comparisons of the TIPs between the patients showed that when the tissue section came from the same block the majority had the same phenotype (83%) (Figure 4.7C) whilst when the section came from a different tissue block from the same patient nearly half the samples (42%) (Figure 4.7D) had a change in phenotype. This may indicate a significant level of heterogeneity within TIP throughout a tumour, and indicates when comparing distant tumour nodules the phenotype may vary. As seen before, the spread of VCAN in both excluded and inflamed tissues showed a significant increase in the VCAN levels in tumour core compared to the outer invasive region (Figure 4.7E), with inflamed tissues also showing a significant increase in VCAN levels between the tumour core and stroma. The percentage area of VCAN in the epithelial zone was significantly higher in the excluded and desert compared to the inflamed. No difference was seen in the stromal zone (Figure 4.7F). As before a significant negative correlation was observed between VCAN and percentage of CD8+ T-cells in the epithelial zone (Figure 4.7G).



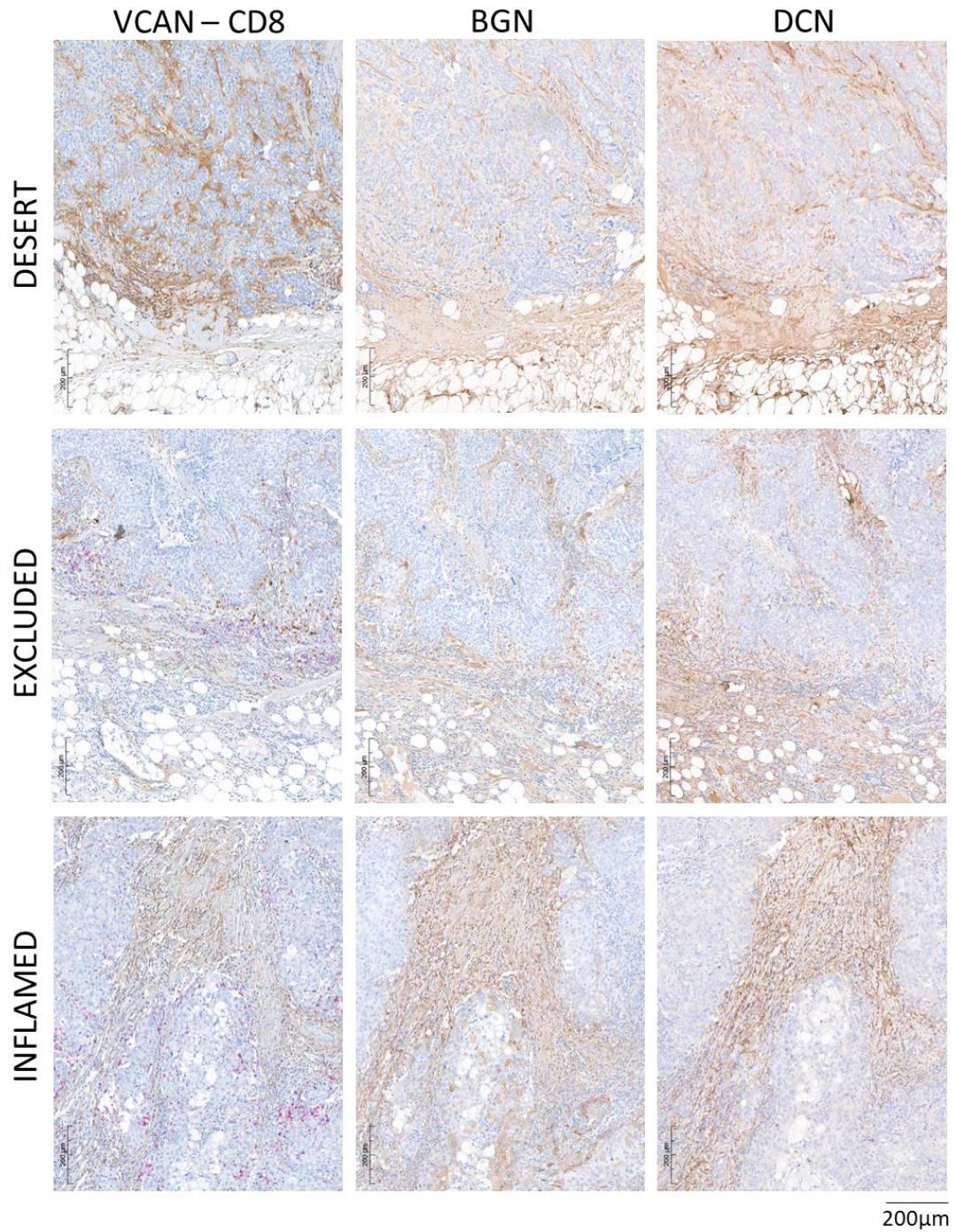
**Figure 4.7. Tissue immune phenotype analysis in second subset of tissues.** A) Pie chart showing CD8 tissue immune phenotypes for the tissues. B) Number of CD8+ cells in each phenotype. ANOVA. New subset compared to old tissues where either the C) same blocks or D) same patient but different block. E) Levels of VCAN were compared in the different regions in excluded and inflamed tissues. RM-ANOVA. F) VCAN was compared between the phenotypes in the different zones. Welch ANOVA. G) Correlation of VCAN% and percentage of CD8 cells in the epithelial zone. Spearman Rank Correlation. \*= $P < 0.05$ , \*\*= $P < 0.01$ .

Following the identification of the different phenotypes and confirming similar associations with VCAN, the slides were analysed for BGN and DCN. The same areas in the tissue were selected for the analysis to allow more reliable comparisons. It was initially seen that the area of BGN accumulation was quite uniform across all tissues independent of the phenotype (Figure 4.8). This observation was confirmed with QuPath analysis showing that the levels of BGN were similar across the phenotypes in the whole area and also in both stromal and epithelial zones. A slight increase was observed in the epithelial zone of excluded tissues compared to the inflamed however this was not significant. The levels of BGN in the stromal zone were equal for all phenotypes (Figure 4.9A). Comparison of the spread of BGN across the regions independent of phenotype identified that the regions in the stromal zone were significantly different to regions of the epithelial zone. Both inflamed and excluded tissues mirrored these differences showing that the pattern of BGN expression is not phenotype specific (Figure 4.9B). BGN was significantly positively correlated to VCAN within the epithelial zone but no correlation was observed in the stromal zone (Figure 4.9C). Levels of BGN in the epithelium did not associate with the percentage CD8+ T cell (Figure 4.9D).

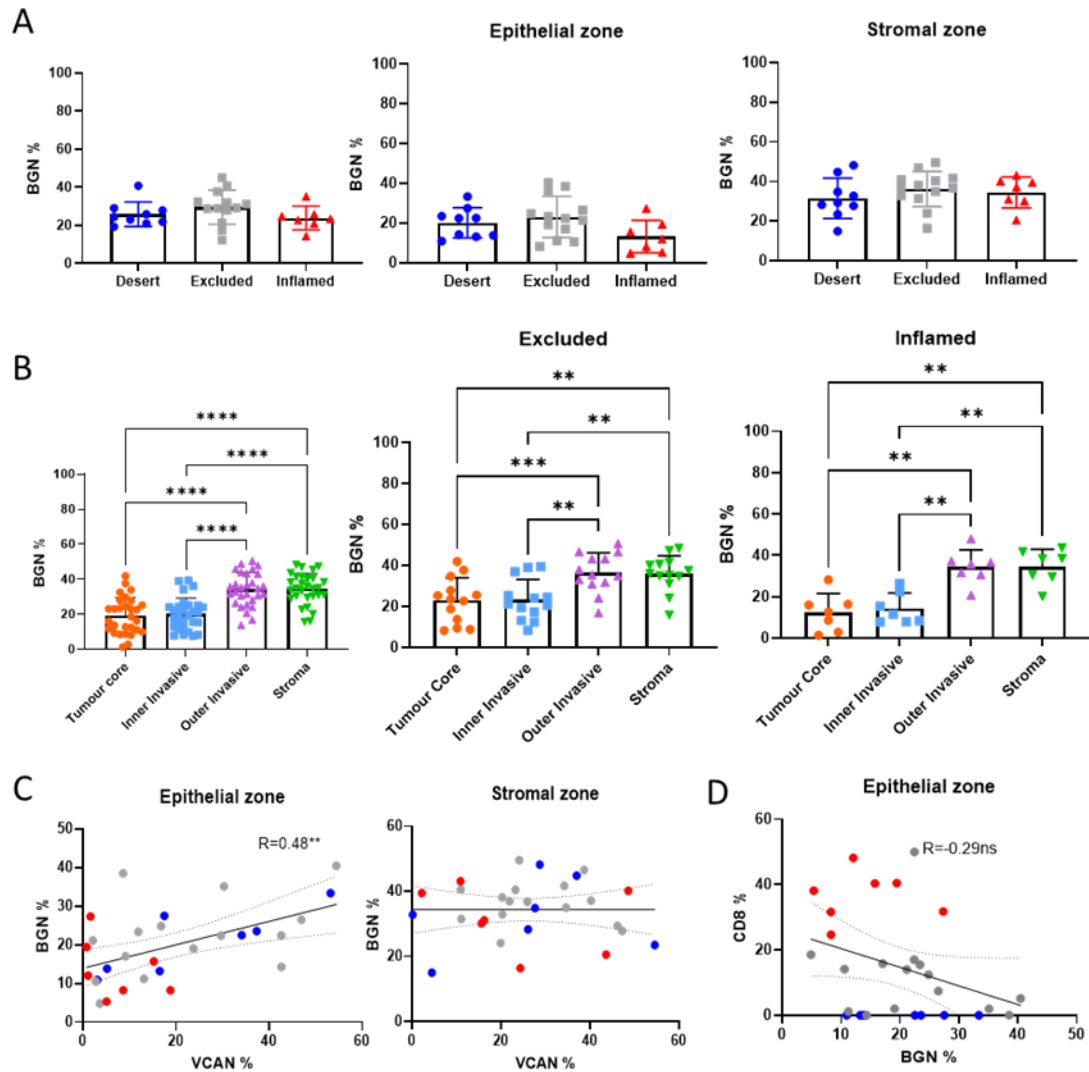
This pattern was also seen with DCN, where the expression was quite uniform across the whole tissue section (Figure 4.8). This confirmed the high levels observed in the proteomics analysis. The same analysis was completed for DCN as with BGN. Likewise with BGN the levels of expression between the phenotypes in the whole area as well as different zones were not significantly different (Figure 4.10A). The levels of DCN in the epithelial zone regions were significantly lower than the regions in the stromal zone (Figure 4.10B). No correlations were observed between VCAN and DCN in any region (Figure 4.10C). There was also no association with the percentage of CD8+ T cells (Figure 4.10D).

In conclusion these data indicate BGN and DCN are not correlated with TIP, and therefore less likely than VCAN to have an active role in immune infiltration. This also indicates that the poor correlation between CS and VCAN observed in the excluded tissues may not be due to other proteoglycans.

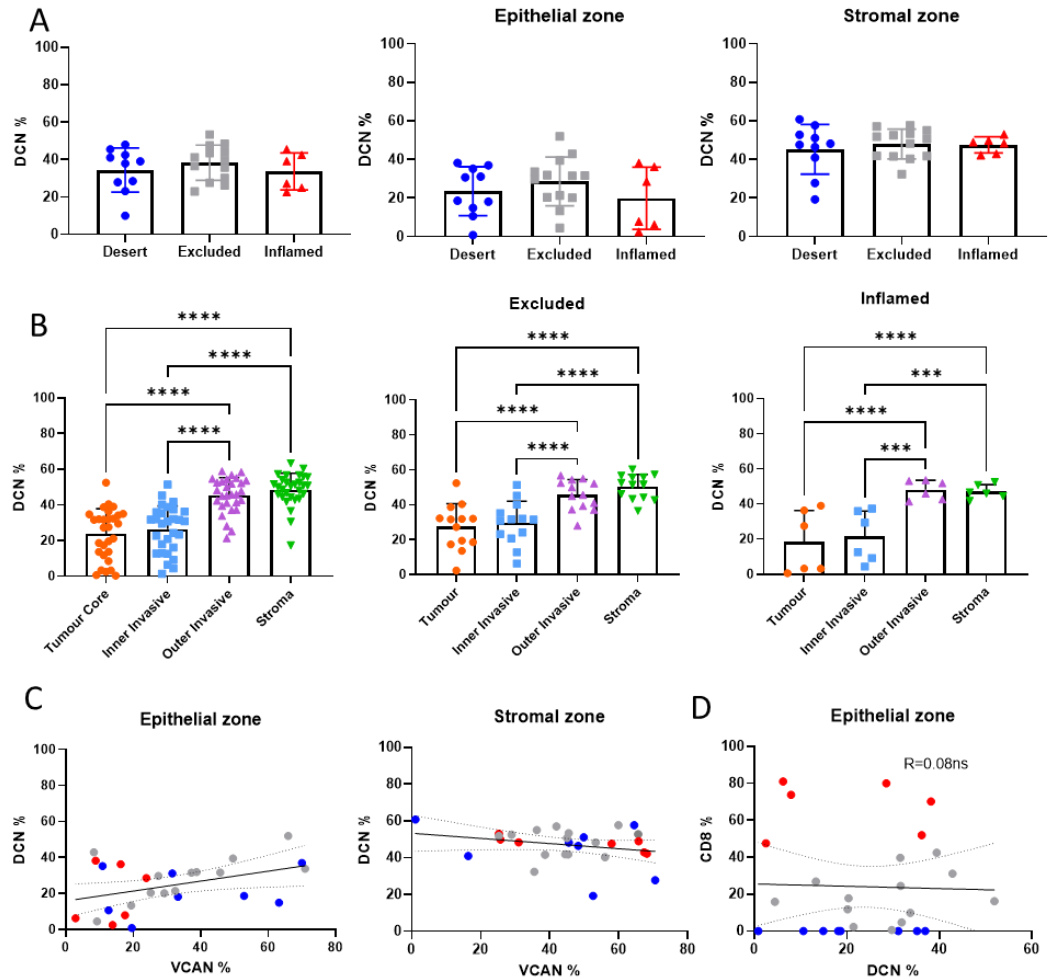




**Figure 4.8. VCAN, BGN and DCN in the different TIPs.** IHC carried out on consecutive slides for the different markers. Dual IHC completed for VCAN and CD8. VCAN, BGN and DCN in brown (DAB), CD8 in pink.



**Figure 4.9. BGN is not associated with TIP.** IHC was completed for BGN and the same areas and regions were analysed as the VCAN-CD8 stains for the immune phenotype. ANOVA A) Comparison of BGN in the different phenotype in the whole tissue and different areas. B) Comparison of BGN across regions in all phenotypes, excluded and inflamed tissues. RM-ANOVA. C) Correlation of VCAN and BGN in the different zones. D) Correlation of percentage of BGN and percentage of CD8 cells in the epithelial zone. Spearman Rank Correlation.  $**=P<0.01$ ,  $***=P<0.005$ ,  $****=P<0.001$ .

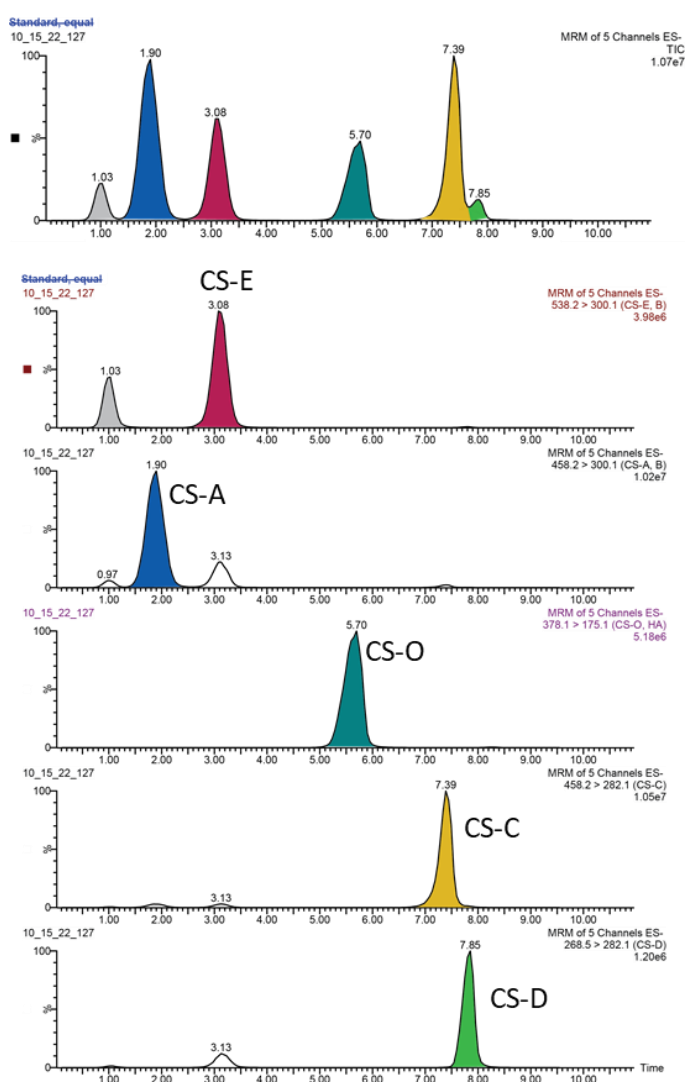


**Figure 4.10. DCN expression is not associated with phenotype.** IHC was completed for DCN and the same areas and regions were analysed as the VCAN-CD8 stains for the immune phenotype. ANOVA A) Comparison of DCN in the different phenotype in the whole tissue and different areas. B) Comparison of DCN across regions in all phenotypes, excluded and inflamed tissues. RM-ANOVA. C) Correlation of VCAN and DCN in the different zones. D) Correlation of percentage of DCN and percentage of CD8 cells in the epithelial zone. Spearman Rank Correlation. \*\*\*= $P < 0.005$ , \*\*\*\*= $P < 0.001$ .

## 4.5. Effect of CS isomer on immune phenotype

CS can be present as multiple isomers based on the sulphation patterns, which in turn alters their biological function. A change in CS isomer pattern may help explain the difference in T-cell location seen between inflamed vs excluded tissues. The CS-O isomer contains no sulphation whilst the mono-sulphated CS-A and CS-C are sulphated at the 4<sup>th</sup> and 6<sup>th</sup> carbon of the GalNAc respectively. The CS-D and CS-E isomers contain 2 sulphations. The CS-D contains a sulphation at the 2<sup>nd</sup> carbon of GlcA and on the 6<sup>th</sup> carbon of GalNAc. The CS-E isomer is sulphated at both the 4<sup>th</sup> and 6<sup>th</sup> carbon of the GalNAc sugar <sup>242</sup>.

Due to the different sulphation locations and the number of sulphations the isomers can be differentiated using mass spectrometry as the mass-charge ratio varies as shown in Figure 4.11.



**Figure 4.11 Chromatogram of CS-MS.** Standards of each CS isomer were run to detect where the peaks are present.

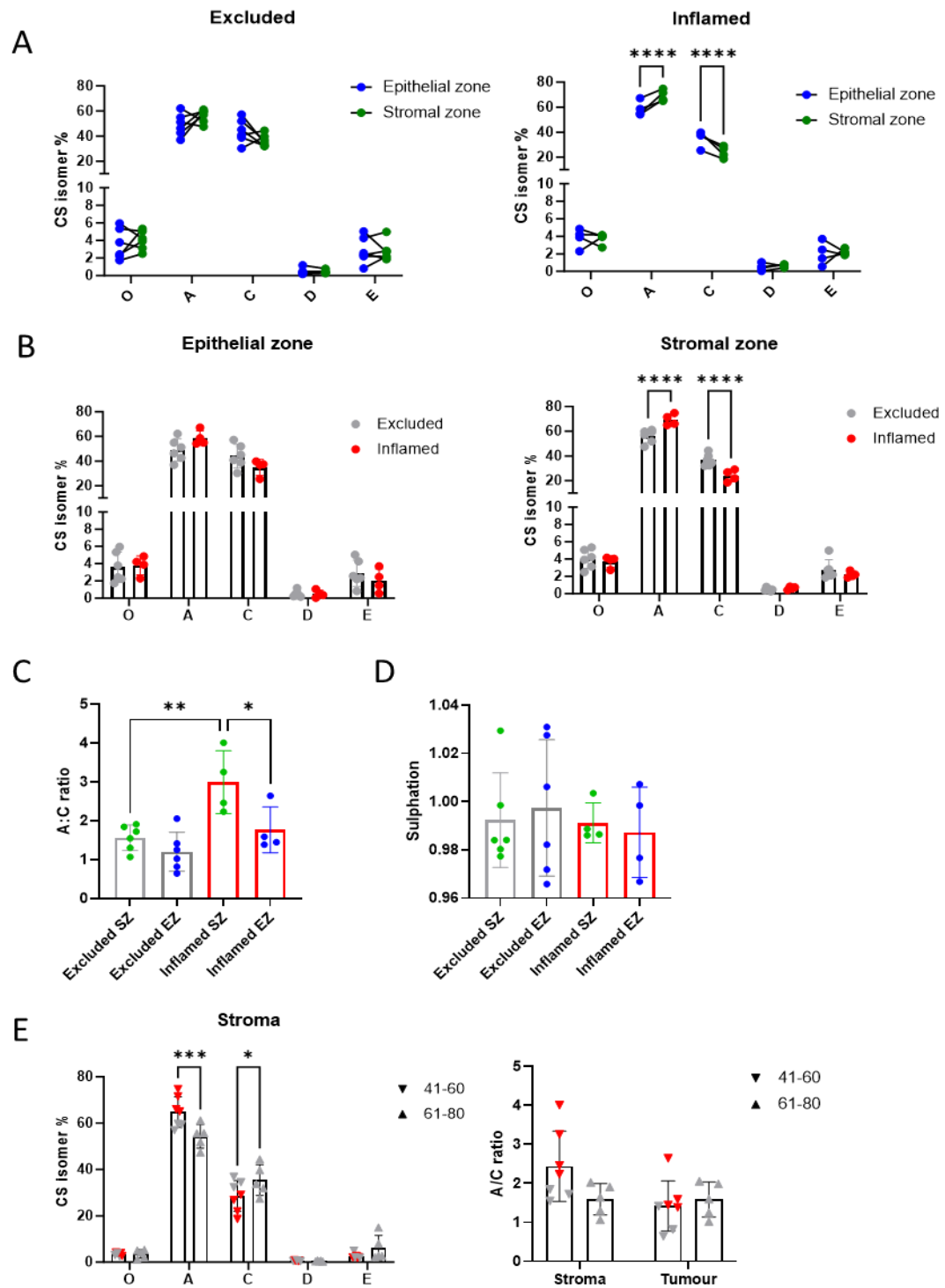
The proportion of different isomers in the epithelial zone and stromal zone were compared across the different phenotypes. 8 excluded and 4 inflamed tissues were taken for analysis. In collaboration with Dr Kim Alonge (University of Seattle, Washington), we optimised a protocol for the detection of CS isomers in tissues and free-flow protein by MS. For the extraction of CS from tissues, 2 sections of tissues were taken for each tissue. As the tissues were sectioned onto slides for IHC and at a thickness of 4µm they could not be segregated for tumour and stroma areas using laser microdissection. The epithelial zone and stromal zone were therefore marked out with a PAP pen. The epithelial zone was defined using Pan-CK staining. CSase was pipetted to these outlined regions to cleave the CS. The solution was then collected and run by MS. The levels of the different isomers were determined by the area of the peaks. Following the run of the samples, 2 excluded tissues did not show detectable levels of CS-D to be able to confidently analyse the samples so they were excluded from the analysis. From the data, the CS-A and CS-C isomers were the most abundant overall. In the excluded tissues there was no significant difference between the isomers in the epithelial zone and stromal zone. When comparing the CS-A and CS-C isomers in the excluded tissues, it was observed that in 4 of the 6 tissues there was an increase in CS-A and a reduction in CS-C from the epithelial zone to the stromal zone, but in 2 of the tissues there was a reduction in CS-A and increase in CS-C. Significant differences were observed in the inflamed tissues for the CS-A and CS-C isomers with an increase in CS-C and a reduction in CS-A in the stromal zone. No changes were observed in the other isomers (Figure 4.12A).

Next the CS isomers were compared between phenotypes. No differences were observed in the epithelial zone. In the stromal zone, there was a significantly higher level of CS-A (56% in excluded and 69% in inflamed) and a significantly lower level of CS-C (37% in excluded and 24% in inflamed) in the inflamed tissues compared to the excluded (Figure 4.12B). These differences in the levels of CS-A and CS-C led to a significantly higher A:C ratio in the stromal zone of inflamed tissues ( $r = 3$ ) compared to the stromal zone of excluded tissues ( $r = 1.6$ ) (Figure 4.12C). The level of sulphation indicates the average number of sulphate ions per CS disaccharide. A wide range was observed between the samples with no associations with phenotype or area (Figure 4.12D).

Age is a reported factor affecting the ratio of CS-A and CS-C with a reduction in CS-C seen with older age <sup>244</sup>. To explore how age was affecting the isomers observed in tissues, the samples were grouped based on age. Two groups were formed: 41-

60 and 61-80. From the age groups a significant difference was observed in the CS-A and CS-C isomers where A decreased with age and C increased with age. This was opposing to what had been observed in healthy human tissue<sup>244</sup>. Identifying the TIPs in these age groups showed that the inflamed tissues were all within the 41-60 age group and the 61-80 age group was occupied by only excluded tissues. When comparing the excluded tissues in both age groups there was a similar level of CS-A and CS-C isomers as well as the CS-C ratio as indicated by the points in grey (Figure 4.12E). These data indicate inflamed patients tend to have a lower age and that the difference seen in CS isomers in TNBC tissues have a greater association with phenotype than age.

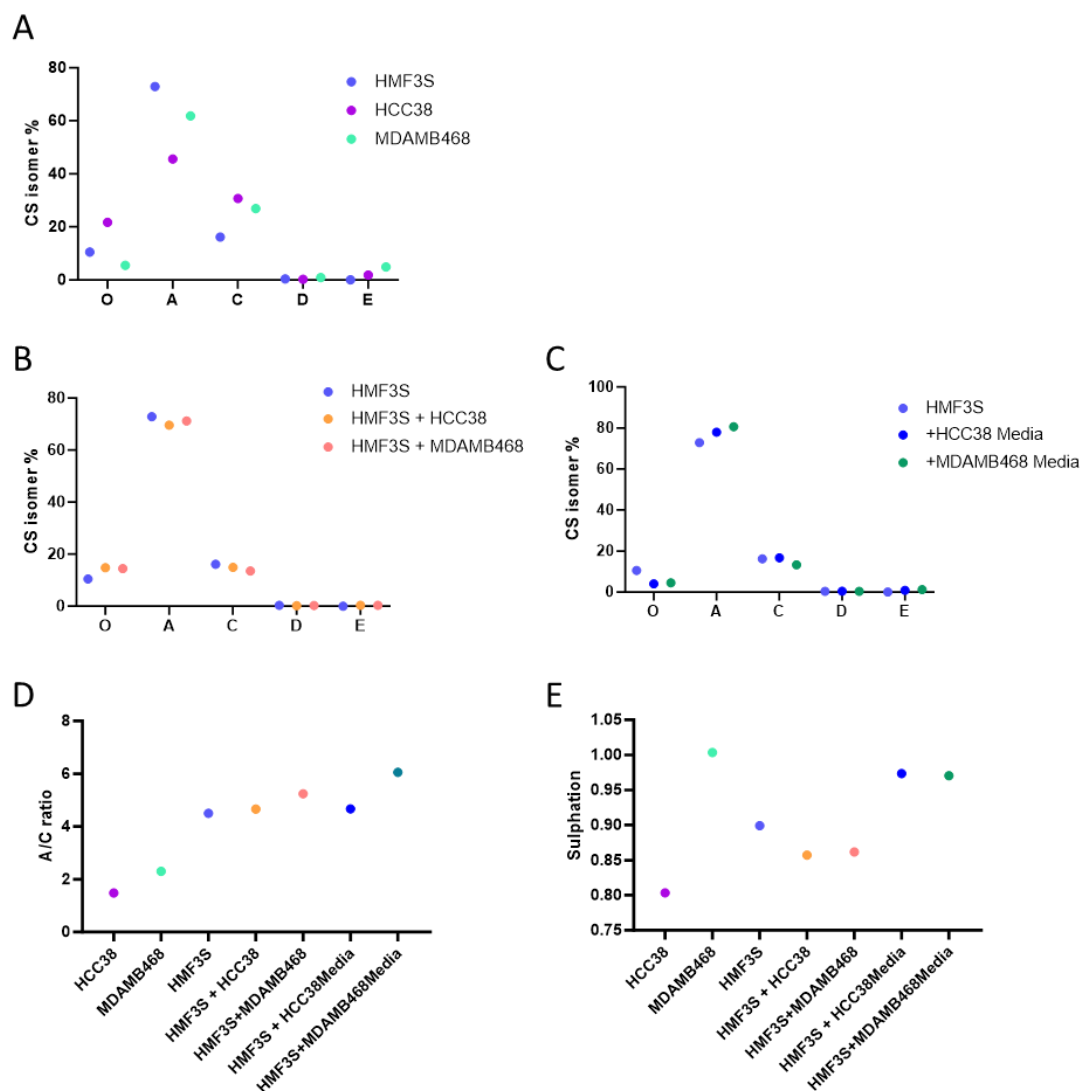




**Figure 4.12. CS isomers A and C are differentially expressed in excluded and inflamed tissues.** CS was extracted from 6 excluded tissues and 4 inflamed tissues in the epithelial and stromal zones. The CS isomers were detected using LCMS/MS. A) Changes in the isomer levels in the different regions in the different phenotypes. Multiple Mann Whitney T test. B) Comparison of isomers between phenotypes in the different regions. Multiple Mann Whitney T test. C) A:C ratio within the stromal zone (SZ) and epithelial zone (EZ) of the excluded and inflamed phenotypes. ANOVA. D) Number of sulphates per CS isomer within the stromal zone (SZ) and epithelial zone (EZ) of the excluded and inflamed phenotypes. ANOVA. E) Proteoglycans in the different regions and phenotypes. F) CS isomers in stromal areas and A:C ratio in different age groups. Tissues shown in grey were identified as excluded and those in red were identified as inflamed. \* = $P < 0.05$ , \*\* = $P < 0.01$ , \*\*\* = $P < 0.005$ , \*\*\*\* = $P < 0.001$ .

The levels of CS isomers were then also explored in cell lines to see if the patterns observed between excluded and inflamed tissue may be linked to a specific cell of origin. HMF3S fibroblasts, and HCC38 & MDAMB468 cell lines were analysed for CS isomer composition from the secretions in the media. CS-O levels varied with HCC38 cells expressing the most. Both tumour cell lines had less CS-A and more CS-C than the fibroblast cell line. CS-E was also higher in the tumour cell lines (Figure 4.13A). The effect of tumour cells towards CS expression by fibroblasts was explored by a co-culture and conditioned media treatment. From the co-culture treatment, very little difference was observed to the level of the CS isomers (Figure 4.13B). This was also the case with the conditioned media (Figure 4.13C). From the A/C ratio of the samples, the tumour cell lines had a much lower ratio. The ratio in the co-cultures did not vary from HMF3S alone, similarly with HCC38 conditioned media treatment. Surprisingly an increase was seen from the MDAMB468 conditioned media treatment (Figure 4.13D). Differences were observed with the level of sulphation where opposing levels were seen from the tumour cell lines. Again co-culturing of the cells had little effect. The conditioned media treatment led to an increase in sulphation irrespective of the tumour cell line (Figure 4.13E). When comparing this data to what has been observed within the tissues, the low A/C ratio with the tumour cell lines matched the low ratio observed in the epithelial zone of both phenotypes. The effects of the co-culture and the media treatment could not be directly linked to the differences observed in the stromal zone of the excluded and inflamed tissues. To confirm the effects of the tumour cell conditioned media towards sulphation levels, repeats will be need to be completed with additions of other cell lines.





**Figure 4.13. CS isomers in cell lines.** CS was extracted from the media of cell cultures. Isomers were detected using LCMS/MS. A) Isomers in monocultures of cells after 7 days. B) Co-cultures of HMF3S with tumour cells 4:1 after 7 days of culture. C) Treatment of HMF3S with tumour conditioned media at a ratio of 1:1. D) A/C ratio of monoculture, co-culture and conditioned media treatment. E) Sulphation per CS disaccharide in monoculture, co-culture and conditioned media treatment samples.

## 4.6. Discussion

The expression of CS is an important factor in the function of VCAN. However as CS is not solely expressed by VCAN it was important to consider how other CS expressing proteoglycans are associated with immune exclusion when exploring the role of CS. There was no significant difference between the level of CS in the excluded and inflamed tissues. As seen with VCAN, the CS levels in inflamed tissues were concentrated in the stromal zone, whilst in excluded tissues there was only a higher level observed around the tumour periphery. From the co-localisation and correlations of VCAN and CS, it became apparent that other CS presenting proteoglycans may be present within the TME.

To explore what these proteoglycans may be, a literature search was carried out to identify which proteins express CS<sup>240,241,245</sup>. Over 30 molecules were found to express CS. Fibroblasts and tumour cells were the predominant cell types expressing these genes. Despite being expressed highly at RNA level, the translation to protein expression is not seen for all proteins found with the exception of *BGN*, *DCN*, *HSPG2* and *VCAN*. *HSPG2* known as perlecan often expresses heparan sulphate, however there have been indications of CS and HS expression<sup>246</sup>. As the focus of the analysis was to look at CS expression around the tumour epithelium, *HSPG2* was not further investigated as a majority of expression was observed by endothelial cells in the single cell RNAseq. From the remaining 3 proteoglycans, similar roles were observed in inflammation where they utilise TLR2 and TLR4 as key receptors for inducing a reaction. Despite the high levels of DCN observed in the proteomics and RNA-seq data, a downregulation of its expression was observed in the ECM in diseased tissue compared to non-diseased tissues<sup>247,248</sup>. This is possibly due to its role as a tumour suppressor through the Met receptor pathway<sup>248</sup>.

To first confirm what was seen from the RNAseq data, the RNA expression of VCAN, DCN and BGN was explored in tissues and cell lines. No link was observed between the levels of each proteoglycan. Previous studies looking at these three genes identified TGF- $\beta$ 1 to increase VCAN and BGN and reduce DCN however there was no correlation observed between the genes<sup>249</sup>. The effect of FGF on the expression of these genes was also explored and it was found that FGF could increase DCN expression, downregulate BGN and have no effect on VCAN<sup>250,251</sup>. The different effects of these growth factors on their expression could justify not observing any associations between the genes in tissues. Different cell lines were then tested. Tumour cell lines, HCC38, showed high levels of expression of VCAN

isoforms whilst very little to no expression of some isoforms in MDA MB 468 cells. The MDA MB 468 cell line was derived from a metastatic pleural effusion whilst the HCC38 cell line was derived from the primary tumour. This may have affected the level of ECM expression by these cell lines. The significant difference in the VCAN levels between the 2 tumour cell lines allows the use of these cells for functional models to test for the effect of VCAN as the expression of the other 2 proteoglycans were similar.

A co-culture model was used to test whether tumour cells would affect the expression of these proteoglycans by fibroblasts. The limitation of testing the gene expression through co-culture is that the RNA from the tumour cells will also be extracted, possibly affecting the overall level of the genes. The expression of VCAN, BGN and DCN by HMF3S fibroblasts was not altered by the co-culture with HCC38 and MDAMB468 despite the differences in the expression by the tumour cells in mono-culture. To understand whether tumour cells have little effect or if the ratios selected in the co-culture do not allow for this difference, other co-culture ratios and treatment with conditioned media could also be completed.

To test the association of the other CSPGs towards immune exclusion, different tissue blocks had to be used for some patients as well as other patient samples, due to the unavailability of these samples. The immune exclusion analysis of these tissues showed that most of the patients showed similar phenotypes. The detection of CD8+ T cells was much lower in the cold tissues, with most of the cold tissues not having any T cells detected. This led to the margin for cold tissues being lower than the previous analysis.

The same analysis was completed for BGN and DCN. BGN was found to be mostly a stromal protein with much less found in the epithelial zone in comparison to VCAN. The pattern for BGN expression was also very similar in tissues, independently of their immune phenotype, showing that BGN is not associated with phenotype by either the amount of expression or the location of its expression. The same was observed for DCN where no association was observed with the different phenotype. This gave the impression that VCAN may have a greater importance in terms of its role towards immune cell exclusion compared to other proteoglycans. DCN only has one CS chain <sup>252</sup> and BGN has 2 chains <sup>253</sup> in comparison to the possible 22 chains <sup>66</sup> on VCAN which may be a factor in the roles they have.

For the CS analysis, the extraction of CS was completed on the slides by isolating areas using a PAP pen. The limitation of this method was the direct tumour-stroma

border as analysed in the image analysis could not be examined. Following the segregation with a PAP pen and CSase treatment, the slides were stained with PanCK to check whether the segregation was done accurately.

Differences in CS isomers were observed between the tumour and stromal regions. This was significant in the inflamed tissues and not within the excluded tissues overall. The stromal zone showed significant differences between the excluded and inflamed tissues. The 2 key isomers found in this analysis were CS-A and CS-C. From the data, the CS-C isomer seemed to be associated with a reduction in T cell infiltration leading to an excluded phenotype.

The binding properties of CS-A (C-4-S) and CS-C (C-6-S) are different due to conformation of CS chain <sup>78</sup>. The CS-C groups have been found to be more accessible to ligands in comparison to the CS-A. IL-8 has been found to bind better to CS-C in comparison to other CS isomers <sup>78,254</sup>. CD44 also binds more avidly to CS-C <sup>254,255</sup>. Therefore, it can be assumed that a greater expression of CS-C in the stroma may be constraining the migration of T cells. The association of CS-C was also explored in colorectal cancer where higher levels in the stroma surrounding the tumour associated with immune exclusion <sup>256</sup>.

The total abundance of CS isomers could not be measured so associations could not be made for the level of CS and the immune phenotype. The levels of VCAN, BGN and DCN did not show to be associated with any isomer. In the epithelial zone, significant differences in VCAN levels between the inflamed and excluded tissues did not make a difference to the CS isomer levels, indicating these may be conserved to the tumour glycocalyx.

To conclude, VCAN is not the only CS proteoglycan that is highly expressed in tumours, however it seems to be the only protein associated with immune exclusion. Cell to cell communication looks to have a role in affecting how VCAN is expressed. The CS isomer results did not explain how the increase in VCAN in the epithelial zone affects T cell exclusion, however it does reveal an association between the type of isoforms in the stroma and the trafficking of T-cells to the tumour epithelium (**Figure 4.14**).

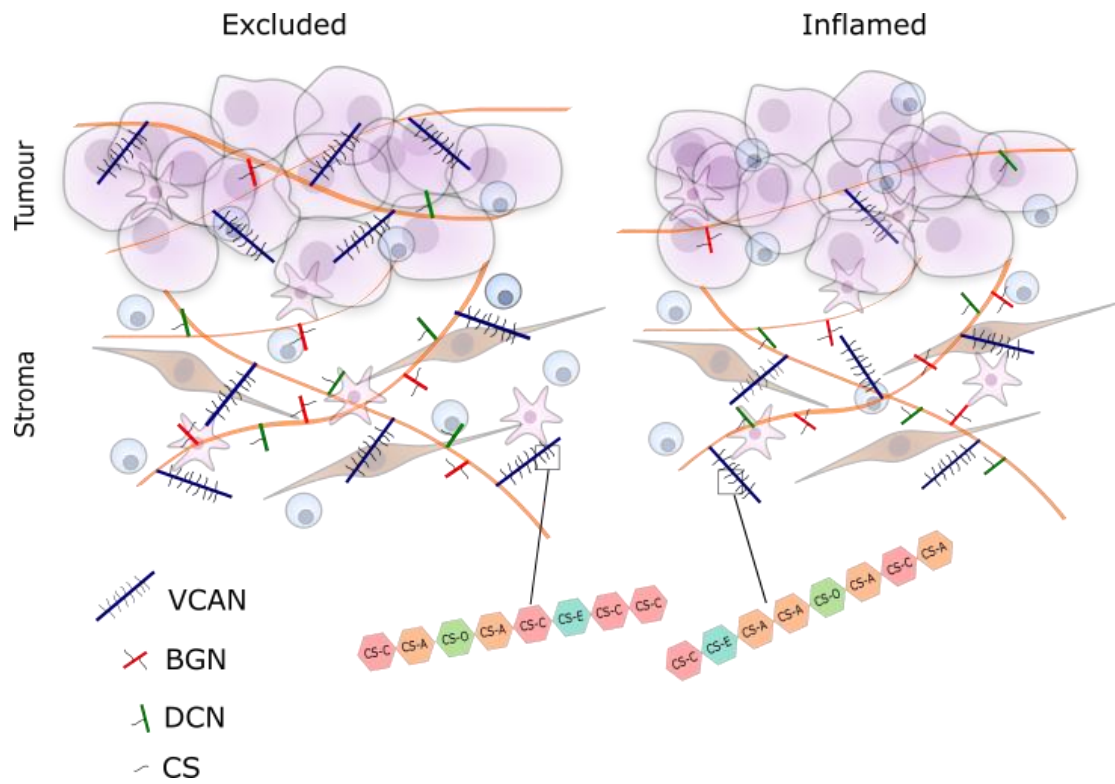


Figure 4.14. Summary of CSPG expression in excluded and inflamed tissues and CS isomer.

## Chapter 5: Enrichment of VCAN from a TNBC and a human mammary fibroblast cell line.

## 5.1. Introduction

In order to further study the structure of VCAN, and carry out functional assays for VCAN's role in T-cell migration, VCAN was isolated from TNBC tumour cells and fibroblasts. Collecting VCAN from these cells provides access to VCAN with post translational modifications of CS that are found in excluded or inflamed tumours as observed in the previous chapter. As shown in Figure 3.19, VCAN expression is the highest in the TNBC cell line HCC38 and human mammary fibroblast cell line HMF3S, therefore these cells were chosen for VCAN isolation.

Previous methods in literature used to isolate VCAN were dependent on the biochemical properties of VCAN. These included the binding of VCAN to HA<sup>257,258</sup> (affinity chromatography), negative charge of VCAN from the CS chains (ion exchange chromatography (IEX)<sup>111,258,259</sup> and density gradient centrifugation<sup>260</sup>) and also the size of VCAN (size exclusion chromatography)<sup>259</sup>.

The binding of VCAN to HA occurs via the G1 domain as detailed in Chapter 1. This domain is present in all isoforms of VCAN and therefore enables the isolation of all isoforms. To isolate VCAN based on the negative charge is dependent on the expression of CS. The number of CS chains can vary across VCAN structures and the V3 isoform contains no GAG domain for CS expression. This means that not all forms of VCAN can be isolated. The limitation of both these techniques is the presence of other proteins that possess these characteristics. Size exclusion chromatography has been used to try to overcome this by taking advantage of the large size of VCAN. However this only accounts for the V0, V1 and possibly V2 isoforms.

In this chapter the enrichment of VCAN using the characteristics mentioned above is presented describing my experiences with the methods and the process of optimisation.

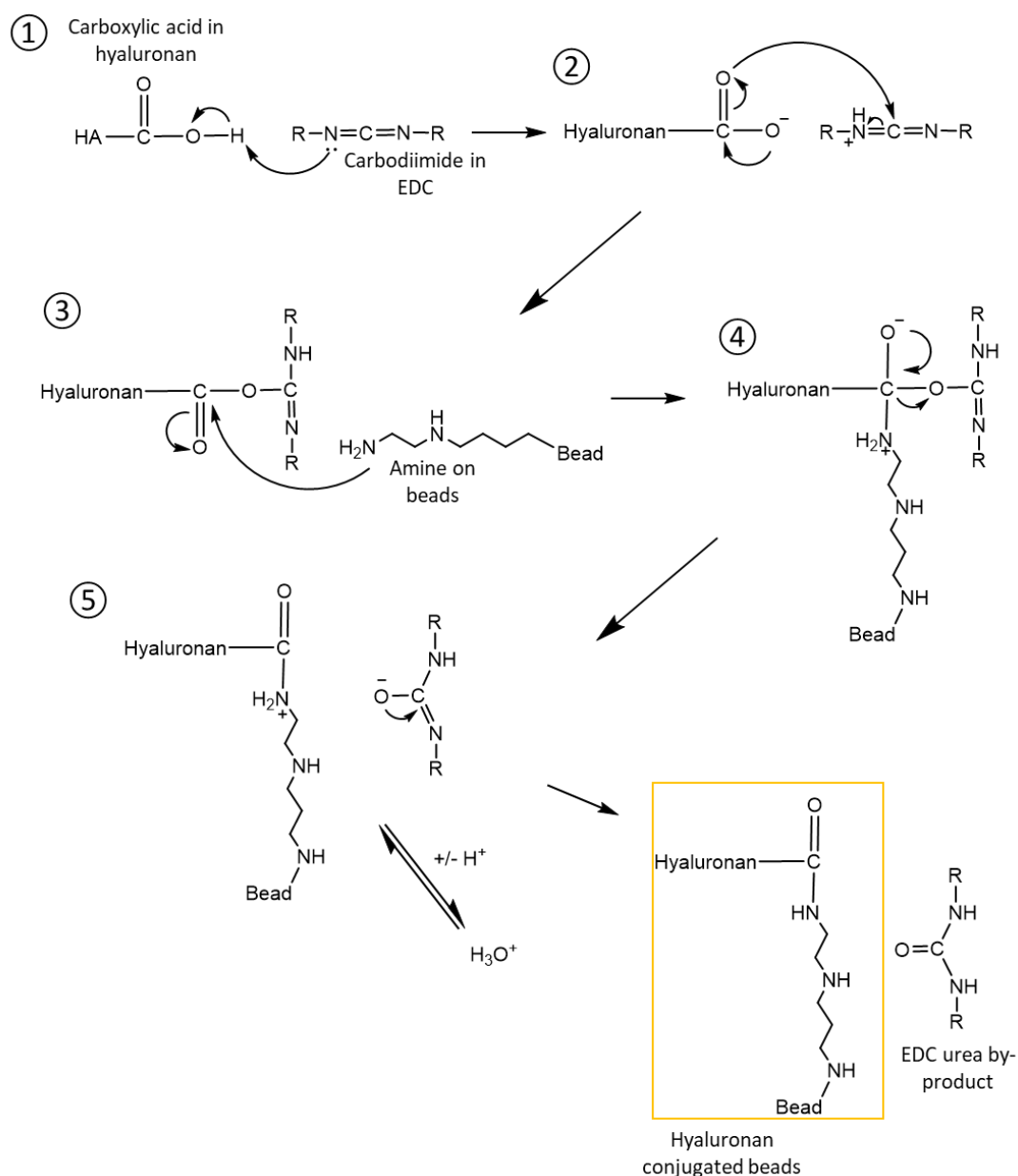
### 5.1.1 Chapter highlights at a glance:

- VCAN can be enriched from other ECM proteins by difference in charge
- VCAN from different cells elute at different stages
- CS proteoglycans elute at different stages

## 5.2. VCAN enrichment via HA-affinity chromatography

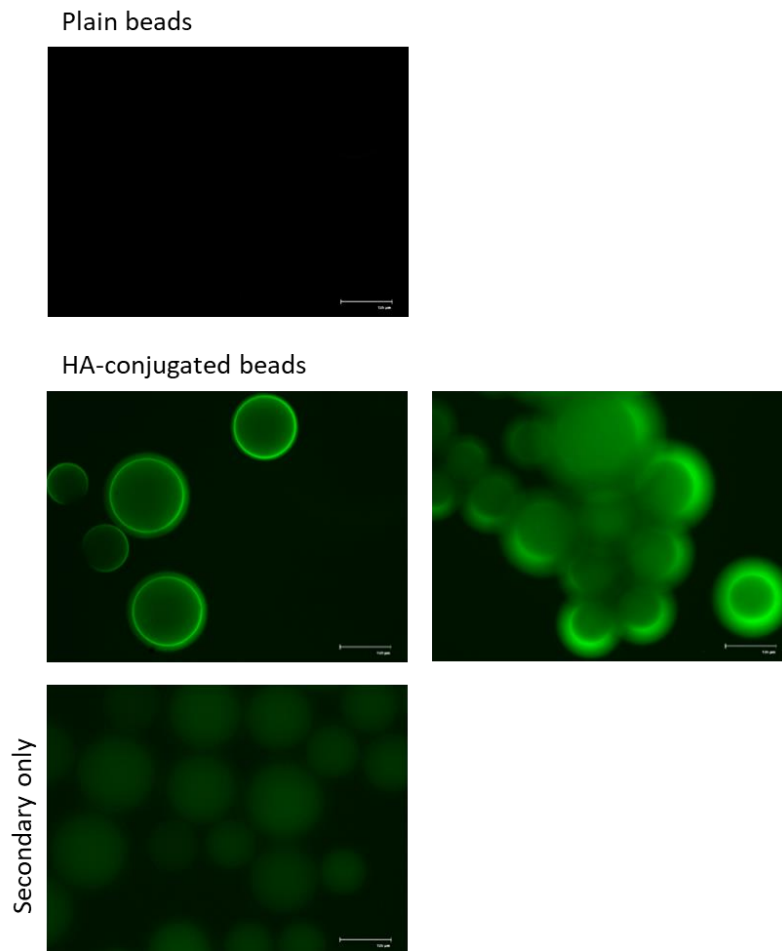
Affinity chromatography with HA was selected as the initial method for isolation of VCAN as it enables isolation of all isoforms. Beads had to first be conjugated to HA. The protocol for binding HA to agarose beads was adapted from Foulcer et al <sup>257</sup>. Carboxylink beads were used to bind to HA. These beads are 4% agarose beads that possess a covalently bound DADPA (diaminodipropylamine) linker which has a free amine at the end. The amine can be covalently bound to the carboxyl group on HA. HA was obtained from Lifecore with a molecular weight of 750K-1M daltons. The reaction binding HA to the amine is catalysed by EDC. The reaction is detailed in Figure 5.1. EDC is functional at a pH around 4.7, so the pH of the solution of beads, HA and EDC was adjusted. The EDC is conjugated to the carboxylate to form the active ester. The primary amine group on the beads displaces the carbodimide via a nucleophilic attack. This leads to the formation of HA conjugated beads and a soluble EDC waste product. The waste product contains urea which leads to an increase in pH. 0.1M MES buffer was used to stabilise the pH of the reaction. The beads were then washed with acetic acid to block unbound amine groups. They were then washed with 1M NaCl to clean the beads and remove metal ions. 0.05M formic acid was used to clean the beads and remove any bound proteins. After this the beads were washed with dH<sub>2</sub>O and then with 0.5M Na-acetate with 0.02% Na-azide, which was also used as the storage buffer. Beads were washed in the wash buffer (Table 5-1) before use.





**Figure 5.1. Binding of HA to carboxyl link beads.** Chemical reaction of the binding of HA to the beads facilitated by EDC. The reaction was carried out at RT at a pH of 4.7, Step 1. Deprotonation of carboxylic acid to make carboxylate salt. Step 2. Nucleophilic attack by the carboxylate to the carbodiimide leading to the formation of an o-acylisourea active ester. Step 3. Amine group from the bead attacks the carboxylic acid. An addition-elimination reaction occurs leaving a protonated amide. Step 4. Oxygen is converted into a good leaving group. Step 5. Stabilisation of compounds through the protonation of water in the buffer. Formation of a HA-conjugated beads and EDC urea by-product.

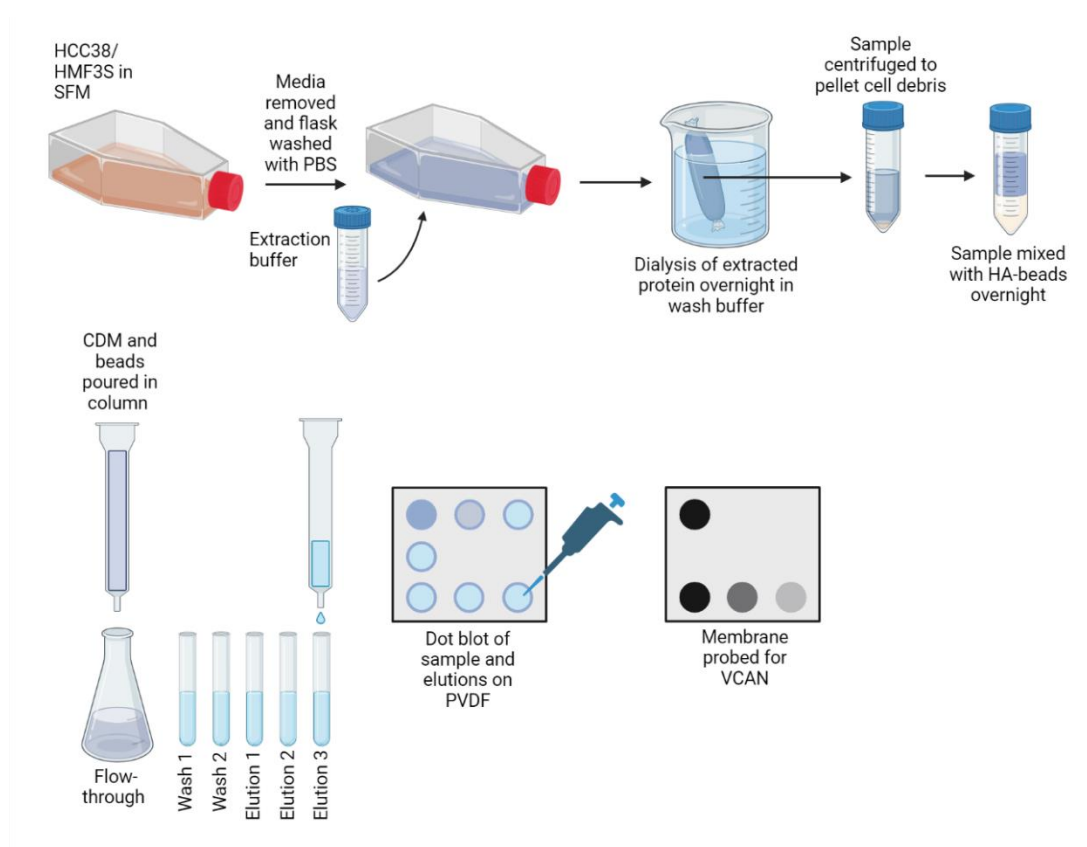
Conjugation of HA to the beads was tested using a biotinylated HA binding protein (HABP). The binding protein is the same structure as the G1 structure of VCAN. The biotin was detected with a streptavidin conjugated to daylight 488. The HA-bound beads were compared to plain beads and beads without the binding protein. A halo of green staining was detected around the HA conjugated beads confirming that the HA had been bound to the beads (Figure 5.2).



**Figure 5.2. IF on beads for HA.** Conjugated and plain beads stained with a biotinylated HABP and fluorescent streptavidin. Beads stained without the HABP were used as a control.

To extract extracellular matrix proteins from cell lines, HCC38 cells were grown in serum free media (SFM) for 7 days. SFM was used so the proteins in the FBS do not interfere with the enrichment. The maximum time of survival for cells in SFM was just over 7 days, therefore the protein was collected at day 7. VCAN is secreted by cells to form a cell derived matrix (CDM) which can be found settled around the cells so an extraction buffer was used which would solubilise the VCAN from the CDM. The protein was extracted directly from the flask by adding the extraction buffer (Table 5-1). Protease inhibitors and EDTA were added to the extracted protein referred to as crude protein. The crude protein was dialysed against wash buffer (Table 5-1) and water to remove the salts from the extraction buffer which may interfere with the ionic binding of VCAN to HA. The sample was then mixed with the HA-conjugated beads before being poured into the gravity flow column for elution. The bed volume of the beads was 5ml so 10ml of buffer was used for each

wash and elution. The process for extraction and isolation is illustrated in **Figure 5.3**.



**Figure 5.3. Protein extraction and isolation procedure.** Cells were grown for 7 days in SFM. The flask was washed and extraction buffer added. The extracted protein was dialysed and then mixed with the HA conjugated beads. VCAN is eluted from the column and tested using a dot blot.

The collected solutions were probed using a dot blot to identify VCAN positive solutions (Figure 5.4A). VCAN was found to be present in only the elutions. No detection of VCAN in the flow through or the washes shows that VCAN is binding to the column and that the bed volume is sufficient for the volume of protein. Ponceau staining of the membrane indicated that the majority of the proteins were present in the elutions, suggesting many protein species were binding to the beads. The binding of proteins to HA is via an ionic interaction, and therefore a salt gradient was introduced to elute protein based on binding strength, with the aim of eluting VCAN at a different point to other proteins. The buffers used to form a salt gradient are shown in Table 5.2

It was also noticed that VCAN is still present within the media, so the media was also included in the extraction. Protease inhibitors and EDTA were added to both the extraction buffer and media. The media and CDM were dialysed before mixing

with the beads and pouring into the column for enrichment. Each elution buffer was added 3 times. The dot blot for the elutions are shown in Figure 5.4B. VCAN was eluted in buffers containing 0.75M NaCl, 1M NaCl and 4M GuHCl. VCAN was the highest in the first of the 3 elutions for each buffer. VCAN positive samples were pooled together and dialysed for use in future experiments. Ponceau staining showed a high protein concentration at EB1, showing that more non-VCAN proteins were removed at this stage.

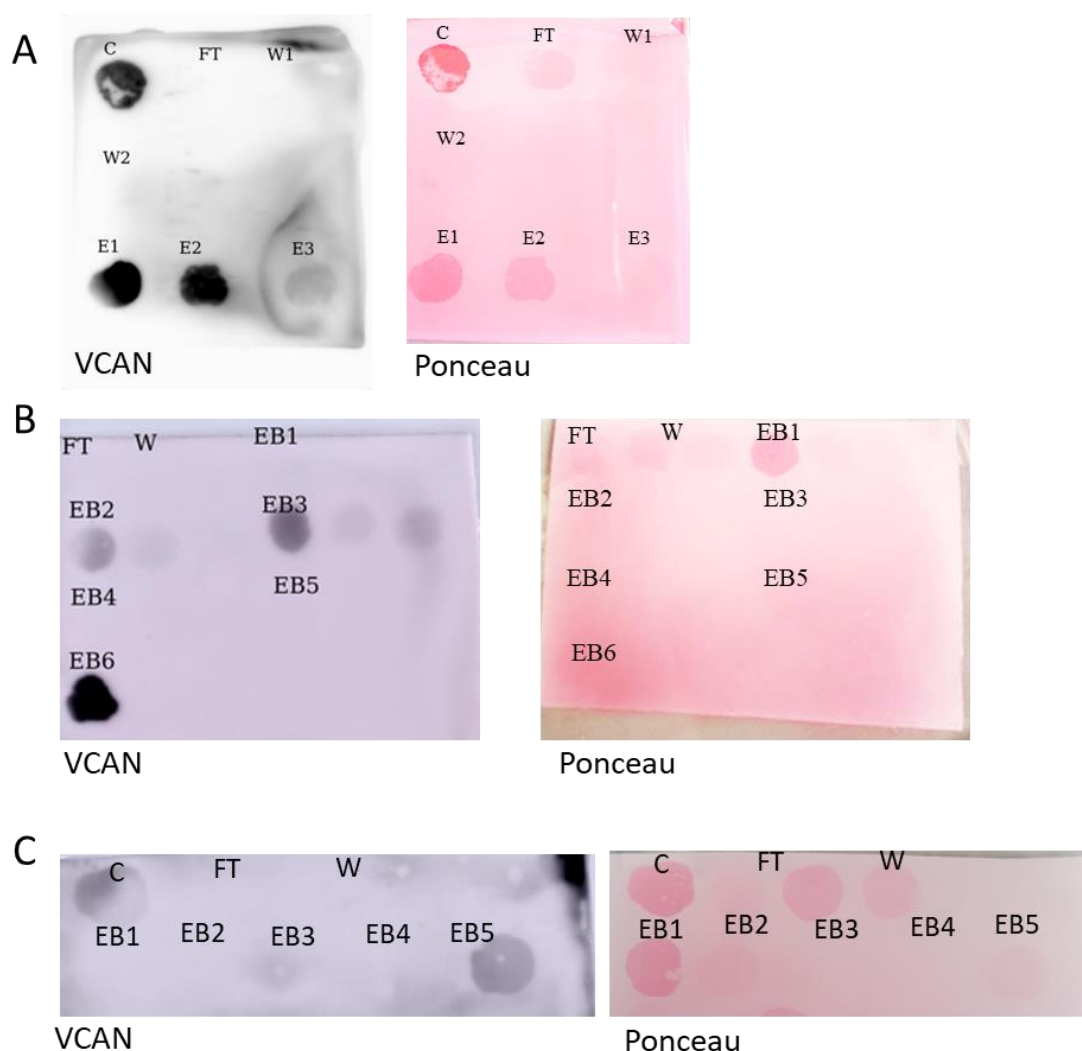
The overall amount of protein within the VCAN enriched fractions was quite low so the amount of crude protein to start with was increased. Six T175 flasks were grown with HMF3S cells, the protein from the media and cell derived matrix was extracted from all flasks. The volume of media was too large to mix with the beads before pouring into the column. Instead 15ml beads were first settled in the column and then the protein solution was carefully poured into the column so as not to disturb the beads. Single elutions were completed for each buffer using twice the bed volume. Dot blot analysis showed similar results to the elutions with less protein, with a majority of the VCAN eluted with 4M GuHCl and a faint detection with 1M NaCl (Figure 5.4C). The positive samples were pooled and dialysed against wash buffer and then water. No VCAN was detected with elution buffers 4 and 5 with very little protein eluted. Therefore it was decided to remove elution buffer 4 as it had little effect to the enrichment process.

**Table 5-1. Buffers composition.** Composition of buffers for the extraction, dialysis, wash and elution of protein.

Buffer	Components
Extraction buffer	8M Urea, 0.05M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Wash buffer	0.15M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer	4M GuHCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA

**Table 5-2. Buffers for salt gradient elutions.** Composition of different buffers for elutions.

Buffer	Components
Elution buffer 1	0.5M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 2	0.75M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 3	1M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 4	1.25M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 5	1.5M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 6	4M GuHCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA



**Figure 5.4. Dot blots of elutions.** Protein from each elution was pipetted onto activated PVDF and stained with Ponceau before probing for VCAN. A) 50ul of crude protein (C), flow through (FT), wash 1 and 2 (W1, W2) and elutions 1,2 and 3 (E1,2 and 3). B) 20ul of flow through (FT), wash (W), Elution buffer 1, 2,3,4,5 and 6 (EB1, 2,3,4,5 and 6). Wash buffer was added to column twice and elution buffers three times. C) 20ul of crude (C), flow through (FT), Elution buffer 1, 2,3,4,5 and 6 (EB1, 2,3,4,5 and 6).

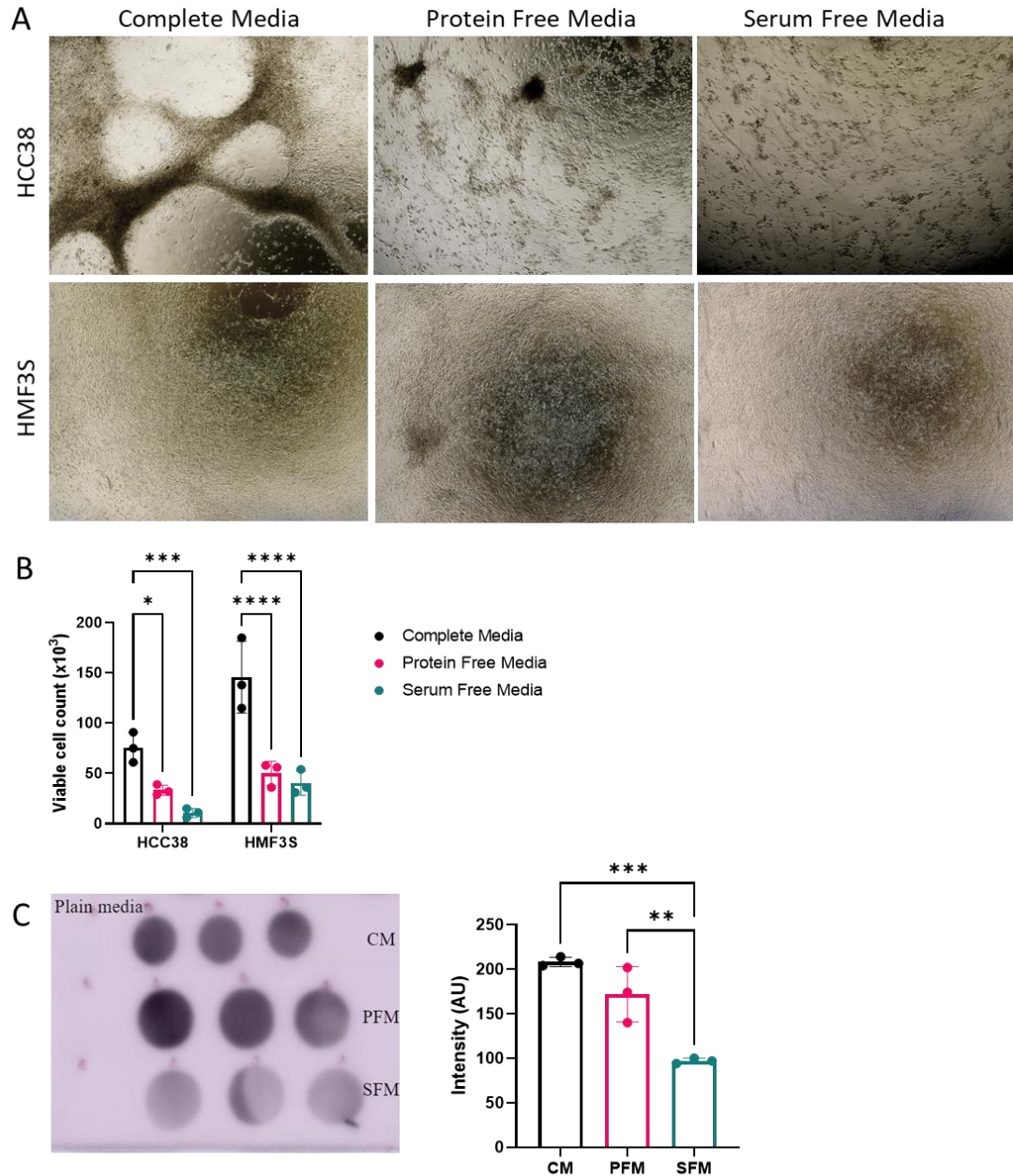
### 5.3. Optimising media for VCAN enrichment

One limitation from the protocol was the length of time that the cells could be cultured for in SFM. This limited the amount of VCAN that could be secreted. RNA analysis in Figure 3.20 showed that there is an increase in the transcription of VCAN at day 7 of culture. To take advantage of the cells producing VCAN at this stage, other media options were trialled to see whether they could improve cell viability. PFM from Lonza was used as it contained other supplements that may support cell growth without FBS. Both HCC38 and HMF3S cells were grown at a density of  $4 \times 10^4$  cells per well of a 24 well plate. The media was changed in the wells to either complete media, PFM or SFM. The cells were grown for 7 days. The media was collected from the wells and the viability of the cells was tested.

Images were taken of the wells at day 7 (Figure 5.5A). For the HCC38 cells, the well containing complete media showed the cells to be confluent and start clumping together. This is a characteristic observed by these cells when the confluency is over 90%. From the PFM some clumping was observed but the confluency was visibly lower. SFM showed no clumping with a lower confluency. For the HMF3S cells, the well with complete media was confluent. The wells with PFM and SFM looked to be confluent however there were high number of floating cells and rounded cells in the wells indicating cell death (Figure 5.5A). The viability of the cells was quantified by detaching the cells from the wells and staining the cells with trypan blue. A haemocytometer was used to count the number of viable cells. A significantly lower viability was observed for both PFM and SFM compared to complete media. A higher number of viable cells was determined for HCC38 using PFM compared to SFM however this was not significant. Compared to the number of cells initially in the well, culturing with complete media led to a fold change of 1.85 for HCC38 and 3.65 for HMF3S. Protein free media saw a reduction in HCC38 with a fold change of 0.8 fold but an increase in HMF3S at 1.25 fold. Serum free media had a fold change of 0.25 for HCC38 and 1 for HMF3S (Figure 5.5B).

The levels of VCAN were then tested in the media samples by dot blot. 20 $\mu$ l of complete media and 40 $\mu$ l of PFM and SFM samples were pipetted onto the activated PVDF. Less volume was used for the complete protein sample as the overall protein level was higher due to the presence of FBS. Densitometry was completed on the samples and the level was normalised for the complete media samples. A significantly higher level of VCAN was seen in the complete media and PFM compared to the SFM (Figure 5.5C). Overall from the comparisons it was observed that PFM did not have a significant advantage over SFM in terms of cell

proliferation. However, less cell death was observed in comparison to SFM. From the levels of VCAN detected it showed that the cells were still producing more VCAN at a lower viability in PFM and therefore it was decided to grow the cells in PFM for future extractions.

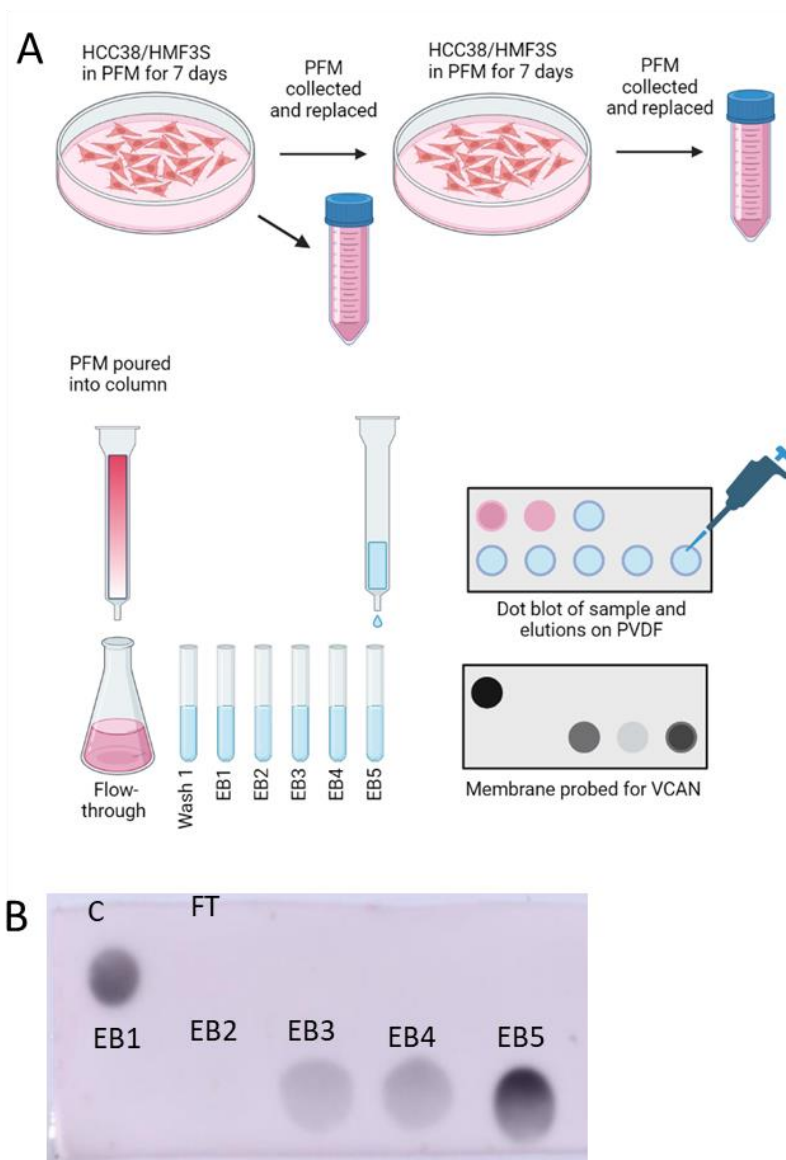


**Figure 5.5. Comparing cell viability and VCAN production in different media.** A) Images were taken of the cells in the 24 well plate after 7 days. B) The number of viable cells were determined at day 7 by staining cells with trypan blue and counting with a haemocytometer. One way ANOVA.  $^* = P < 0.05$ ,  $^{***} = P < 0.005$ ,  $^{****} = P < 0.001$ . C) Media was collected from the wells at day 7. 20 $\mu$ l of complete media (CM) and 40 $\mu$ l of protein free media (PFM) and serum free media (SFM) was pipetted onto activated PVDF. The membrane was probed for VCAN. Intensity of bands was determined using image J. ANOVA.  $^{**} = P < 0.01$ ,  $^{***} = P < 0.005$ . N=3

The effect of PFM towards the enrichment of VCAN was then tested. Instead of T175 flasks, cells were grown in 10cm<sup>2</sup> dishes as it was easier to culture large volumes. HCC38 cells were grown in the dishes for 2-3 days to increase confluency. The media was then changed to 15ml PFM and left for 7 days. From RNA analysis VCAN expression increased after 3 days, therefore collection of secreted VCAN after this point would allow a greater level of VCAN to be collected. Cells were observed under a microscope and there were still viable cells present so more PFM was added and collected 7 days later. After each collection of media, protease inhibitors and EDTA was added and the protein was stored at -20 before the enrichment process. 25ml of the PFM was poured into the column and eluted as previously completed (Figure 5.6A). Elution buffers are shown in

Table **5-3**. A dot blot was completed for the eluants and most of VCAN was eluted in EB5 with low levels in EB3 and EB4 (Figure 5.6B).





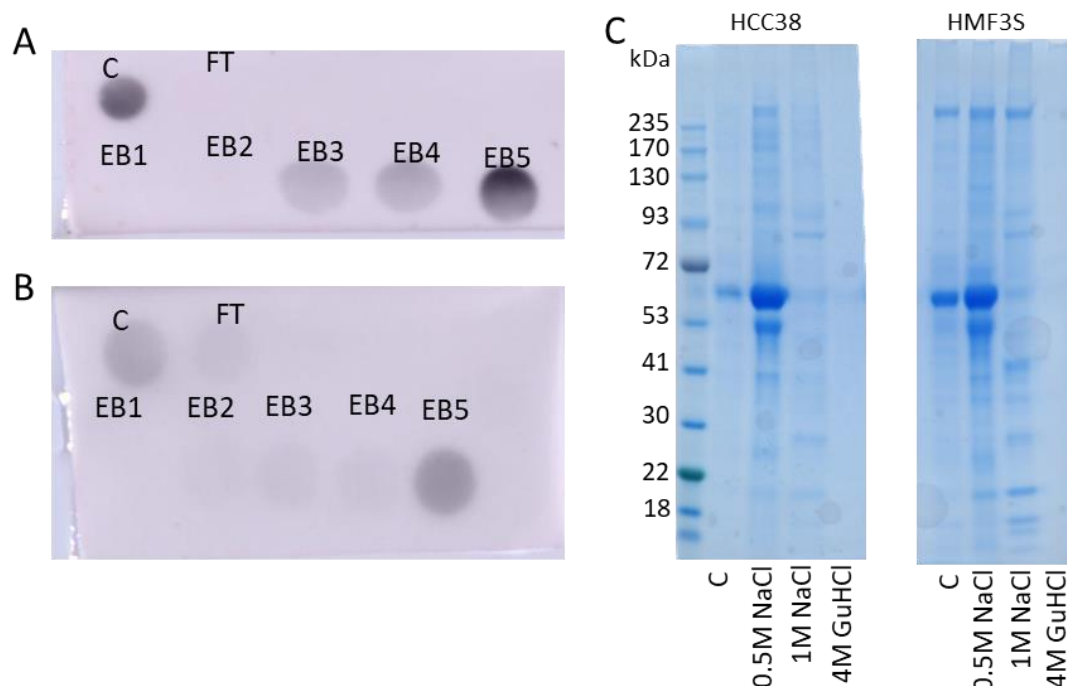
**Figure 5.6. Enrichment process of VCAN using PFM and salt gradient with HA-affinity chromatography.** A) Process of media collection and elution. B) Dot blot completed with crude (C), FT and each elution and probed with anti-VCAN.

**Table 5-3. Elution buffers for updated salt gradient.** Concentration of components in different buffers for elutions.

Buffer	Components
Elution buffer 1	0.5M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 2	0.75M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 3	1M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 4	1.5M NaCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA
Elution buffer 5	4M GuHCl, 0.05M Tris HCl (pH7.6), 0.001M EDTA

#### 5.4. Variations in elutions from different cell lines

It was noticed that when eluting the protein from HCC38 and HMF3S cells, VCAN would be eluted at different points. From HCC38 cells, VCAN was eluted with the 1M, 1.5M NaCl and 4M GuHCl buffers (Figure 5.7A) but with HMF3S cells, VCAN was only detected at the final elution with 4M GuHCl (Figure 5.7B). Gel electrophoresis with coomassie staining was used to compare the differences in proteins at each elution. Samples were dialysed and concentrated using 30kDa centrifugal filters. The crude protein was run along with the elutions from buffers 0.5M NaCl, 1M NaCl and 4M GuHCl. The crude protein showed very little bands for both samples which may be due to the low protein concentration. A thick band was detected around 70kDa in both samples and also in the 0.5M NaCl elutions. This may be histones or heatshock protein. A thick band for larger proteins were observed in HMF3S in the crude and earlier elutions. It was evident that high levels of protein were eluted with 0.5M NaCl buffer with the most bands observed in this well. More proteins were eluted with 1M NaCl with HMF3S compared to HCC38. The darkest bands for both samples seemed to be in the same places. With no VCAN detected at this point for HMF3S, these bands were concluded to not be VCAN. For GuHCl very little protein was detected (Figure 5.7C). This may be due to the levels of protein being low but also the difficulty to run samples in the GuHCl buffer. GuHCl salts can react with the LDS sample buffer leading to precipitation which can affect the running of the protein in the gel.



**Figure 5.7. Comparison of elutions with HCC38 and HMF3S.** Dot blots were completed for crude (C), FT and elution samples from A) HCC38 and B) HMF3S elutions. Membranes were probed for anti-VCAN. C) Samples from C, EB1 (0.5M NaCl), EB3 (1M NaCl) and EB5 (4M GuHCl). Were concentrated and dialysed. Samples were run by gel electrophoresis in a 4-12% gel and stained with coomassie.

Mass spectrometry was completed on the two samples. The aim of the mass spectrometry analysis was to test whether VCAN had been isolated and also identify CS binding sites on VCAN peptides. Enzymes were compared for the digestion of VCAN using the ExPASy PeptideMass online tool to identify which would give the highest coverage. The parameters for the search were iodoacetamide treatment of cysteines, no oxidation of methionines and peptides between 750 and 3000Da. From the search, trypsin had a 47% coverage with 105 peptides, endoproteinase Asp-N has a coverage of 48.3% with 98 peptides and endoproteinase Glu-C has a 65.5% coverage with 205 peptides. As it had the greatest coverage, Glu-C was used to digest the proteins.

VCAN positive samples from the elutions from HCC38 and HMF3S were dialysed and concentrated with a 100kDa centrifugal filter. A 100kDa filter was used to remove any smaller fragments of protein and the smaller versions of VCAN where no CS chains are present. The amount of protein in each sample was determined by nanodrop. 20µg of protein was taken and brought to a volume of 100µl with 1M HEPES buffer. The process of protein digestion was adapted from a protocol to detect extracellular matrix proteins in our lab<sup>23</sup>. The protein was denatured with DTT

and treated with iodoacetamide to convert cysteines to carbamidomethyl-cysteine. PNGase F was used to remove N-glycans. The protein was then digested with GluC and CS chains removed with CSase treatment. The removal of CS leaves behind the linkage tetrasachharide (Figure 1.6) and one CS disaccharide which could be used to detect CS binding sites.

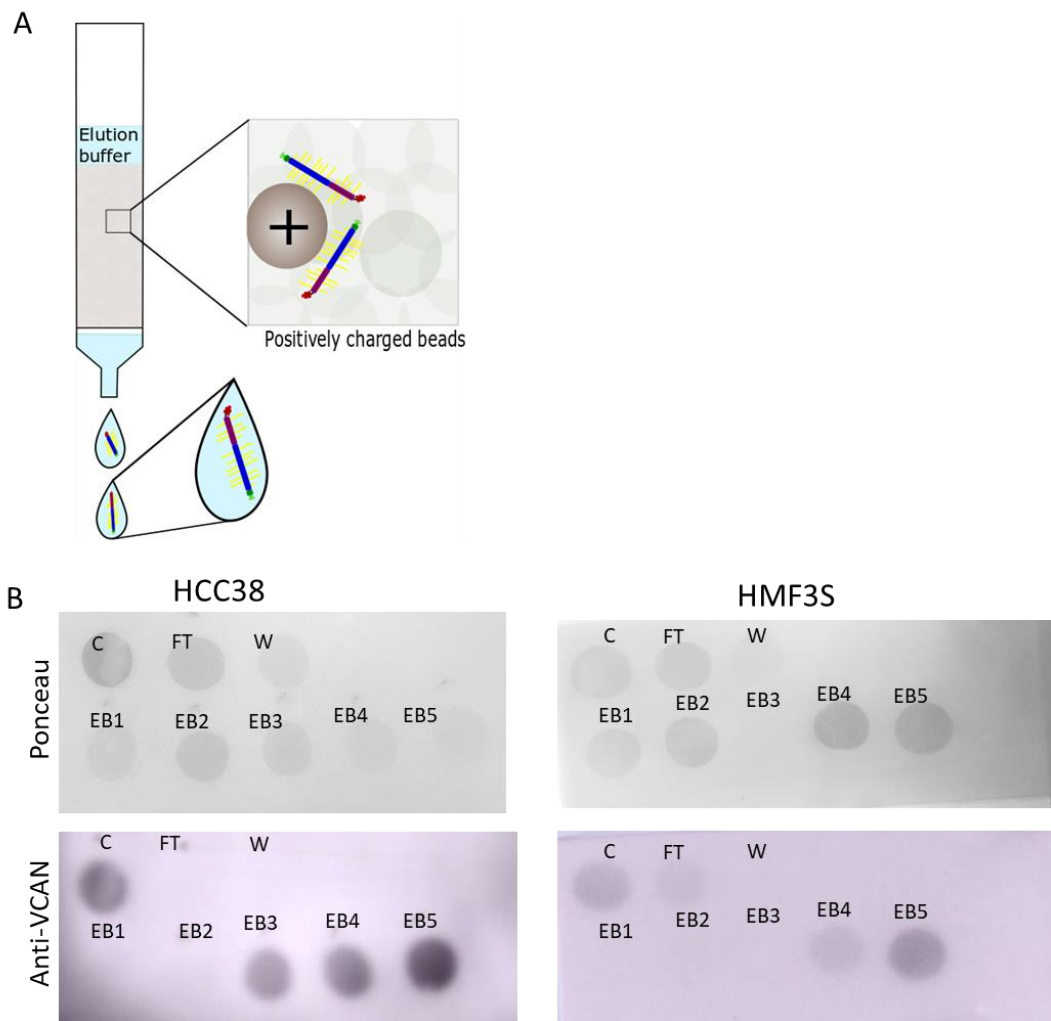
After the sample was run on the mass spectrometry machine, the data was analysed via the Mascot analysis server provides both the mascot score and the emPAI score. The mascot score is the measure of how reliable the identification of a protein is. Therefore the higher the score the more confidence there is of the protein being present and the peptides identified being from that protein. The emPAI stands for the exponentially modified protein abundance index and provides an approximation of the relative quantification of the proteins in the mixture <sup>261</sup>. This is based on the protein coverage from the peptides detected compared to the calculated number of detectable peptides. ECM proteins were extracted from the Mascot analysis to show the number of ECM proteins in each sample and mascot scores. Comparisons of the enrichments from HCC38 and HMF3S showed differences in the number of proteins and the abundance of VCAN within the sample. VCAN isolated from HCC38 had a mascot score of 101.5 and an emPAI score of 0.105 whilst VCAN isolated from HMF3S had a mascot score of 48.5 and an emPAI score of 0.05. Within both samples the proteoglycan BGN was detected at high levels. DCN was detected at high levels in the HMF3S protein but was not detected in HCC38 protein. qRTPCR analysis from the cell lines showed that there was a higher level of DCN expression in HMF3S cells compared to HCC38. Other non-HA binding PGs like AGRN were also detected (Figure 5.8). AGRN is decorated with heparan sulphate GAG chains and is a constituent of basement membrane <sup>262</sup>. The identification of these proteins indicated that the enrichment was not specific to HA-binding proteins and that some proteins may be binding to the agarose beads via ionic interactions. High levels of cellular proteins were also observed. Therefore we decided to adapt the approach based on what had been learnt to improve the enrichment.



## 5.5. Ion exchange chromatography

As the VCAN was enriched in a manner similar to IEX, it was decided to switch to IEX chromatography for isolation. This method would mean that there would be less V3 enriched, however from the RNA analysis, the V3 levels in the cells were quite low. IEX has the advantage over affinity approaches as the resin does not need to be prepared and the beads would be more selective for the larger isoforms with more CS chains.

For the IEX chromatography, DEAE beads were used. These beads have a positive charge to attract negatively charged proteins (Figure 5.9A). Methacrylate beads were selected instead of agarose beads. This was to limit the background proteins which may bind to agarose. The DEAE methacrylate beads were washed with deionised water and wash buffer before adding the protein free medium to the samples. Initially the bead bed was set in the column to 5ml and 20ml media was added to the beads, however it was observed that the phenol red and other proteins in the media were binding to the beads at the top of the bed which prevented the remaining buffer moving through the column. The protocol was adapted and the media was mixed with the beads on a roller for 2hrs at 4°C before being poured into the column. The salt gradient for the elution was kept the same as the previous experiment, however EDTA was removed from the buffers as it was interfering with functional assays using the eluted proteins (Table 5.4). Interestingly VCAN was eluted at similar salt buffers which further confirmed that the protein was binding by charge to the HA-beads (Figure 5.9B).



**Figure 5.9. IEX chromatography for the enrichment of VCAN.** Schematic showing VCAN binding to the positively charged beads during IEX. B) Dot blots of the elutions of HCC38 and HMF3S using IEX. Membrane stained with Ponceau and probed with anti-VCAN.

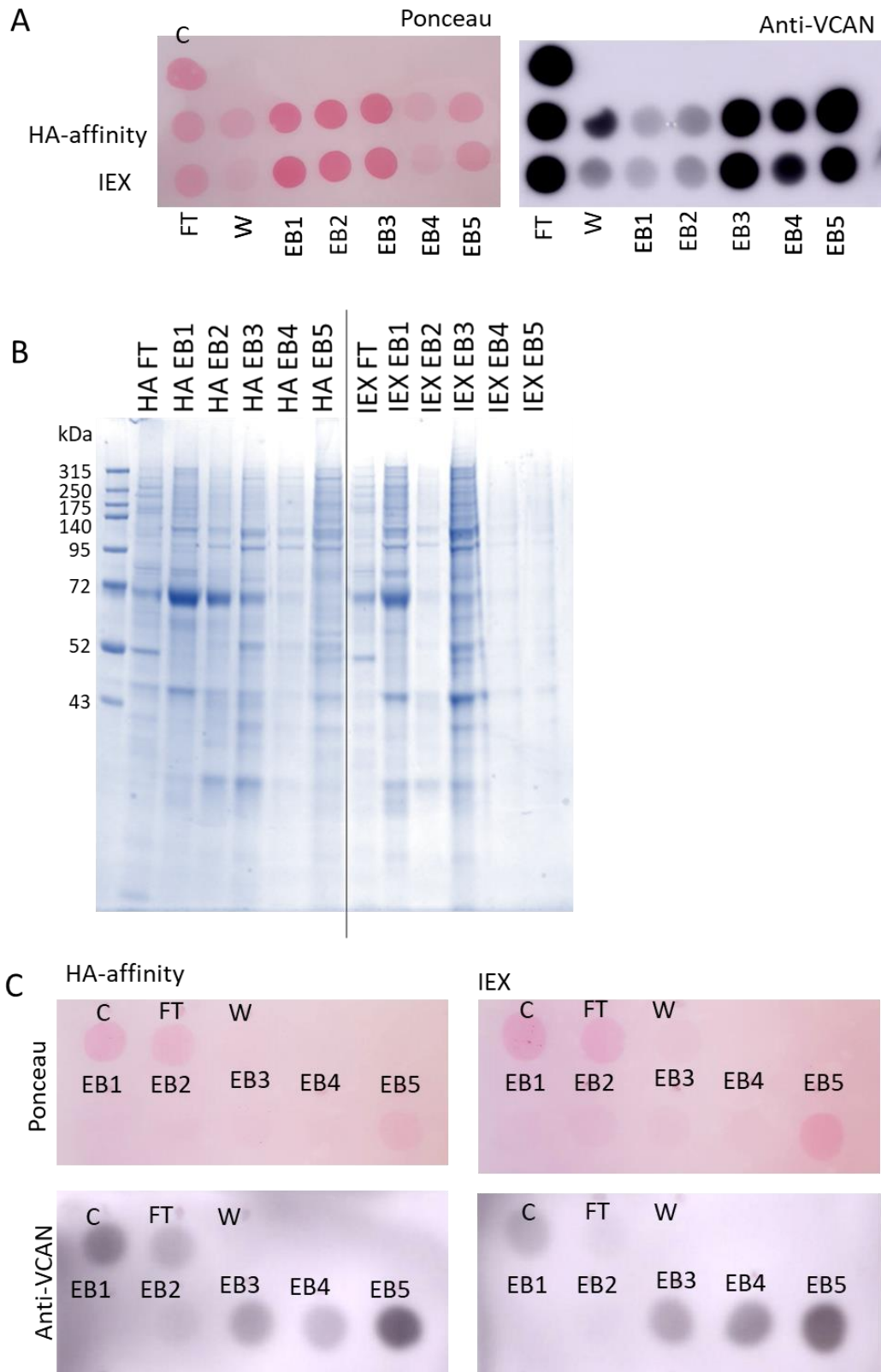
**Table 5-4. Elution buffers used for IEX chromatography.** Concentration of components in different buffers for elutions.

Buffer	Components
Elution buffer 1	0.3M NaCl, 0.05M Tris HCl (pH7.6)
Elution buffer 2	0.5M NaCl, 0.05M Tris HCl (pH7.6)
Elution buffer 3	1M NaCl, 0.05M Tris HCl (pH7.6)
Elution buffer 4	1.5M NaCl, 0.05M Tris HCl (pH7.6)
Elution buffer 5	4M GuHCl, 0.05M Tris HCl (pH7.6)

The two different methods were compared to identify whether there were any improvements in the enrichment by using IEX over HA-affinity. PFM from the culturing of HCC38 cells was taken and added to a falcon tube with 1ml of HA-beads or 1ml of DEAE beads. The PFM was mixed with the beads before being poured into a column and the protein eluted. 2ml elutions were collected and a dot blot was completed to compare the elutions. The ponceau staining and the elutions were found to be very similar with VCAN detected at the same stages and the intensity of the ponceau also equal at each elution (Figure 5.10A). 100µl of each elution was precipitated using ethanol to be able to run the samples in a gel. The samples were run in a 10% acrylamide gel and the gel was stained with coomassie. The high level of protein seen from the ponceau stains translated to the coomassie stains for the HA eluted samples however this was not seen with the IEX samples where in EB2 a very low level of protein was detected. This may be due to the incomplete solubilisation of protein after precipitation. From the stain it was visible that there is a difference in the proteins eluted from each column. The highest level of protein was seen in EB3 for the IEX beads but EB5 for the HA beads. The more prominent bands showing the highest concentrated proteins were quite similar between the two samples for each elution. (Figure 5.10B). Overall from the dot blot and coomassie staining differences between the elutions of IEX and HA-affinity chromatography could not be made.

The FT from each sample was then run again on the gels to see if a better distinction could be made. With this elution the most protein was seen in the FT and EB5. For the DEAE beads VCAN was eluted at EB3, 4 and 5. For the HA samples VCAN was still detected in the FT (Figure 5.10C). This indicated that the capacity for VCAN binding was higher in the DEAE beads (IEX) compared to the HA beads.

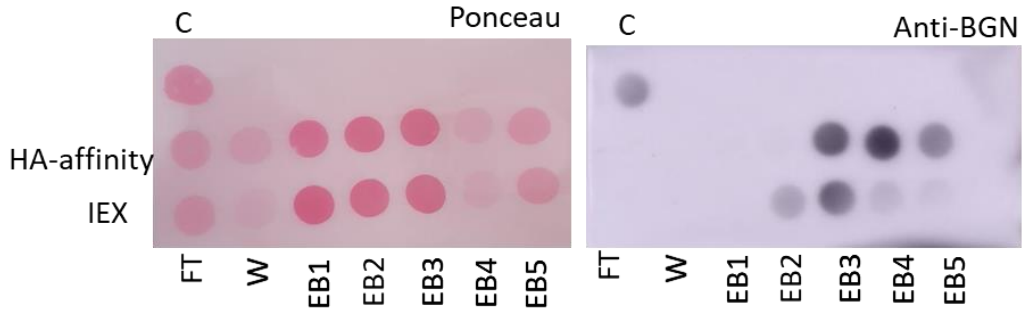




**Figure 5.10. Comparing IEX and HA- affinity chromatography.** A) Dot blot of crude(C), FT and elutions using both IEX and affinity chromatography. Membrane stained with Ponceau and then probed for VCAN. B) Elutions precipitated and run on a 10% gel. Gel stained with coomassie G250. C) FT from previous elutions added to beads. Dot blot of crude(C), FT and elutions completed, Membrane stained with Ponceau and then probed for VCAN.

The VCAN positive elutions in the initial enrichment were pooled together to test by mass spectrometry. The sample was dialysed against water and concentrated using a 100KDa centrifugal filter. The protein concentration for the samples was determined using the nanodrop and 20µg was taken for mass spectrometry. The samples were processed using the same protocol as before. From the analysis more proteins were detected by HA-affinity chromatography compared to the IEX, indicating that there is a better enrichment with the IEX column. The mascot scores and the emPAI score for VCAN (Blue arrow) were lower than expected for both enrichments, this may be due to the oversaturation of the column or possibly poor enzyme degradation so less peptides were detected. The mascot score for VCAN in the affinity chromatography was 35 whilst the score for IEX was 28. Both scores are below the normal cut off of 50, indicating low confidence in the detection of the protein. The emPAI score was the lowest for both techniques (Figure 5.11). From the data there is little to suggest that VCAN has been enriched in the samples, however from the overall number of proteins it can be interpreted that the level of VCAN per µg of protein is greater in the IEX enriched samples, due to the overall lower number of proteins. It was also interesting to note that other CS proteoglycans like BGN were only detected in the affinity chromatography sample and not in the IEX. The binding of VCAN to the IEX column may be stronger than BGN. This might be due to the higher number of CS chains present.





**Figure 5.12. BGN detections in HA-affinity chromatography and IEX.** Dot blot was completed on the crude, FT and elutions for both columns. Membrane was stained with ponceau and probed with anti-BGN.

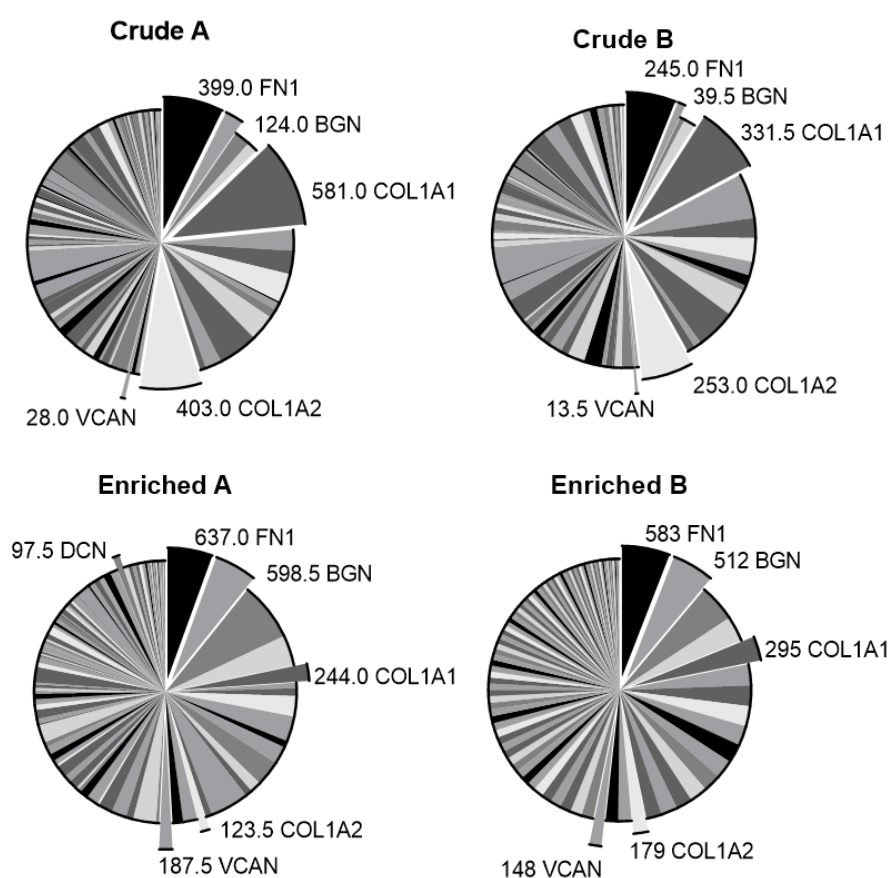
From this data, it could be concluded that HA-beads were not enriching VCAN as expected. The comparison of affinity chromatography with IEX chromatography showed that VCAN was being enriched in similar ways between the two techniques. However, IEX was shown to have a better binding capacity for VCAN and have lower levels of other CSPGs present, making it a better procedure to move forward with.

## 5.6. Determining level of VCAN enrichment

HCC38 cells were grown on five 10cm<sup>2</sup> dishes for VCAN enrichment. As described previously the medium was changed to PFM on day 3 and grown for 7 days before the medium was collected. The VCAN positive samples were identified as EB3, EB4 and EB5 as seen in previous elutions. In sterile conditions, the positive samples were pooled together and filter sterilised using 0.45µm filters. Sterile 100kDa centrifugal filters were used to dialyse and concentrate the samples. Samples were dialysed against sterile water. Crude protein was filter sterilised to be used in comparisons to the enriched in functional assays. The level of protein in the sterile samples was quantified by taking a small aliquot and determining the concentration using a BCA assay. 20µg of proteins were taken for MS analysis. The same method was used to digest and process the samples as previously stated. Two different VCAN enriched protein samples were compared to see if the components within enrichments are similar across biological repeats.

From the MASCOT analysis it was found that the enriched samples contained more proteins than the crude samples (Appendix 6,7). This was unexpected as proteins were seen to be removed from the enriched sample as they were detected in the

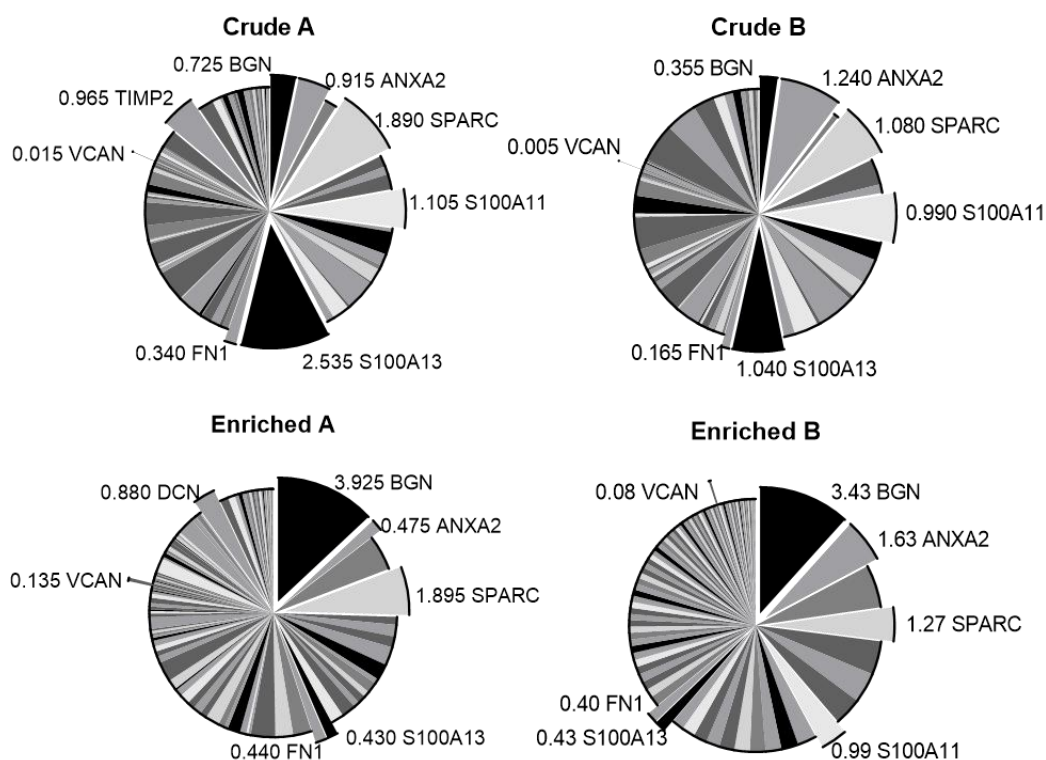
flow through and earlier elutions. The mascot score and emPAI scores were analysed for both the samples. Higher mascot scores indicate greater confidence in the detection of the protein. The mascot scores in the crude were the highest for COL1A1, COL1A2 and FN1. In the enriched, the scores increased for FN1 but dropped with the collagens. VCAN was seen to increase in the mascot scores for both samples with the score being 28 and 13.5 for the crude samples and 187.5 and 148 for the enriched. BGN was also seen to have a higher mascot score in the enriched samples (Figure 5.13, Appendix 6,7).



**Figure 5.13. Mascot scores of proteins in the crude and enrichments.** Pie charts showing proteins detected by mass spectrometry and the mascot scores. Highest detected proteins and CS proteoglycans highlighted.

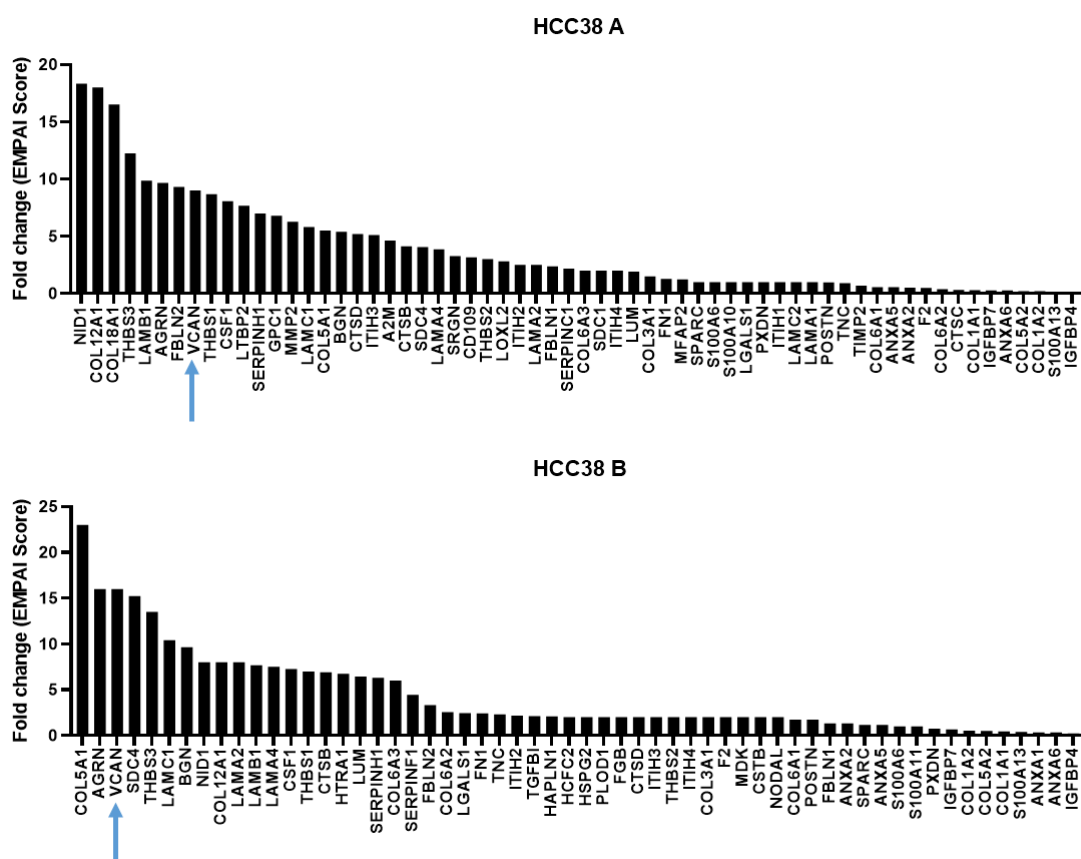
The emPAI score gives an indication of the abundance of the proteins. Based on the emPAI scores ANXA2, SPARC, S100A13 and S100A11 had the highest abundance in the crude. Differences were observed between samples for the changes in the scores after enrichment. In the enriched A sample, a reduction was seen in these proteins whilst in Enriched B, ANXA2 and SPARC increased. The emPAI score for VCAN was seen to increase in both samples. Other CSPGs such

as BGN and DCN were also seen to increase in the enriched samples (Figure 5.14, Appendix 8,9).



**Figure 5.14. emPAI score of proteins in the crude and enrichments.** Pie charts showing proteins detected by mass spectrometry and the emPAI scores. Highest detected proteins and CS proteoglycans highlighted.

The fold change was determined by comparing the emPAI score from the crude and enriched. Through this the level of enrichment could be determined for each protein. VCAN was determined to have a fold change of 9 in sample A and 16 in sample B. BGN had a fold change of 5.4 in A and 9.6 in B. VCAN was found to have the highest fold change in comparison to other CS proteoglycans (Figure 5.15) .



**Figure 5.15. Fold change in the emPAI scores of the proteins.** Fold change calculated by comparing the emPAI score in the crude and enrichments of each samples.

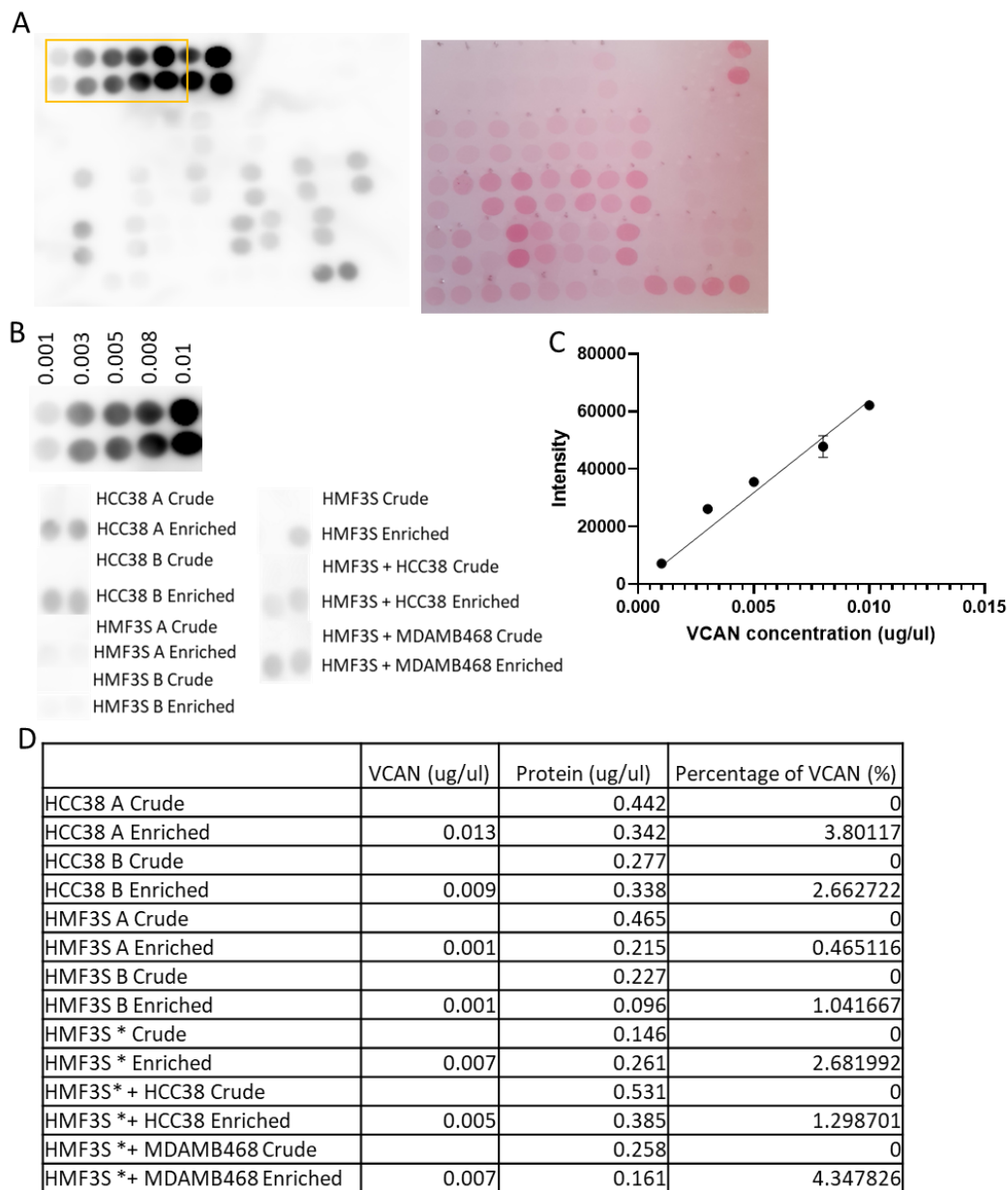
Overall, VCAN was found to be enriched around 12 fold in the enriched samples from HCC38 secretome. Variations in proteins detected and levels were observed across samples, which is common in mass spectrometry with multiple runs of the same sample producing different results. With similar patterns observed in the highest detected proteins and overall number of proteins, it shows that there is potential for reproducibility of the enrichments.

To be able to quantify the amount of VCAN in the samples as  $\mu\text{g}/\mu\text{l}$ , I used a similar principle from protein quantification assays where the amounts are determined by comparing to standards. A dot blot was completed using commercially made VCAN peptides containing the G1 domain (Abcam, ab39784). Concentrations of 0.001, 0.003, 0.005, 0.008, 0.1, 0.2 and 0.5  $\mu\text{g}/\mu\text{l}$  were used as standards. 5  $\mu\text{l}$  of each sample was then pipetted onto the membrane. Each standard and sample was pipetted in duplicate. Ponceau staining was completed to compare the protein concentrations within each sample. The membrane was probed for VCAN. (Figure 5.16A, B). Densitometry was used to determine the intensity of the standards which

could be plotted as a standard curve. The standard curve was then used to interpolate the levels of VCAN in the samples based on the intensity. For the standard curve, standards up to 0.1 were used as the samples with a concentration above had become overexposed and therefore not accurate to read (Figure 5.16C).

Samples included the crude and enrichments of HCC38 and HMF3S in mono-culture and co-cultures of HMF3S with TNBC cell lines. From the crude samples, the level of VCAN was below the level of detection and therefore the VCAN concentration could not be determined. This confirmed that the levels of VCAN were higher in enriched samples. To understand how the level of VCAN related to the level of overall protein in the sample, the protein concentration for each sample was determined via BCA assay. Higher levels of VCAN were determined in HCC38 enriched protein compared to HMF3S, as seen in the mass spectrometry data in Figure 5.8). For the cultures of HMF3S with TNBC cell lines, the TNBC cell lines had opposing effects. The addition of HCC38 reduced the overall VCAN level whilst MDA MB468 increased VCAN levels. This may be due to the presence of other ECM proteins that are secreted by HCC38 and possibly the stimulation of the secretion of these other proteins by HMF3S (Figure 5.16D).

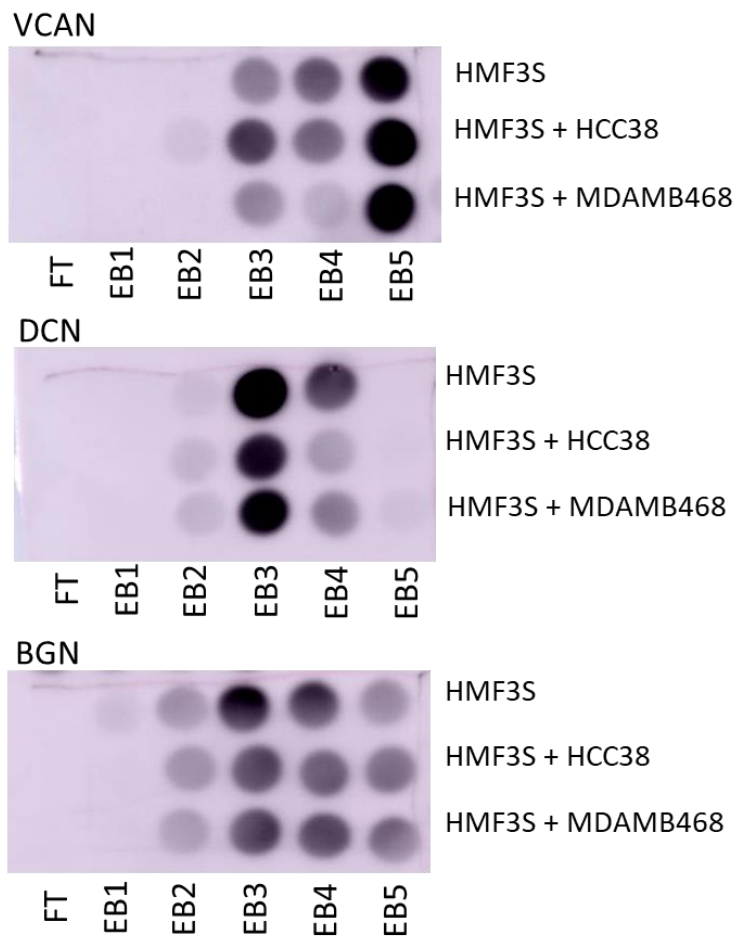




**Figure 5.16. Quantification of VCAN by comparing to standards.** A) Dot blot completed of the standards and the samples and probed for anti-VCAN. B) Standards and samples blotted. C) Standard curve based on the intensity of the blots and the VCAN concentrations. D) Interpolation of VCAN concentrations from the standard curve. Table showing level of VCAN, overall protein concentration of samples and the percentage of VCAN within the sample.

## 5.7. Proteoglycans in enrichments

IEX-beads will bind negatively charged proteins that includes VCAN and other proteoglycans, as seen by MS analysis previously discussed. From the mass spectrometry analysis in Figure 5.13 and 5.17, it is visible that other proteoglycans like DCN and BGN are present in the sample. From the IHC analysis in Chapter 4, CS was seen to be a potential factor in immune exclusion. When studying the role of CS via functional assays using these enriched samples, the presence of other proteoglycans need to be taken into account. Therefore the elution of the other detected proteoglycans was compared. Whilst VCAN was found to be eluted at predominantly EB5, DCN was predominantly eluted at EB3. BGN was equally eluted at EB3, EB4, and EB5 (Figure 5.17). It may be possible to strategically remove DCN and BGN by selecting the elutions with lower levels of these proteins such as EB6. However this will means sacrificing VCAN that is eluted in the other fractions. Size exclusion chromatography could also be applied to remove DCN and BGN, however smaller VCAN proteins may be lost in the process.



**Figure 5.17. Elutions of different CS proteoglycans.** Protein from HMF3S alone and in co-culture was enriched using IEX. 15µl of the flowthrough (FT) and elutions were blotted onto activated PVDF and probed for anti-VCAN, anti-DCN and anti-BGN.

## 5.8. Discussion

The isolation of VCAN has been trialled by different labs using multiple different methods. The size of VCAN has made it a difficult protein to be isolated but various techniques have been trialled by other labs <sup>111,257,259</sup>. The main methods being HA-affinity chromatography, IEX chromatography and size exclusion chromatography. The determination of the technique to use was based on the yield, enrichment, and which forms of VCAN would be isolated.

HA-affinity chromatography was selected as the initial method of isolation in order to isolate all VCAN isoforms. The limitation of this technique is that all proteins which bind to HA will be selected for. This includes other proteoglycans and also versikine. HA-affinity columns have been previously used to isolate hyaluronan binding proteins <sup>263</sup>. HA-beads have also been used to study the binding of proteins to hyaluronan <sup>264</sup>. When trialling using HA-beads to try to isolate VCAN, it was deemed unsuccessful with high levels of other proteins eluted along with VCAN. Agarose is the common support for affinity chromatography and is known to have low background binding. However, from the plain beads control it was seen that VCAN was able to bind to the beads with no HA present. This may be due to the positive charge of the amide group at the end of the linker regions attracting negatively charged proteins. Trying to increase the HA binding to the beads to avoid this did not seem to impact VCAN binding and it seemed that VCAN was not binding sufficiently to the HA chains. This may be due to the conformation of VCAN in the solution. In tissues the binding of VCAN to HA is strengthened by the presence of a link protein <sup>60,86</sup>. The formation of this binding complex does not occur in column isolation and may have affected the effectiveness of the HA-beads.

Ion exchange chromatography was then tested as VCAN seemed to bind to the beads by charge. The charge of a protein is dependent on the pH of the stationary resin and the isoelectric point of the protein. The isoelectric point of VCAN is 4.42. For the solvents a pH of 7.6 was used at which VCAN will have a negative charge. VCAN was found to elute from the column at 3 different salt concentrations (1M, 1.5M NaCl and 4M GuHCl). The elutions were run in a gel however protein bands visualised with coomassie appeared comparable across the three fractions. Western blots were trialled for VCAN however the level of protein was too low for detection, this also suggested that the bands detected in the coomassie may not represent VCAN. The effect of CS chain was also explored. Chain length analysis was trialled by isolating the GAG chains from each elution and running the samples by gel electrophoresis. The gel was stained with alcian blue but other GAGs such as

heparan sulphate were also present which limited the ability to distinguish CS chains between elutions. The impact of the number of CS chains could be tested from the DCN and BGN elutions. DCN has 1 CS chain <sup>252</sup> and BGN has 2 <sup>253</sup>. High levels of DCN were seen to elute at 1M NaCl compared to the 4M GuHCl for VCAN and BGN. However BGN was seen to elute at multiple different salt concentrations. Therefore no conclusions could be made with number of chains and ionic charge.

Another method for isolating proteins is immunoprecipitation. This requires the use of an antibody bound to beads which will only bind to the desired protein. For VCAN there were no commercially available antibodies that had been trialled with immunoprecipitation. This method was briefly trialled with a few antibodies that were used for IHC and immunoblotting. VCAN was not detected after eluting the beads. This may be because of the size of the protein and the also the conformation in solution preventing antibody binding.

Mass spectrometry was used to check for the detection of VCAN peptides and examine the level of enrichment. The outcome of unlabelled mass spectrometry data can be varied and dependent upon multiple factors such as enzymes used for digestion, charge used for detection and also software used for analysis. The enzyme selection was based on the ability to detect VCAN peptides to identify points of CS modification. This may have also led to the enrichment of the detection of other proteins. From the analysis on the various samples multiple proteins were detected where only one sequence was found. Low mascot scores and single sequence detections can be seen as false positives. These factors have to be taken into account during the interpretation of the mass spectrometry data. From the mass spectrometry data from the enrichments, having a higher mascot score and emPAI score gave confidence that there was a higher level of VCAN present. Other proteins also showed an increase in these scored indicating they were also enriched. The peptides of VCAN detected were from the GAG- $\alpha$ , GAG- $\beta$  and G3 domains, indicating that all VCAN isoforms were potentially present within the sample. For further enrichment of VCAN, negative isolation techniques could be used to remove the other proteins that were enriched in the sample. For example FN1 can be removed using a gelatin based affinity chromatography <sup>265</sup>.

To conclude, from the isolation process, an enrichment of VCAN was achieved through the use of IEX chromatography. To achieve a pure isolation of VCAN, a sequence of isolation procedures would be required. Isolating all isoforms of VCAN

in a pure sample would not be possible and therefore the techniques would need to be based on the biochemical properties of these isoforms.

## Chapter 6: Exploring the structural variations of VCAN and their role in T cell trafficking.

## 6.1. Introduction

Having observed associations with VCAN and T cell localisation in tissues, I next tested whether VCAN has a functional role in T cell trafficking using *in vitro* models. The models used were designed to look at the impact of high VCAN levels in the epithelial zone as observed in Chapter 3 and the effect of CS chains as shown in Chapter 4.

Collagen gel models were selected to build the different regions of the tissue. Collagen gels are made with Collagen I and can be cross-linked to form 3D gels. Multiple cell types have been grown in collagen gels and the presence of ECM expression has been detected<sup>266</sup>. The limitation of using collagen gels is that the levels of collagen present may eclipse the effect of VCAN, therefore this had to be taken into consideration when adding protein to the gels. Transwell models were also explored to look at the direct effect of VCAN protein to T cells without other cell types such as tumour cells and fibroblasts that may have an effect on T cell phenotype and migration.

The exploration of the function of VCAN structure towards T cell migration was done using enzymatic degradation to attempt to differentiate function controlled by the protein backbone vs the CS decorations. CSase cleaves the CS chains in between the different disaccharides leaving behind the linkage tetrasaccharide and a single CS disaccharide<sup>267</sup>. ADAMTS enzymes can cleave VCAN to form the versikine fragments. For the analysis, ADAMTS4 was selected as it has been found to degrade V0, V1 and V2<sup>117,268,269</sup>.

In this chapter, hydrogel and transwell models were made which incorporated enriched amounts of VCAN, done using enriched protein, chemokines to stimulate production and also plasmid overexpression. The effect on T cell invasion was then explored, with a focus on the impact of CS chains and the degradation of VCAN to versikine.

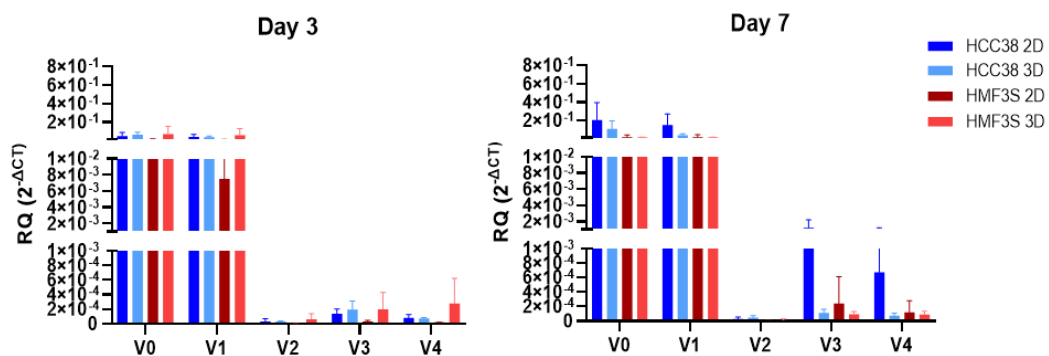
### 6.1.1. Chapter highlights at a glance:

- TGFβ3 stimulates VCAN production and can affect T cell invasion
- Removal of CS from CS-C high VCAN increases T cell invasion in a transwell assay and a collagen gel model.
- Versikine correlated with VCAN levels in the stroma of inflamed tissues.



## 6.2. Exploring VCAN expression in 3D

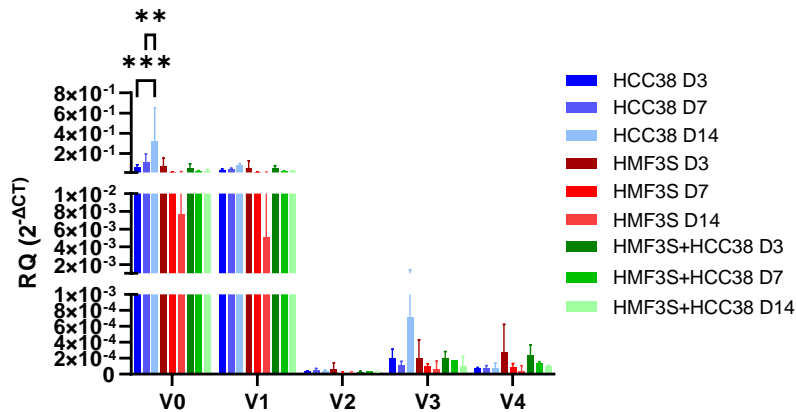
Collagen was used as a base for the hydrogel model. Previous experiments had been completed in the centre making collagen based gels and the protocols were adapted for the cell lines HCC38 and HMF3S. Initially the RNA expression of the cell lines grown in 3D gels and in 2D monoculture were compared. At both Day 3 and Day 7 there was no significant difference between the amount of VCAN transcribed in 2D and 3D culture. The level of V3 and V4 in 2D did seem to be higher however this was not significant (Figure 6.1).



**Figure 6.1. VCAN isoform expression in cells grown in 2D and 3D.** HCC38 and HMF3S cells were grown in a 24 well plate and in a collagen gel. The RNA was extracted at day 3 and day 7. qRT-PCR was carried out for the VCAN isoforms (V0-V4) and normalised to RPS13 expression. N=3.

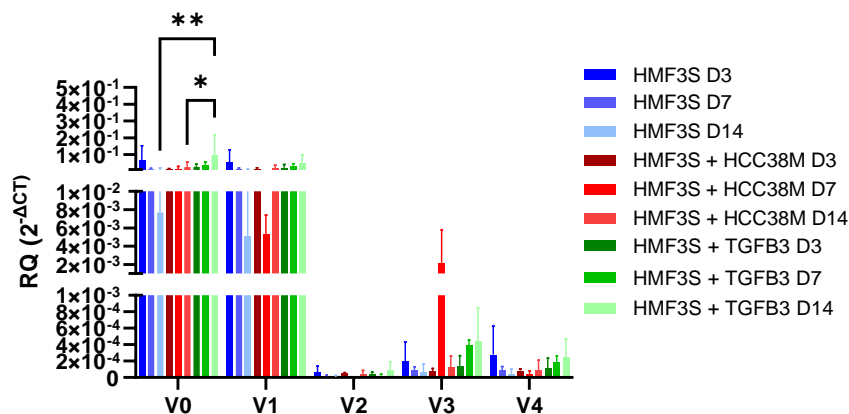
Various conditions were tested to test whether we could increase VCAN production. These included growing cells in the gels as a mono-culture and co-culture, as well as treating HMF3S with 10ng/ml TGFβ3 and HCC38 conditioned media. For the co-culture a 2:1 ratio was selected for HMF3S to HCC38. TGFβ3 was selected as it was previously shown to activate high grade serous ovarian cancer fibroblasts to produce greater levels of VCAN <sup>266</sup>. The RNA was extracted from the gels at different time points to have an understanding at what point VCAN production would be at the highest to take into account when building a model to study VCAN and T cell migration.

There was a significant increase in the V0 expression in HCC38 cells in the gels at day 14. An increase was also seen at D14 in the V3 isoform but this was not significant. No difference was found in the other isoforms. The HMF3S cells did not show any significant difference for any isoform, however a reduction in the levels of V0 and V1 were observed at day 14. Within the co-culture gels no difference was detected across the time points (Figure 6.2).



**Figure 6.2. VCAN isoform expression at different time points.** HCC38 and HMF3S cells were grown in a in a collagen gel as mono-cultures and a co-culture (1:2). The RNA was extracted at day 3, 7 and 14. qRT-PCR was carried out for the VCAN isoforms (V0-V4) and normalised to RPS13 expression. N=3

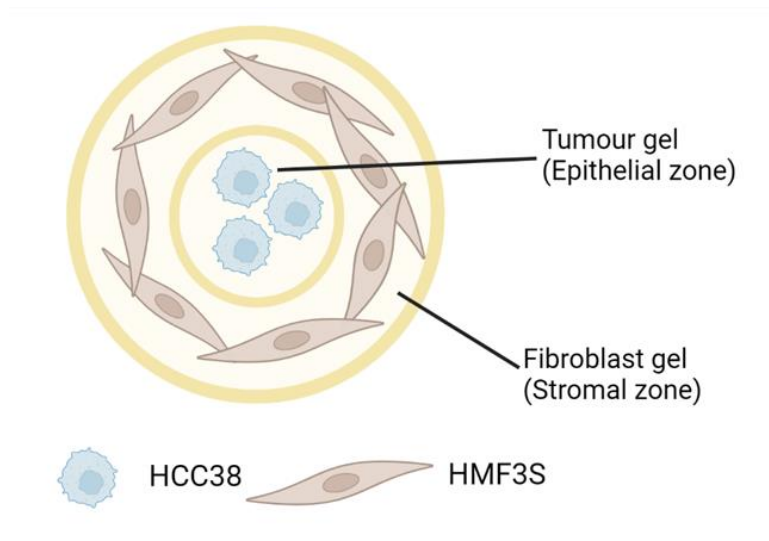
The treatment of HMF3S cells with HCC38 conditioned media did not seem to have a significant effect on the level of the different isoforms. At day 7 a reduction was seen in the V1 isoform and an increase in the V3 isoform, at day 14 the levels reverted to the same as day 3. TGF $\beta$ 3 did not increase the levels of VCAN isoforms at day 3 or day 7 but there was a significantly higher level at day 14 in comparison to HMF3S alone and with conditioned media (Figure 6.3). From this analysis TGF $\beta$ 3 did not seem to increase the expression of VCAN, but did appear to maintain VCAN production over longer periods of two weeks.



**Figure 6.3. Effect of treatment of HMF3S cells towards VCAN expression.** HMF3S cells were grown in a in a collagen gel and treated with HCC38 conditioned media (HCC38M) and TGF $\beta$ 3. Untreated HMF3S were used as a control. The RNA was extracted at day 3, 7 and 14. qRT-PCR was carried out for the VCAN isoforms (V0-V4) and normalised to RPS13 expression. N=3

### 6.3. Building a tumour-stroma 3D model.

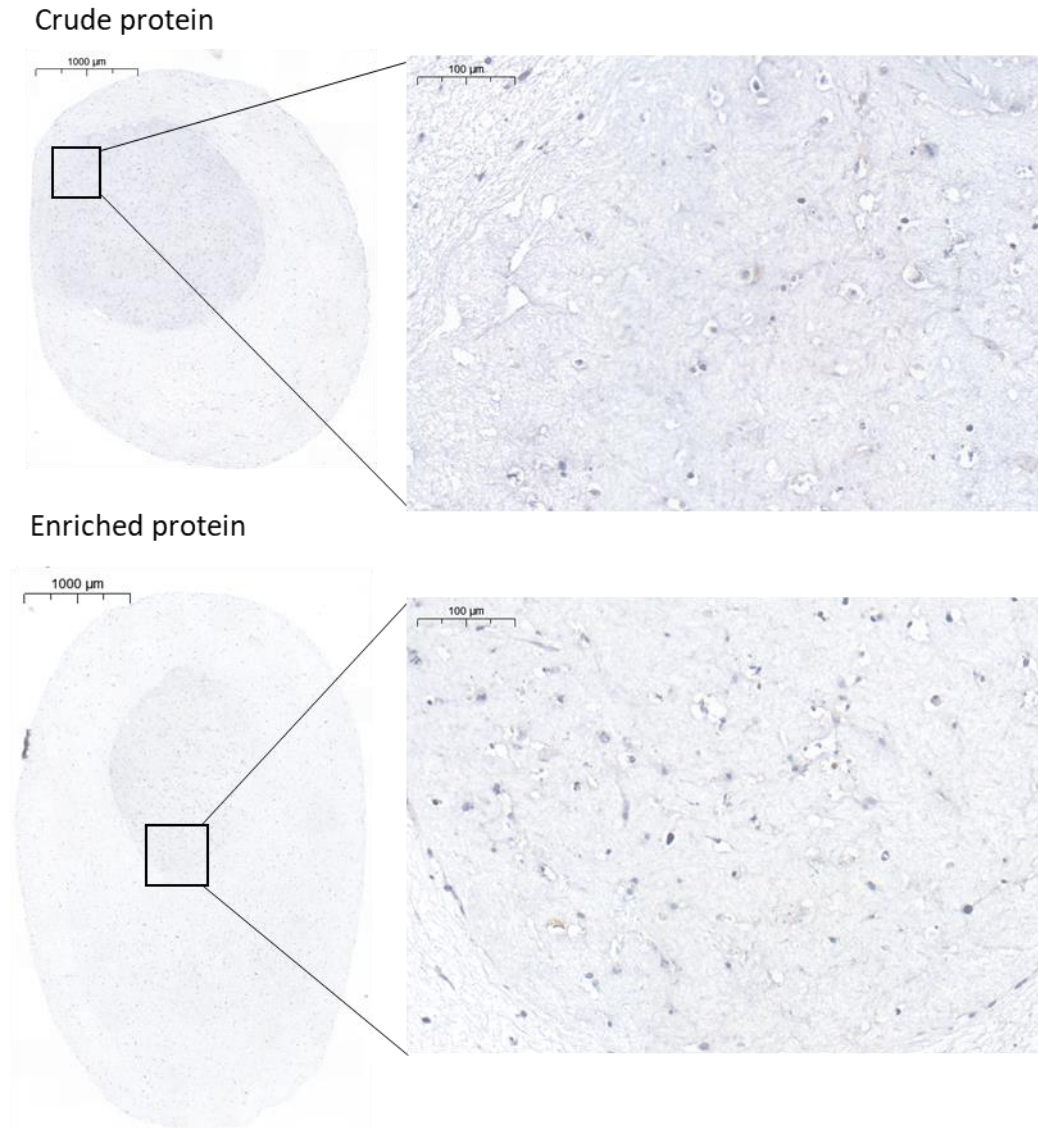
From the RNA data it was observed that for the highest VCAN production by HCC38 was up to day 14 day, and for HMF3S cells a culture of up to 7 days. Taking this into account a model was designed to try to replicate what was observed in the tissues. The aim of this model was to have both the epithelial zone and the stromal zone to mimic what was observed at the tumour margin. To achieve this a gel representing the epithelial zone with HCC38 cells was embedded into a fibroblast gel containing HMF3S cells to represent the stromal zone (Figure 6.4).



**Figure 6.4. Schematic of the tumour-stroma 3D model.** Tumour gel containing HCC38 cells is embedded into the stromal gel containing the HMF3S cells. (Created with BioRender)

For the gels to replicate both excluded and inflamed tissues, we hypothesized the model representing the excluded tissue required a higher level of VCAN in the epithelial zone gel in comparison to the model representing the inflamed tissue. To increase the level of VCAN in the tumour gel, 50µg of VCAN enriched protein was added to the gel. To control for the effects of adding protein, unenriched crude protein was added to the other model. The tumour cells were left to grow for 7 days before embedding in collagen containing 600,000 HMF3S cells. The gels were incubated for 7 days before fixing and embedding in agarose followed by paraffin embedding and sectioning. The gels were stained for VCAN but only low levels were detected. Comparison of the gel with crude protein vs enriched protein showed a comparable amount of VCAN (Figure 6.5). This might be due to the enriched protein getting diluted in the gel and the concentration of protein not being high enough. In order to increase the amount of VCAN detected more protein would

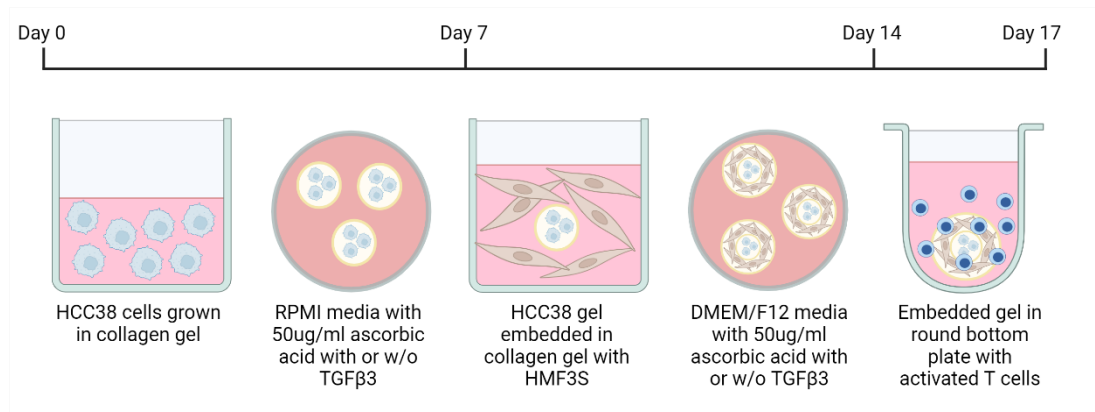
need to isolated and concentrated to a high level. This level of isolation was not feasible and therefore to try to overcome this issue, the protocol was adapted to try to increase the amount of VCAN produced by the cells.



**Figure 6.5. Addition of crude and enriched protein to tumour gel.** HCC38 cells were grown in a collagen gel with either crude or VCAN enriched protein for 7 days. The gel was embedded into a gel with HMF3S cells and grown for 7 days. The gels were fixed and paraffin embedded. Sections were stained were VCAN.

TGF $\beta$ 3 was found to be able to increase VCAN expression by HMF3S so it was predicted that it may have the same effect on VCAN expression by HCC38 cells. The effect was tested by comparing gels with and without TGF $\beta$ 3 in the media when culturing the tumour gel and the fibroblast gel. T cells were then seeded on top of the gels to test the level of infiltration. This is outlined in Figure 6.6. This led to 4

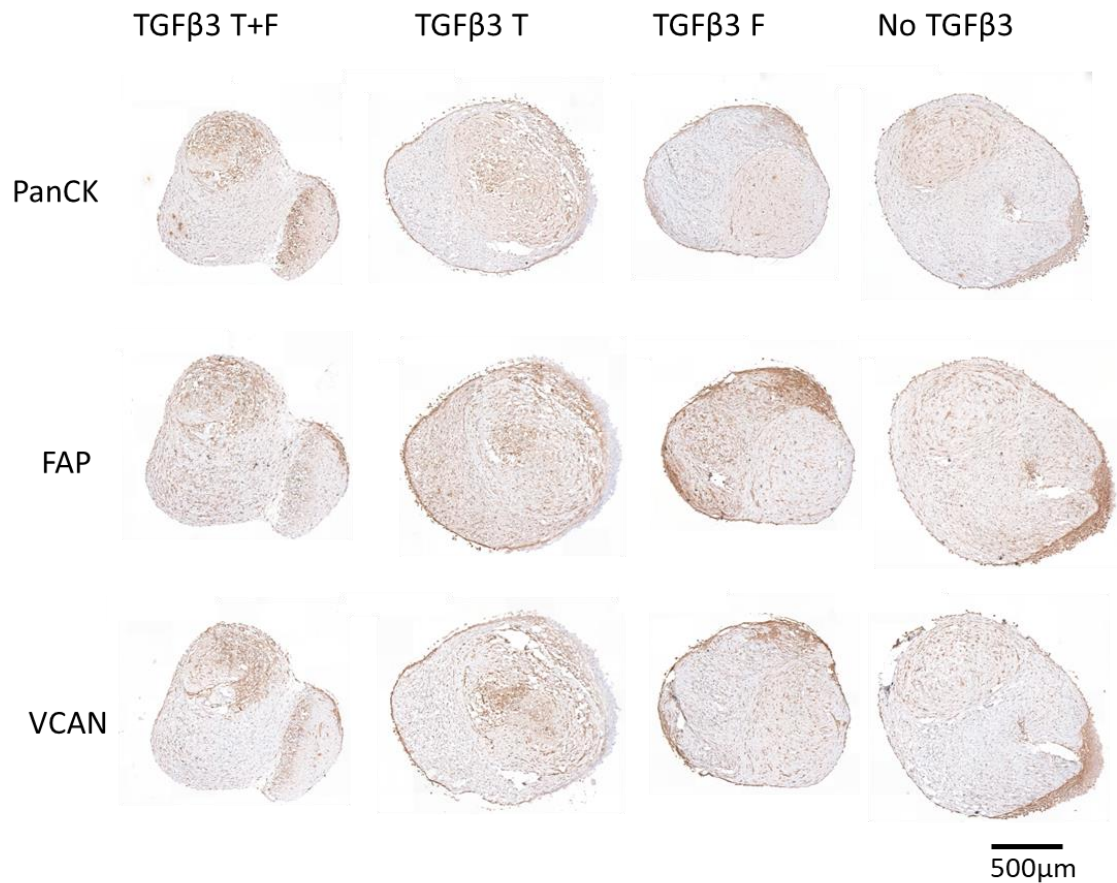
conditions: no TGFβ3; only the tumour gel was exposed to TGFβ3; only the fibroblast gel was exposed to TGFβ3, or both gels were exposed to TGFβ3.



**Figure 6.6. Procedure for growing embedded gels.** HCC38 cells were grown in a collagen gel and grown in media with 50ug/ml ascorbic acid with or without 10ng/ml TGFβ3. The gels were embedded into a gel with HMF3S cells. These gels were grown in media with 50ug/ml ascorbic acid with or without 10ng/ml TGFβ3. The gels were then put in a round bottom plate and 200,000 activated T cells were seeded on top to invade.

The gels were fixed and paraffin embedded for sectioning to be stained by IHC. Gels were made in triplicate for each condition, however during sectioning some gels were found not to be in the same plane and therefore did not get sectioned. Gels were stained with PanCK and FAP to detect the tumour cells and fibroblasts. HCC38 cells were found to be invasive with some cells detected in the fibroblast gel or at the edge of the whole model. The HCC38 cells did not grow in clumps as seen in 2D cultures but invaded out when dividing. This may be due to the limited media available to the cells within the gel after constriction occurs. Necrotic HCC38 cells were also detected within the centre of the gel. A similar pattern was seen with the HMF3S cells where they were detected bordering around the edge of the gel. The level of invasion did not seem to be associated with the TGFβ3 treatment with invasion visible in all gels (Figure 6.7).

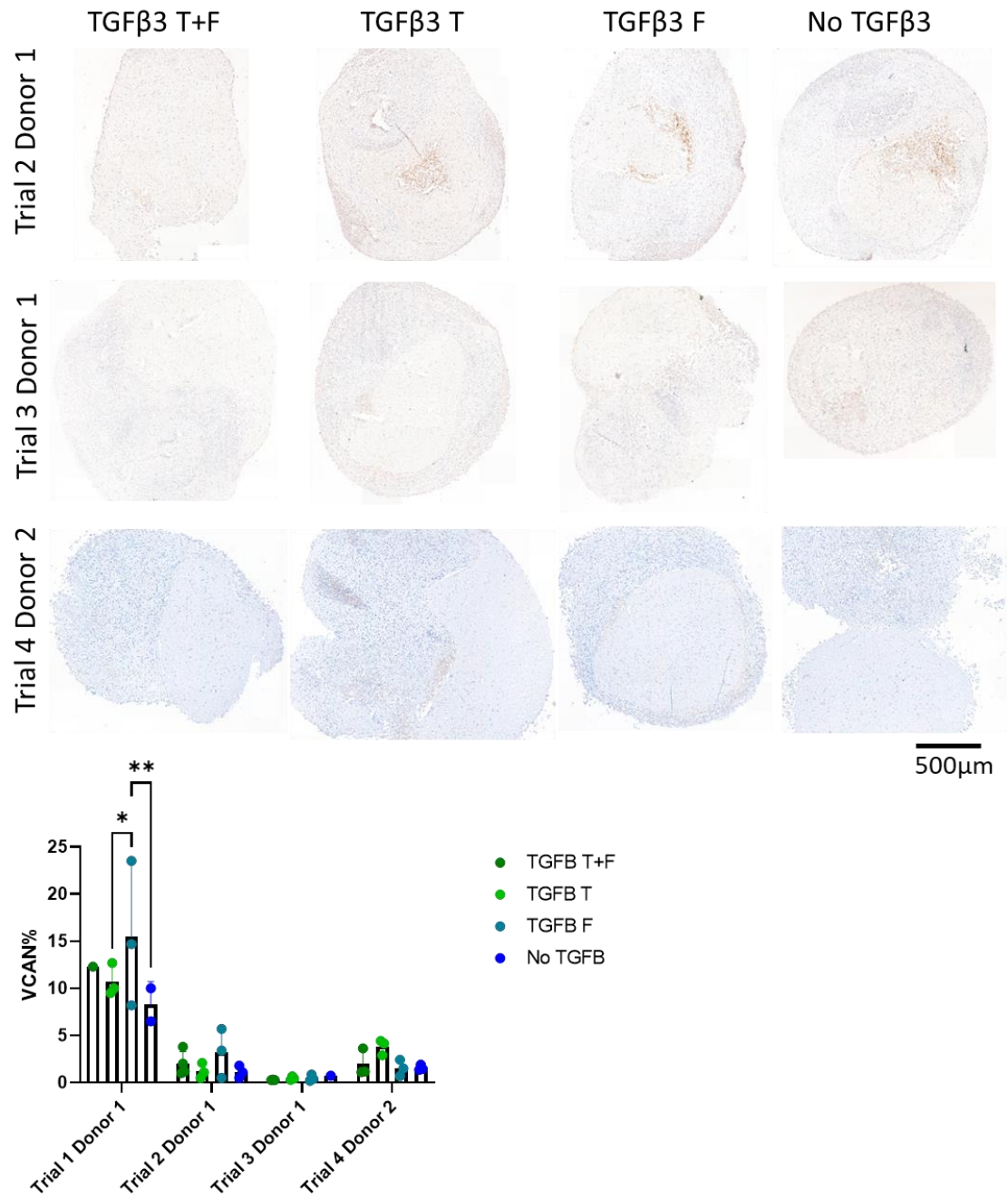
From the VCAN staining it was observed that VCAN remains around the cells following secretion and does not form matrix structures as observed in the human tissues (Figure 6.7). With HCC38 and HMF3S cells invading between gels, there was no representation of the epithelial zone or stromal zone, so the analysis could not be completed as desired, however I could still calculate VCAN abundance for the whole gel.



**Figure 6.7. Location of tumour cells, fibroblasts and VCAN within the embedded gels.** Gels were paraffin embedded and sectioned. IHC was completed for PanCK (tumour cells), FAP (fibroblasts) and VCAN.

The experiment was repeated 3 times with the same T cell donor (Trial 1-3) to form biological repeats of the gels. A second donor was then tested (Trial 4) to see whether there would be a donor specific response. In the repeats of the experiment the levels of VCAN were seen to reduce. In the first trial a significant increase in VCAN levels was only observed in the model where the only the fibroblast gel was exposed to TGFβ3. In the other trials an increase was seen with TGFβ3 treatment but it was not significant. This may have been due to culture conditions or possible effects of the passage of the cells (Figure 6.8).





**Figure 6.8. VCAN expression in the different gel conditions.** Gels from the repeats were paraffin embedded and sectioned for IHC staining. IHC was completed for VCAN. Quantification of VCAN completed using QuPath. 2-way ANOVA. \*=P<0.05, \*\*=P<0.01.

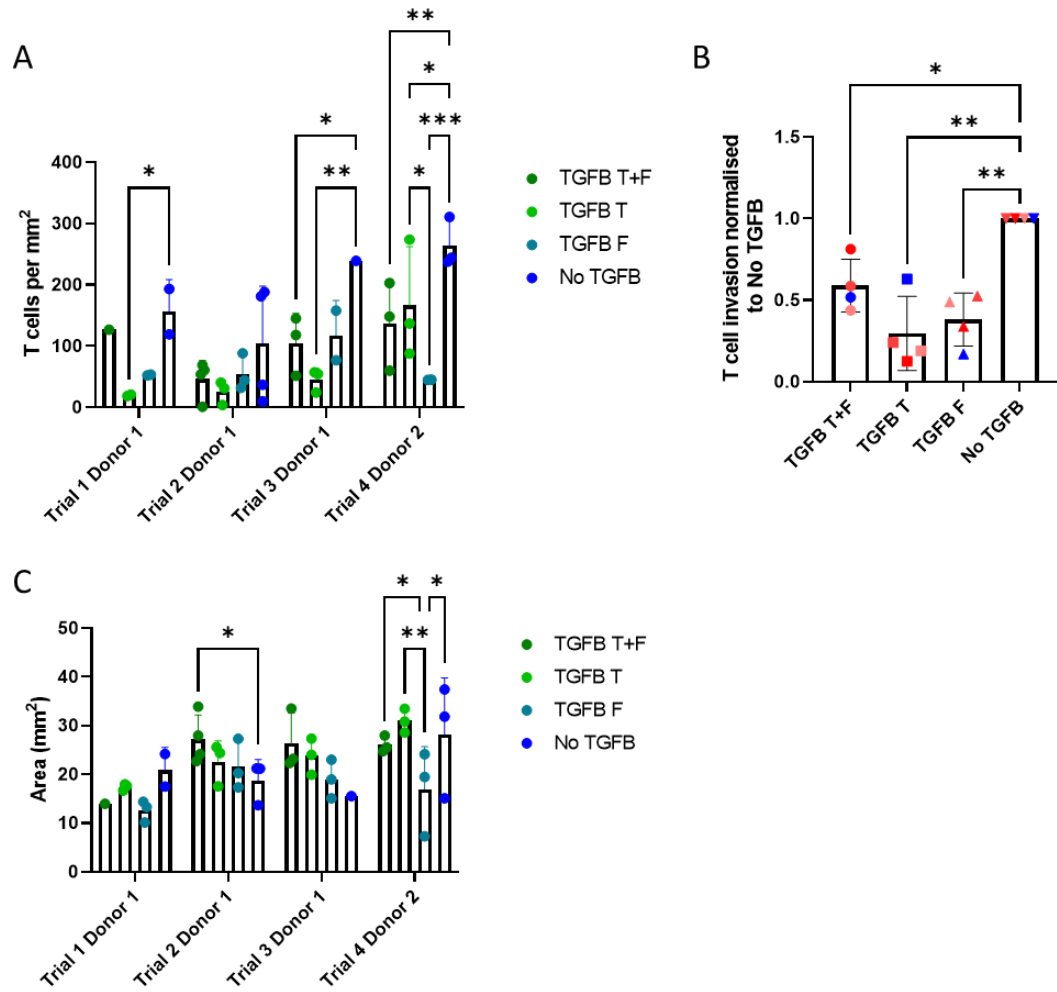
As mentioned previously the HCC38 cells invaded out of the tumour gel into fibroblast gel so an epithelial core could not be defined. Therefore as with the VCAN staining the T cell count was made for the whole model as opposed to a separate value for the tumour gel and the fibroblast gel. CD45 was used to stain for the T cells as both CD4+ and CD8+ T cells were used. CD3 could not be used as a marker as it was used to activate the cells. Models unexposed to TGFβ3 were found to have the highest level of T cell invasion while the exposure of only the tumour gel to TGFβ3 led to the least infiltration. This led to the suggestion that the treatment of HCC38 cells with TGFβ3 may be stimulating the cells to produce higher levels of specific proteins or may be affecting the immunogenicity of the cells. In the trial with the second T cell donor, TGFβ3 was still found to limit T cell invasion. Similar levels of T cells invaded into the models where the tumour gel was exposed or both the tumour and fibroblast gel were exposed. The least infiltration was observed where only the fibroblast gel was exposed. The overall number of T cells that invaded into the gels were greater in the second donor (Figure 6.9A). This could be associated with the level of activation of the gels.

The number of T cells invaded into the TGFβ3 treated models were normalised to the number invaded into the untreated models (no TGFβ3). This showed that overall, TGFβ3 reduced T cell infiltration. No significant differences could be observed between the treatment conditions (Figure 6.9B).

The area of the gel could affect the number of cells invading as more constriction of gels could lead to less space for T cells to invade through. Significant differences were only observed in trial 2 and 4. In trial 2, the model without TGFβ3 had a significantly lower area than the model that was exposed to TGFβ3 twice. This was the opposite of what was expected. In trial 4, the gel with TGFβ3 in only the fibroblast gel media was significantly smaller than the other gels. This might explain the low T cell infiltration seen in this gel (Figure 6.9C).

From the VCAN staining it could not be concluded that VCAN was the matrix component leading to this barrier for T cell invasion. Overall from the experiment it could be concluded that TGFβ3 treatment leads to the expression of proteins in the matrix that can effect T cell invasion, which correlated with VCAN expression in the model.





**Figure 6.9. Invasion of T cells into the embedded gel model.** A) CD45 staining was completed on sections of the gel to detect the T cells. Number of cells counted with QuPath. 2-way ANOVA. B) Number of cells in each condition normalised to the number of cells invaded with no TGFB $\beta$ 3. RM-One way ANOVA. 4 technical replicates, 2 biological replicates. C) The area of the gel was determined using QuPath. 2-way ANOVA. \*= $P < 0.05$ , \*\*= $P < 0.01$ , \*\*\*= $P < 0.005$ .

## 6.4. Effect of VCAN and T cell invasion

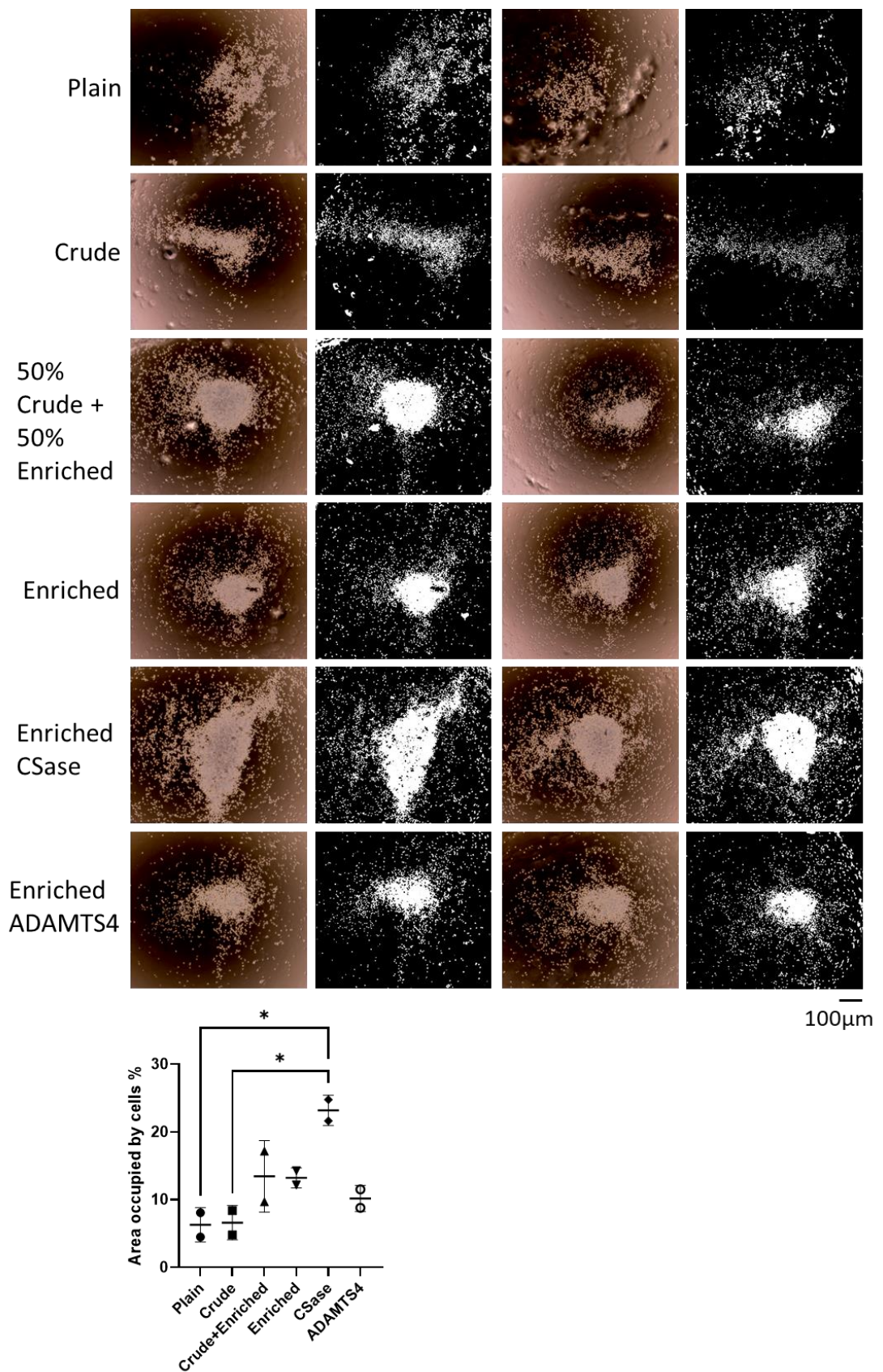
### 6.4.1 Invasion in a transwell model

To further examine whether VCAN was having an effect on T cell invasion, a transwell model was used to look at the direct effects of VCAN towards T cells. The VCAN enriched protein described in the previous chapter was used, and the unenriched crude preparation (which contains approx. 10-fold less VCAN than the enriched preparation) was used as a control. Enzyme treatments to degrade CS (CSase) and the peptide structure (ADAMTS4) were used to determine the effects that different parts of VCANs structure has on T cell trafficking.

Initial experiments were completed using 3µm transwell inserts in 24 well plates. The protein was found not to stick to the well properly therefore it was added into a collagen gel and the gel was set onto the transwell. A 50µl gel with 20µg protein was set onto each transwell. From literature, it has been shown that T cells can migrate across HA strands and that VCAN binding to HA limits migration<sup>162</sup>. Therefore HA was added to the gels at an equal concentration to the collagen. To direct migration through the gel 100ng/ml SDF1 was added to the media in the bottom of the well. 100,000 activated T cells were seeded onto each well. A plain gel was used with just collagen and HA to be able to observe what effect is caused by the crude or enriched protein. Crude and enriched protein were taken from HCC38 samples. A 50:50 crude and enriched protein condition was also tested to see if the addition of enriched protein was dose dependent. CSase treatment and ADAMTS4 treatment were also completed on the enriched protein.

One of the limitations of using transwell inserts was the time point for ending the experiment. If the time point selected was too early, the level of invasion may not have reached a measurable level. The number of cells in the bottom of the well were checked under the microscope after 1hr, 6hrs and then 24hrs. After 24hr it was observed that T cells had invaded through the transwell so the experiment was ended at this point. It was not possible to count the invaded cells by staining the transwell as T cells do not adhere to the membrane. Counting the number of cells in the wells also was difficult as the cell count was low. Alternatively, to quantify the invaded cells the bottom of the well was imaged and the area occupied by the cells was interpreted as the amount of T cells migrated. The images of the bottom well were transformed in image J to form a binary image to highlight cells as white pixels, which could then be quantified to determine the area of occupied by the cells. In comparison to the plain gel, the addition of crude protein showed no effect

on T cell migration whilst the enriched protein increased migration. The effect of the enriched protein was not dose dependent, with equal levels seen with 50% enriched. When the enriched protein was treated with CSase there was a significant increase in the level of invasion (Figure 6.10), suggesting that the CS chains may be impeding migration. Degrading VCAN with ADAMTS4 reduced invasion (Figure 6.10), indicating that the peptide structure of VCAN is important for T cell migration. Overall, it was found that the data contradicted what was seen in the literature<sup>162</sup> where adding protein to HA increased T cell invasion.

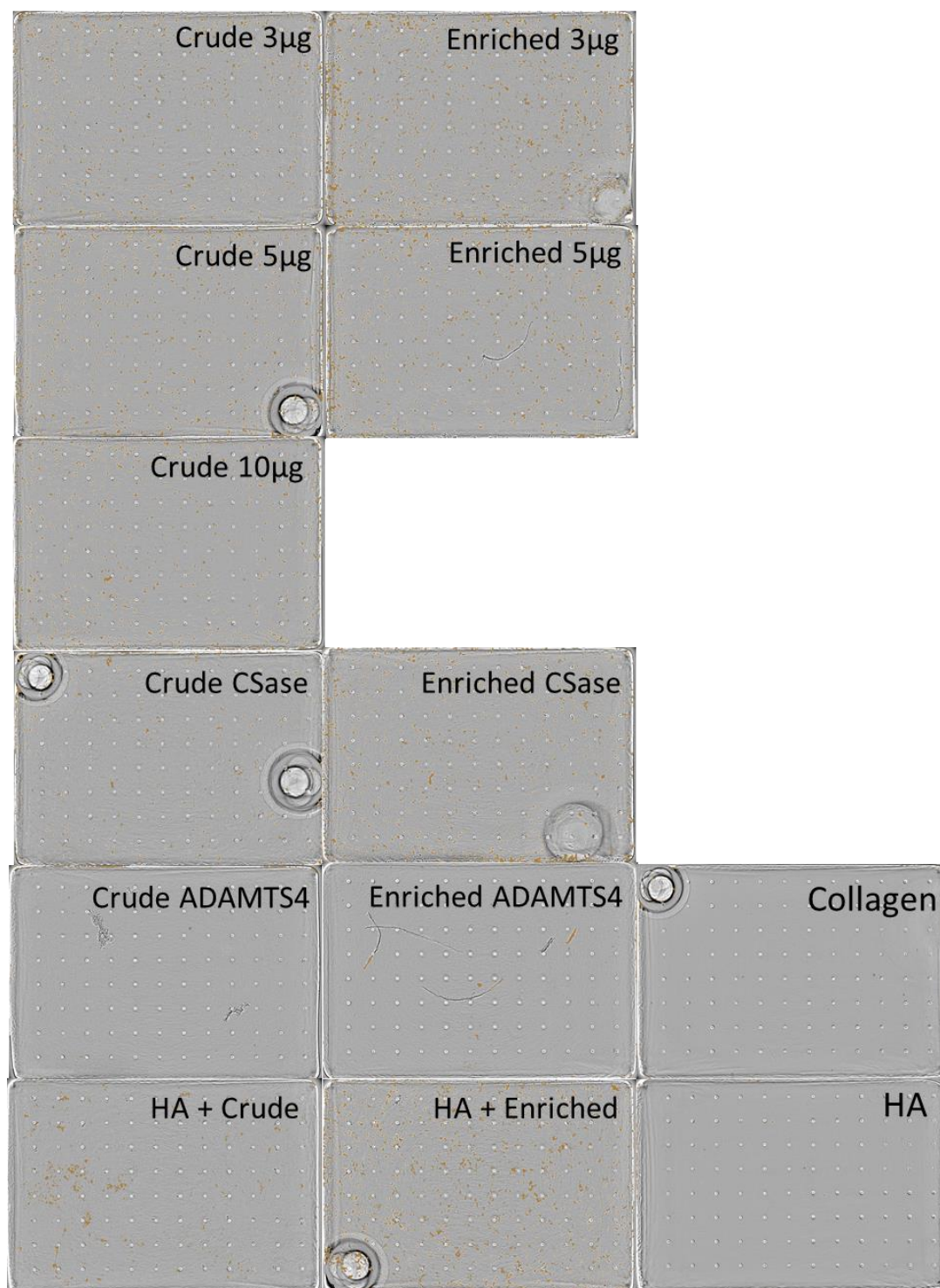


**Figure 6.10. T cell migration to the bottom well of the transwell.** Images of T cells migrated to the bottom well of the transwell taken after 24hrs culture. The images were made into binary with image J and the area of white pixels determined to represent area occupied by cells. 2 technical replicates, 1 biological replicate. One-way ANOVA.  $*=P<0.05$ .

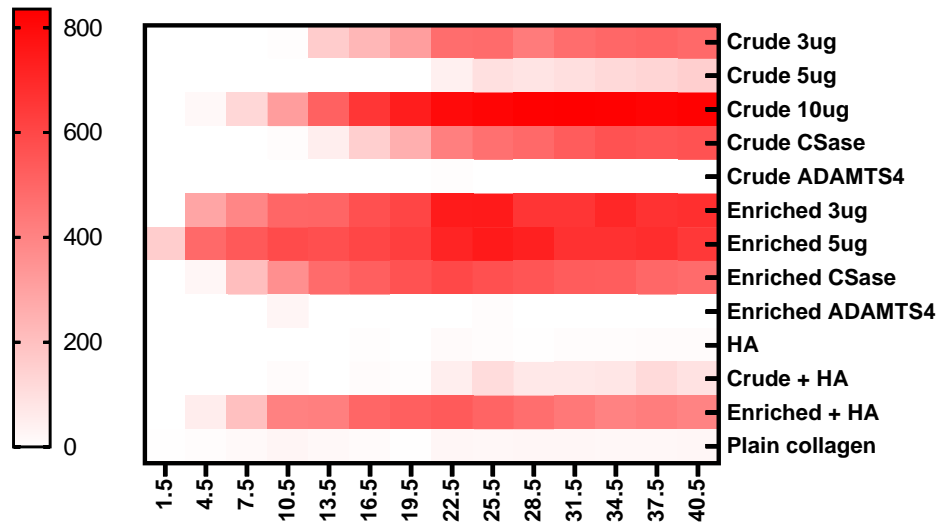
The transwell assay was then repeated using chemotaxis plates. The 96 well chemotaxis plates have a pore size of 8µm but can be imaged on the incucyte, therefore allowing the tracking of T cell invasion over a time course instead of a single endpoint as well as the ability to count the number of cells more accurately. 15µl gels were set on top of the transwell, with different concentrations of the crude or enriched proteins. Both the crude and enriched were treated with CSase and ADAMTS4 to test whether the CS or protein structure were involved in the difference in T cell trafficking as seen previously. To confirm what was seen in the previous assay with HA, conditions were made with HA only and HA with crude or enriched protein. 100ng/ml SDF1 was also added to the media in the bottom wells. 3 wells were made per condition. Images were taken of the membrane of the transwell every 1.5hrs. After 2 days no T cells were detected in the bottom well but the number of T cells at the membrane increased over time as they invaded through the collagen gel. This indicated that the cells were not able to migrate through the pores. Possibly coating of the bottom well was required to facilitate migration, however we decided to simply measure the migration through the gel instead since this approach allows the same analysis to be conducted. The cells on the membrane are shown in Figure 6.11 highlighted in yellow.

The heatmap (Figure 6.12) shows the number of cells at each timepoint between crude and enriched and enzyme treated preparations. For the crude protein a reduction in T cell invasion was seen from 3µg to 5µg, an increase was then observed at 10ug. Increasing the concentration of enriched from 3µg to 5µg improved migration (Figure 6.12). This showed that more protein increases migration levels. Enzyme treatment was done on 3µg of crude and enriched protein preparations, and in comparison to the non-treated protein a slight reduction in invasion was observed. ADAMTS4 treatment stopped any invasion of T cells. The addition of HA had no effect on T cell invasion, but then the addition of 3µg crude and 3µg enriched protein to the HA increased invasion (Figure 6.12) further confirming that the increase in protein concentration increases migration levels.



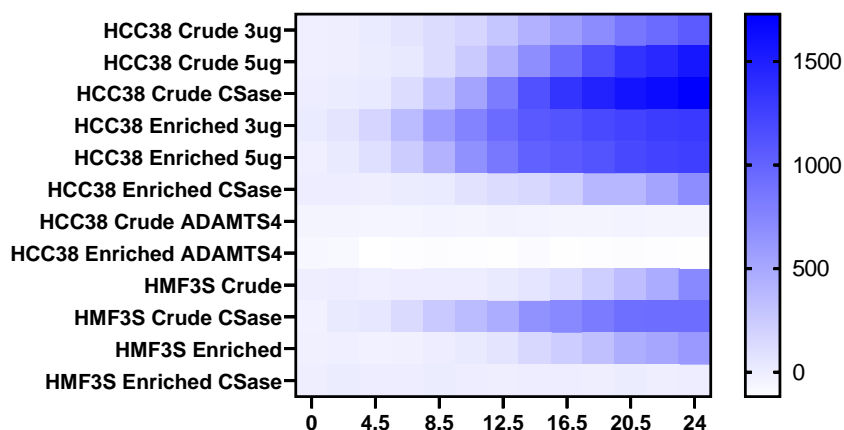


**Figure 6.11. Cells on the transwell membrane.** Images of the transwell taken after 19hrs from the incucyte. Incucyte analysis software used to detect T cells. Positive detections shown in yellow. 1 biological repeat, 3 technical repeats.



**Figure 6.12. Effect of protein level, enzymes and HA on T cell invasion.** Number of T cells at the membrane counted every 1.5hrs. Incucyte software used to count the cells. Gradient represents number of T cells. 1 biological repeat, 3 technical repeats.

The experiment was then repeated again also using protein from HMF3S cells. An increase in protein concentration of the HCC38 enriched protein did not seem to increase invasion levels as previously observed. For the HMF3S sample 3µg of protein was used. Enriched protein from HMF3S increased invasion in comparison to the crude. CSase treatment of enriched protein from both cell lines reduced invasion but there was an increase observed with the crude protein. The effect of ADAMTS4 was the same as the previous experiment (Figure 6.13). From these two experiments, it became apparent that enrichment of protein increases invasion and ADAMTS4 diminishes invasion. The effect of CSase was still not clear with contrasting results seen with the crude and enriched in both trials.



**Figure 6.13. Effect of cell specific protein and enzymes on T cell invasion.** Number of T cells at the membrane counted every 1.5hrs. Incucyte software used to count the cells. Gradient represents number of T cells. 1 biological repeat, 3 technical repeats.

The previous experiment was repeated with the addition of a second enrichment of HCC38 to see if there is a difference in enrichments. The effect of the protein from different enrichments were then tested in addition to the role of CSase. From the initial experiment it was seen that the enrichment of VCAN from HCC38 did not increase invasion contrasting to what was seen with previous assays. This was the case for both enrichments. The enrichment of HMF3S was found to increase invasion. CSase treatment increased the invasion in both crude and enriched samples (Figure 6.14A).

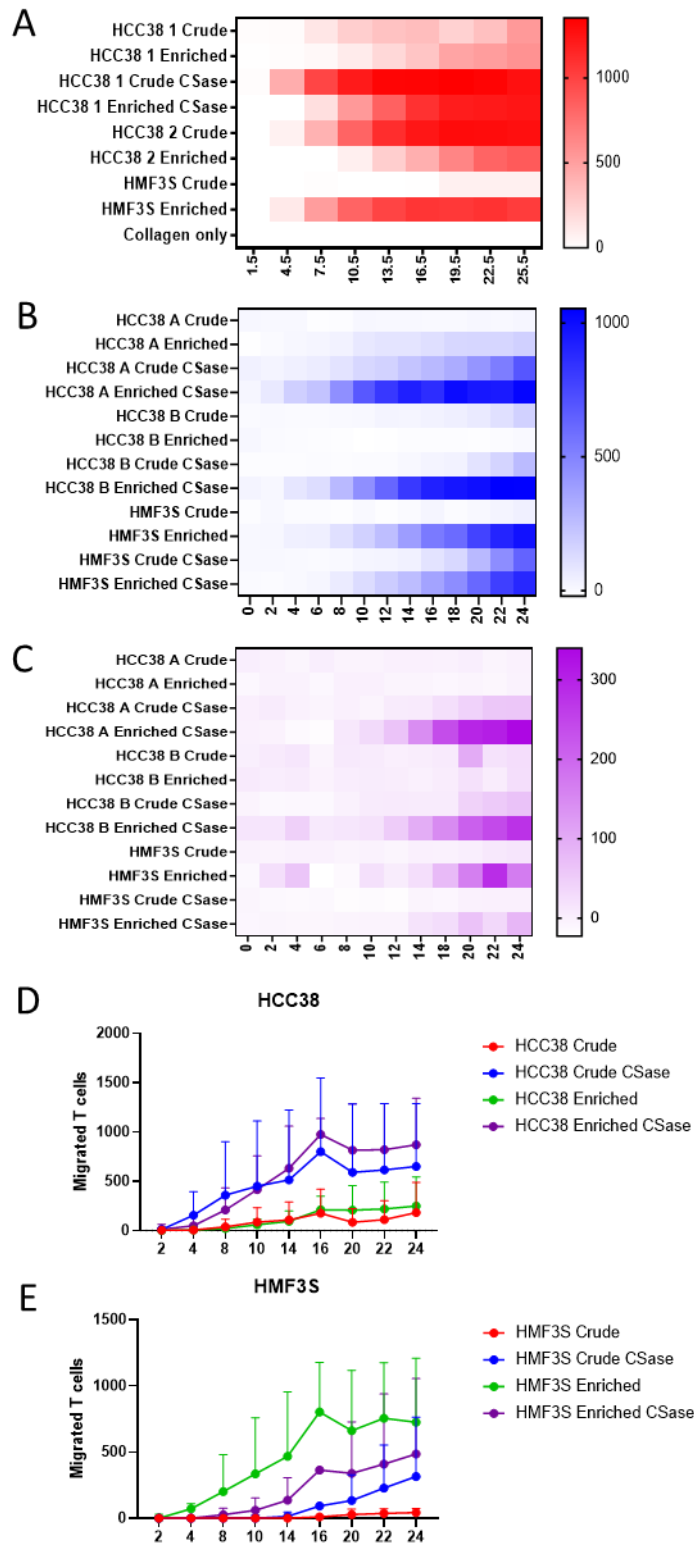
The assay was repeated using 2 different HCC38 samples and also using 2 different T cell donors. Experiments with donor 2 showed a slight increase in invasion with HCC38 enriched sample A. No increase was seen with sample B but very low levels were detected with the crude. The CSase treatment increased the invasion of both the crude and enriched in both HCC38 samples. With the HMF3S sample, a high increase in invasion was observed with the enriched compared to the crude. CSase treatment was seen to not have an impact on either the crude or enriched protein (Figure 6.14B).

The repeat with donor 3 T cells showed much less T cell invasion. This may be due to poor activation of the T cells. The pattern observed was similar to what was seen with donor 2 where the CSase treatment of the HCC38 enriched protein increased invasion and the enriched protein of HMF3S improved invasion but CSase treatment led a slight reduction (Figure 6.14C).

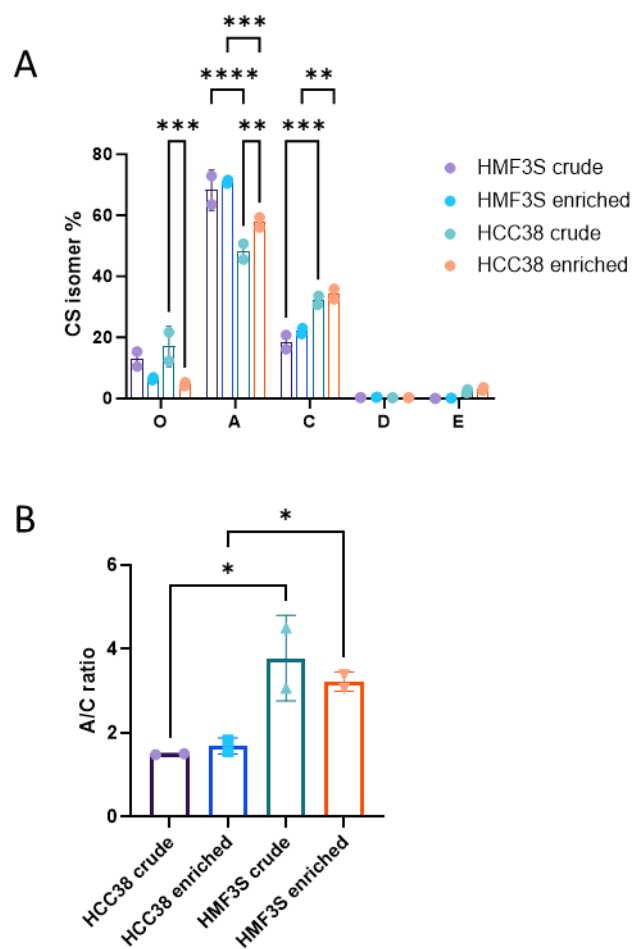


Overall, HCC38 enriched protein had a slight increase in invasion but a greater increase was seen after CSase treatment (Figure 6.14D). HMF3S enriched protein improved T cell infiltration where CSase treatment was seen to cause a reduction in invasion (Figure 6.14E).

With the CSase treatment having different effects on the cell lines, CS isomer analysis was completed on the crude and enriched protein samples used in the transwell model. From the analysis it was observed that there was a lower level of CS-O in the enriched samples, which was significant in HCC38. CS-O disaccharides have no sulphation and therefore the negative charge is very low weakening the binding to the beads during the enrichment process. The CS-A isomer was significantly higher in HMF3S compared to HCC38, whilst the inverse was observed with CS-C where it was significantly higher in HCC38 samples compared to HMF3S (Figure 6.15A). This effect can also be seen in the CS-A/C ratio where the HCC38 samples has a significantly lower ratio than the HMF3S samples (Figure 6.15B). This associates with what was observed in Chapter 4 where the excluded tissues had a lower CS-A/C ratio compared to the inflamed. With the removal of CS with CSase treatment having a greater effect on the HCC38 samples, it suggests that the levels of these CS isomers may be having an impact on T cell invasion. The CS-D and CS-E isomer were detected at low proportions with similar levels of CS-D in the protein from both cell line and a slight increase in CS-E in the HCC38 samples, however there was no significance in the level of these isoforms and less indication therefore that they are involved in T cell invasion.



**Figure 6.14. Invasion with different T cell donors.** 10,000 activated T cells from A) Donor 1. B) Donor 2, C) Donor 3, were seeded onto the collagen gels in the transwell. Collagen gels contained protein with different conditions. T cells were counted and analysed using the incucyte every 1.5-2hrs. Gradient represents T cell count. The average number of invaded T cells from the 3 donors were calculated and graphed for protein from D) HCC38 and E) HMF3S. 3 biological repeats, 3 technical repeats.



**Figure 6.15. CS isomer analysis of the crude and enriched samples of HCC38 and HMF3S.** A) Percentage of the different CS isomers in HMF3S and HCC38 crude and enriched protein. 2-way ANOVA B) A/C ratio of the HCC38 and HMF3S crude and enriched. One-way ANOVA. \*= $P < 0.05$ , \*\*= $P < 0.01$ , \*\*\*= $P < 0.005$ , \*\*\*\*= $P < 0.001$ . 2 biological repeats

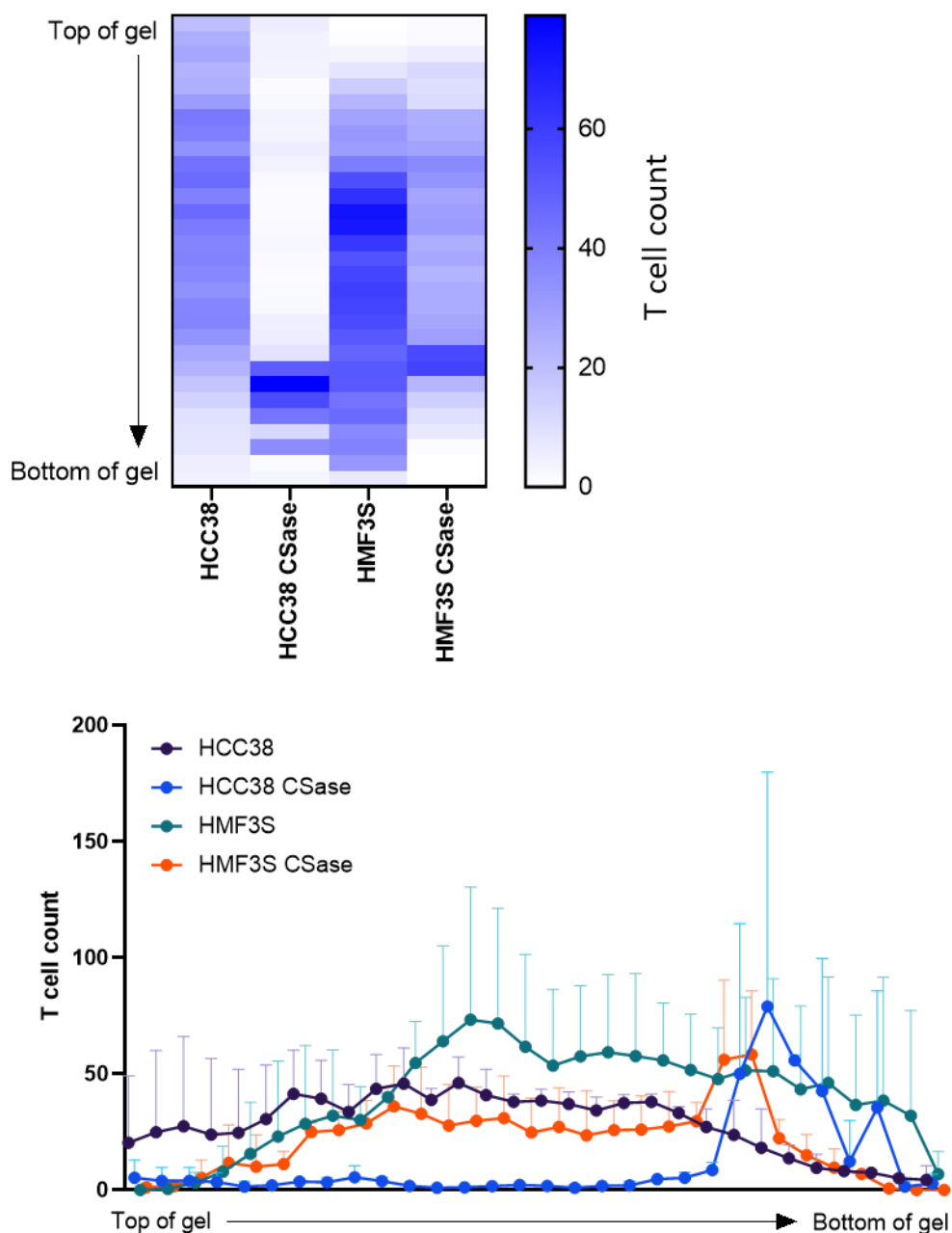
#### 6.4.2 Invasion in a gel

From the transwell it was only possible to look at the T cells that migrated to the bottom of the gel. I then looked at how the T cells were interacting with the collagen gel by counting the cells through the gel. Wells were coated with SDF1 to form a gradient to attract T cells through the gel. A 50µl collagen gel was made with enriched protein and CSase treated enriched protein. CSase buffer was added to all samples due to the pH of the buffer possibly having effects with the cross-linking of the gel. T cells were activated and 25,000 cells were seeded onto the gels and left to invade over 3 days. They were then fixed and stained with DAPI. The gels were scanned across 30 images, 26µm apart. As seen with the transwell model, the CSase treatment of HCC38 enriched protein led to more T cells invading to the bottom of gel, whilst the CSase treatment of HMF3S enriched protein seemed to reduce invasion. Within the HCC38 enriched sample, most of the invaded T cells were found near the top end of the gel. More T cells invaded the gel with HMF3S enriched protein, but most of these cells were found in the centre of the gel. The overall number of cells invaded into the gel looked to be less for both CSase treated samples from the area under the line (Figure 6.16).

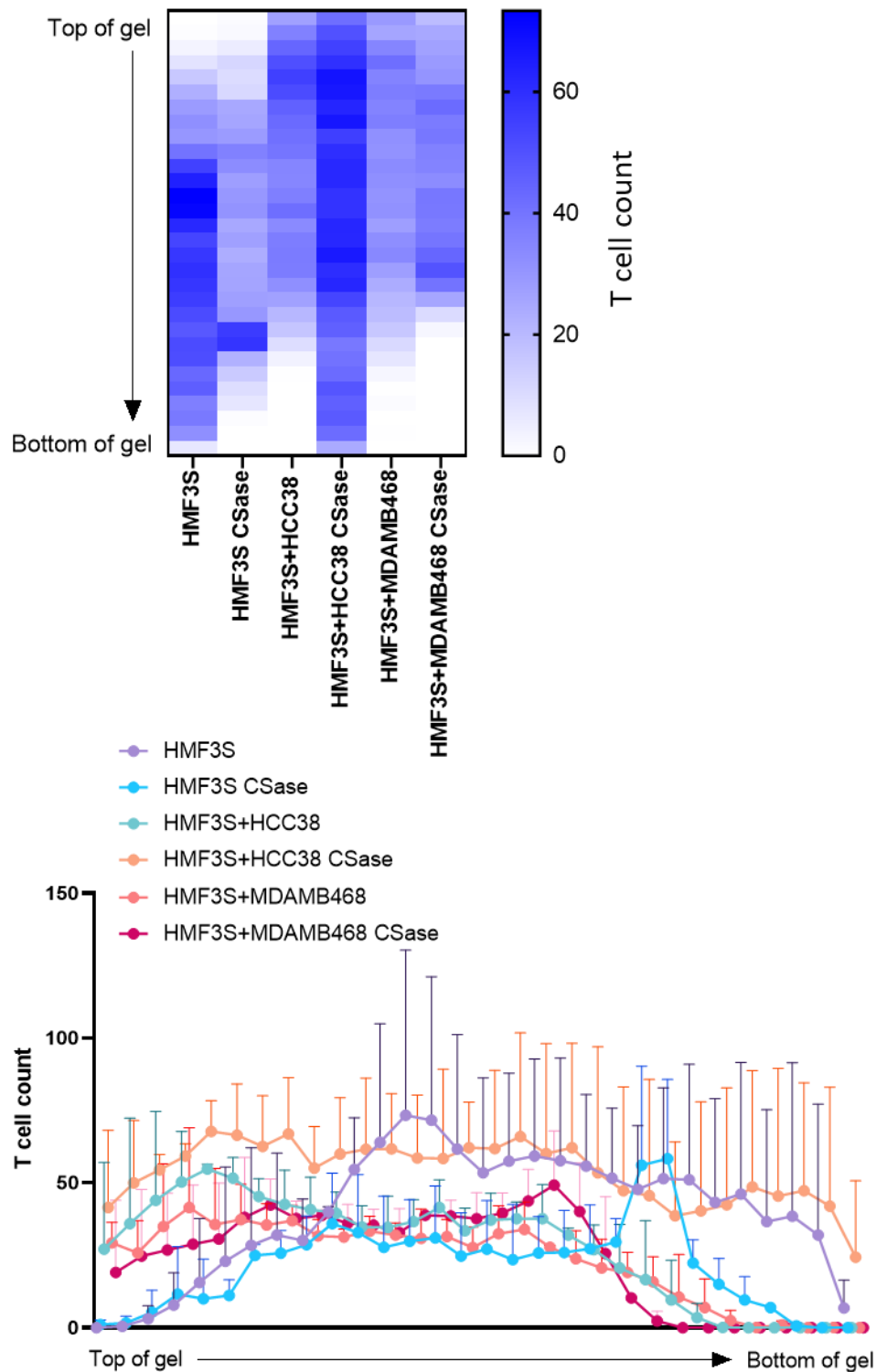
I then questioned whether the differences seen in the HCC38 and HMF3S enriched samples would also be seen with the enrichment following the co-culture of HMF3S with the TNBC cell lines HCC38 and HMF3S. The enriched protein from the co-cultures was seen to have less invasion in comparison to HMF3S alone. Based on the analysis to detect the amount of VCAN in the sample, the most VCAN was present in the HMF3S+MDAMB468 sample (Figure 5.16) and the least in the HMF3S+HCC38 sample but this did not seem to have an effect on invasion. Most of the T cells were found in the top half of the gel. CSase treatment of the HMF3S+HCC38 enriched protein led to T cells invaded to the bottom of the gel as well as more cells invading. No effect was observed following CSase treatment of the HMF3S+MDAMB468 enriched protein (Figure 6.17).

Overall, from these assays I was able to confirm what was seen with the transwell model where the CSase treatment of HCC38 enriched protein can lead to better invasion, but little impact is seen with treatment of HMF3S enriched protein. The amount of VCAN did not seem to have an effect on the invasion from the co-culture samples. Cs isomer analysis was only completed on the crude samples for the co-cultures and not the enriched so correlations could not be made towards this.

Further repeats exploring what the proteins are interacting with in the gel and with different T cell donors will be required to make final conclusions.



**Figure 6.16. Effect of CSase towards invasion through a collagen gel.** 50µl collagen gels with treated and untreated enriched protein from HCC38 and HMF3S set onto a 96 well plate. Wells of 96 well plate were pre-coated with 10ng/ml SDF-1. 25,000 T cells were seeded onto the gel and the gels were fixed after 3 days. Cells were stained with DAPI and images taken at different depths of the gel. The number of cells at the different depths represented as a heatmap and line graph. Gradient on heatmap represent number of cells. 1 biological repeat, 2 technical repeats.

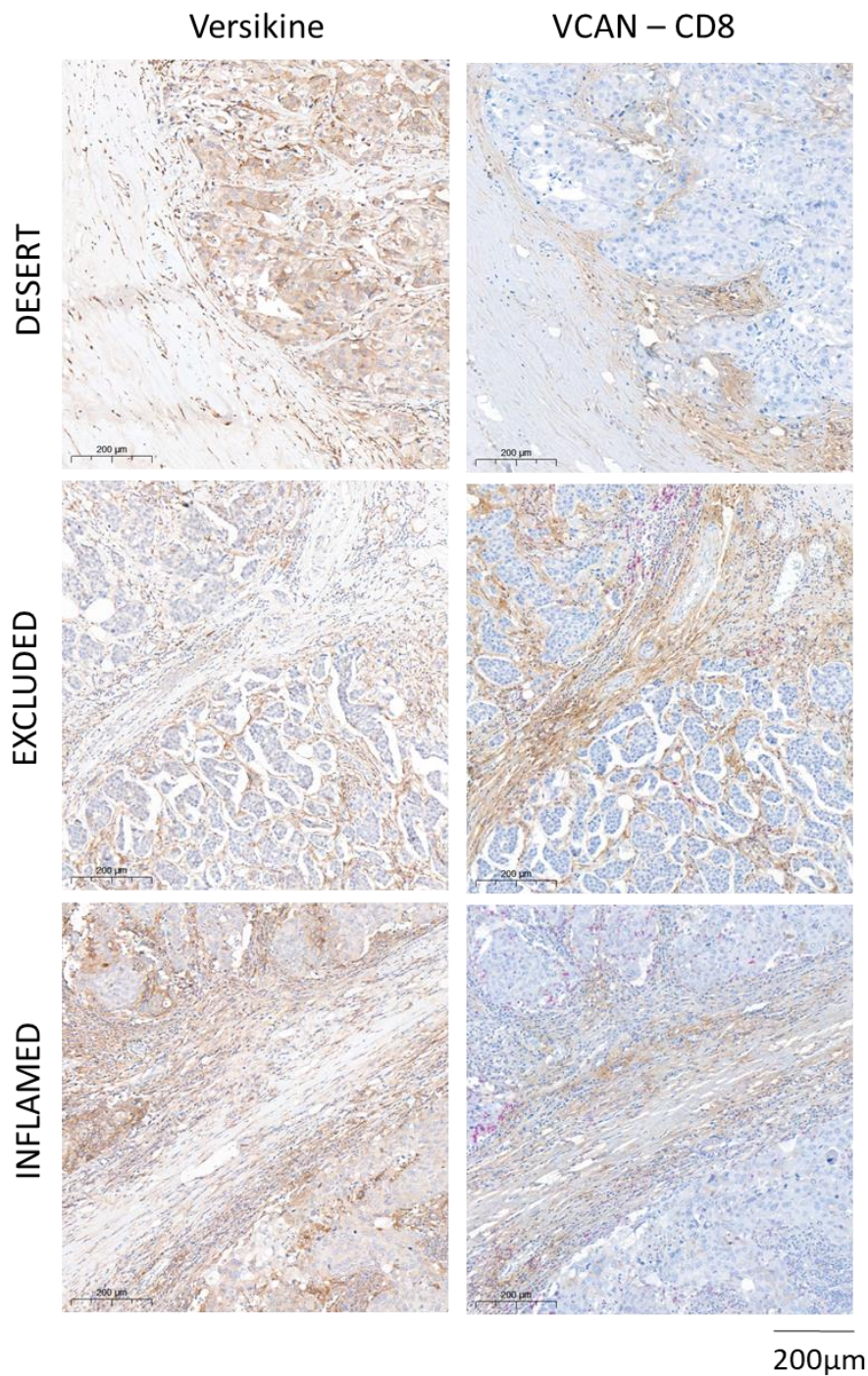


**Figure 6.17 Effect of HMF3S co-culture with HCC38 and MDAMB468 towards invasion through a collagen gel.** 50 $\mu$ l collagen gels with treated and untreated enriched protein from HCC38 and HMF3S set onto a 96 well plate. Wells of 96 well plate were pre-coated with 10ng/ml SDF-1. 25,000 T cells were seeded onto the gel and the gels were fixed after 3 days. Cells were stained with DAPI and images taken at different depths of the gel. The number of cells at the different depths represented as a heatmap and line graph. Gradient on heatmap represent number of cells. 1 biological repeat, 2 technical repeats.

## 6.5. Effect of VCAN degradation towards T cell invasion

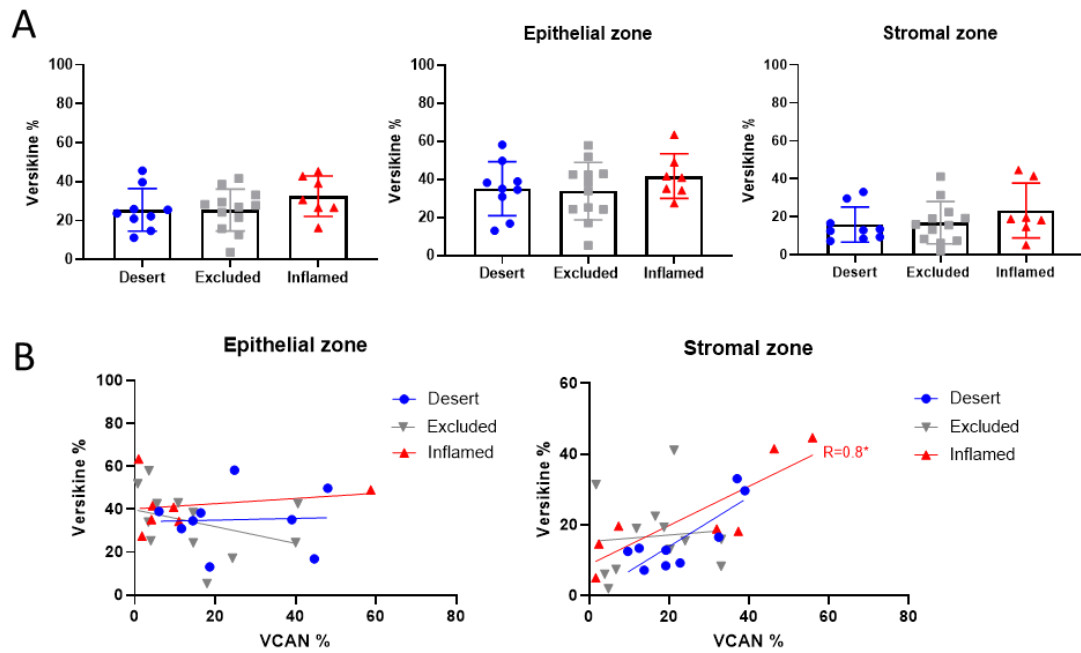
The breakdown of VCAN to versikine is a known factor to effect T cell invasion <sup>20</sup>. From the transwell assays, ADAMTS4 treatment reduced T cell invasion. To explore whether the breakdown of VCAN to versikine is associated with TIPs, the level of versikine was explored between TIPs. Tissues were stained for versikine by using an antibody that targets the epitope (DPEEAE) that is present at the end of the peptide following cleavage. The staining was then compared to the VCAN staining (Figure 6.18). When comparing the level of versikine between the phenotypes and in the different regions, I saw that there was no significant difference between the levels of versikine (Figure 6.19A). The correlation was then determined between VCAN and versikine in the epithelial zone and stromal zone for the different phenotypes. No correlations were found in the epithelial zone. In the stromal zone there a positive correlation ( $R=0.8^*$ ) between VCAN and versikine in the inflamed tissues (Figure 6.19B). With the VCAN antibody detecting the G1 domain which includes the detection of versikine fragments, a positive correlation suggested that a majority of the VCAN detected in the inflamed stroma is versikine.

To try to associate this to the invasion assays completed, the expression of VCAN degrading enzymes were determined. Within most tissues, the enzymes responsible for VCAN cleavage are the ADAMTS enzymes, specifically ADAMTS1, ADAMTS4 and ADAMTS5. These enzymes can cleave VCAN in multiple places including the DPEEAE site <sup>118</sup>. In the cell lines, HMF3S was seen to express the highest level of each enzyme, implying that the enrichment of VCAN from these cells may contain greater levels of versikine in comparison to the other samples Figure 6.20.

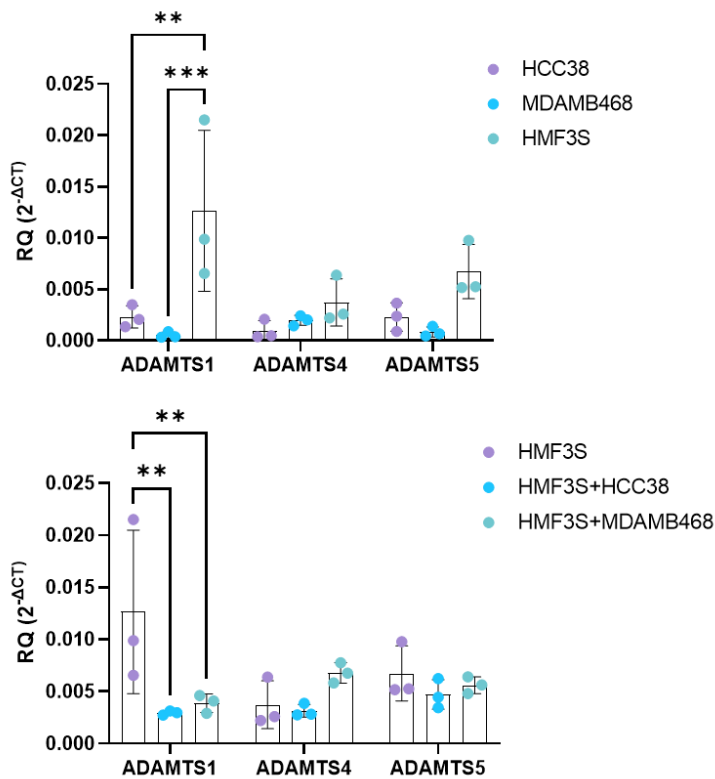


**Figure 6.18. Comparison of versikine and VCAN in the different TIPs.** IHC was completed for versikine and VCAN on 30 TNBC tissues. VCAN-CD8 was completed as a dual IHC. Versikine and VCAN in brown (DAB), CD8 in pink.





**Figure 6.19. Versikine analysis in the different TIPs.** QuPath was used to quantify the level of versikine in the same areas as where VCAN was quantified. The quantification was completed across the tumour epithelial border looking at the epithelial zone and stromal zone. A) Comparison of versikine in the TIPs overall and in the different areas. One way ANOVA. B) Correlation of versikine with VCAN in the different areas and different TIPs. Spearman Rank Correlation.  $^* = P < 0.05$ .

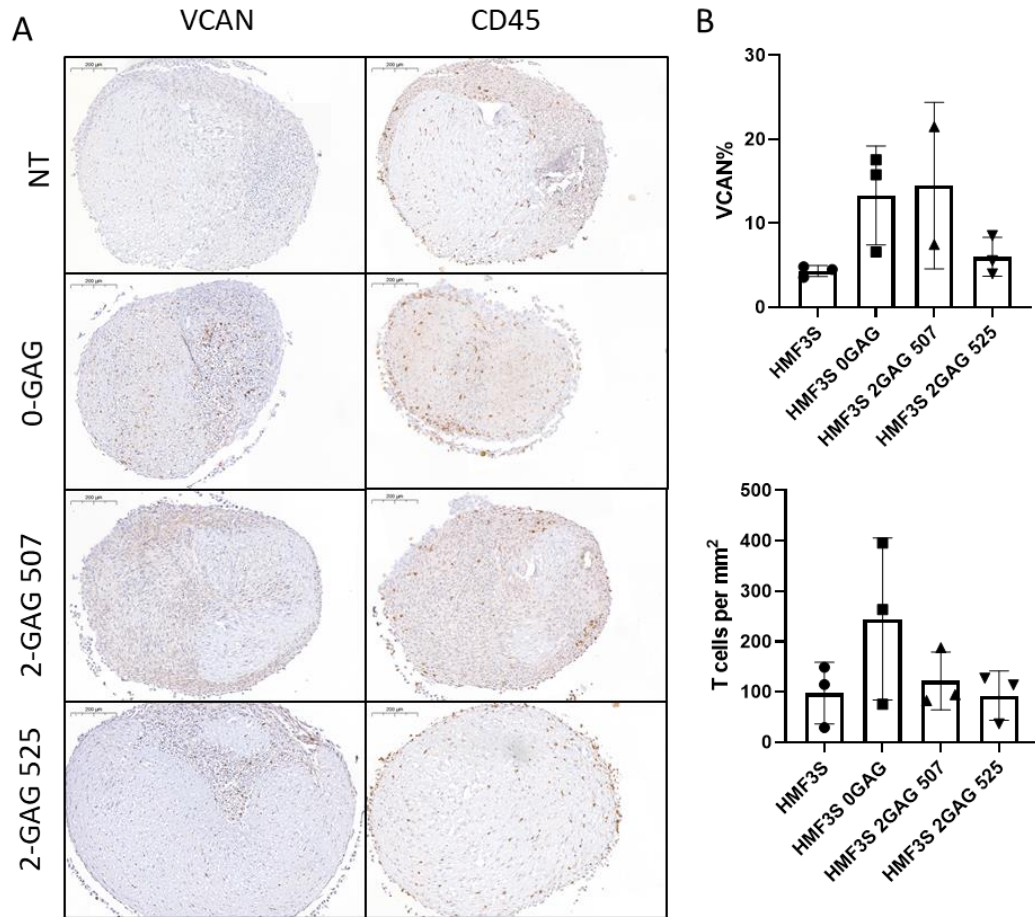


**Figure 6.20. ADAMTS enzyme expression in the cell lines and in co-cultures.** RNA was extracted from TNBC cell lines HCC38 and MDA MB468, as well as the immortalised mammary fibroblast HMF3S grown in mono-culture and in co-culture with the TNBC cell lines (4:1). qRTPCR was completed for ADAMTS1, ADAMTS4 and ADAMTS5 genes. 2-way ANOVA.  $^{**} = P < 0.01$ ,  $^{***} = P < 0.005$ . 3 biological repeats.

To further explore the effects of the different sizes of VCAN towards invasion, cells were transfected with VCAN plasmids which contain either 0 or 1 CS chain. These plasmids were generously given by Dr Salvatore Santamaria. The 0-GAG plasmid replicates the structure seen with versikine, where the protein contains the G1 and  $\beta$ -GAG domain but is truncated before the CS binding sites. The plasmids containing one chain are the 2GAG 507 and 2GAG 525 constructs. These plasmids also contain the G1 domain and the  $\beta$ -GAG domain but are truncated after the first two CS binding sites on the  $\beta$ -GAG domain. The 2GAG 507 plasmid is mutated at the CS binding site S507 and 2GAG 525 is mutated at the CS binding site S525, this means only 1 CS chain can be formed on each of these constructs.

The initial analysis was to look at whether there were differences between having no CS chains and 1 CS chain as well as whether the location of the CS chain has an impact. The transfection was completed in all three cell lines, HMF3S, HCC38 and MDAMB468, however the transfection did not seem to work in the HCC38 cell line. The transfected cells were selected with bleomycin and collagen gels were made with each condition. The gels were incubated for 7 days and then 200,000 activated T cells were added to each gel. The gels were paraffin embedded, sectioned and then stained for VCAN and CD45 (Figure 6.21A). Gels were not embedded and sectioned properly for the MDA MB468 samples so they were not analysed.

From the HMF3S transfected cells it was visible that the gels with the 0GAG construct contracted more as smaller gels were seen. The 2GAG507 construct gel also contracted more compared to the non-transfected (NT) gel. The 2GAG 525 gel was found to be a similar size to the NT gel. The levels of VCAN and CD45 were determined in each gel through analysis using QuPath. The VCAN antibody detects the G1 domain which is present in all constructs. The 0GAG and 2GAG507 gels had the greatest level of VCAN present. The levels of VCAN in the 2GAG525 were comparable to the NT gel, which suggests that the cells did not produce the transfected construct. The number of T cells invading the gel were then compared and it was found that the 0GAG gel had the greatest level of infiltration (Figure 6.21B), showing that VCAN with no CS chains in the form of versikine may be helping T cell invasion.



**Figure 6.21. Effect of CS expression on VCAN to T cell invasion.** HMF3S cells transfected with the 0GAG, 2GAG 508 and 2-GAG 525 plasmids. Gels were paraffin embedded and sections. A) Sections were stained with VCAN and CD45. B) QuPath was used to quantify the percentage area of VCAN and the number of T cells. 1 biological repeat, 3 technical repeats.

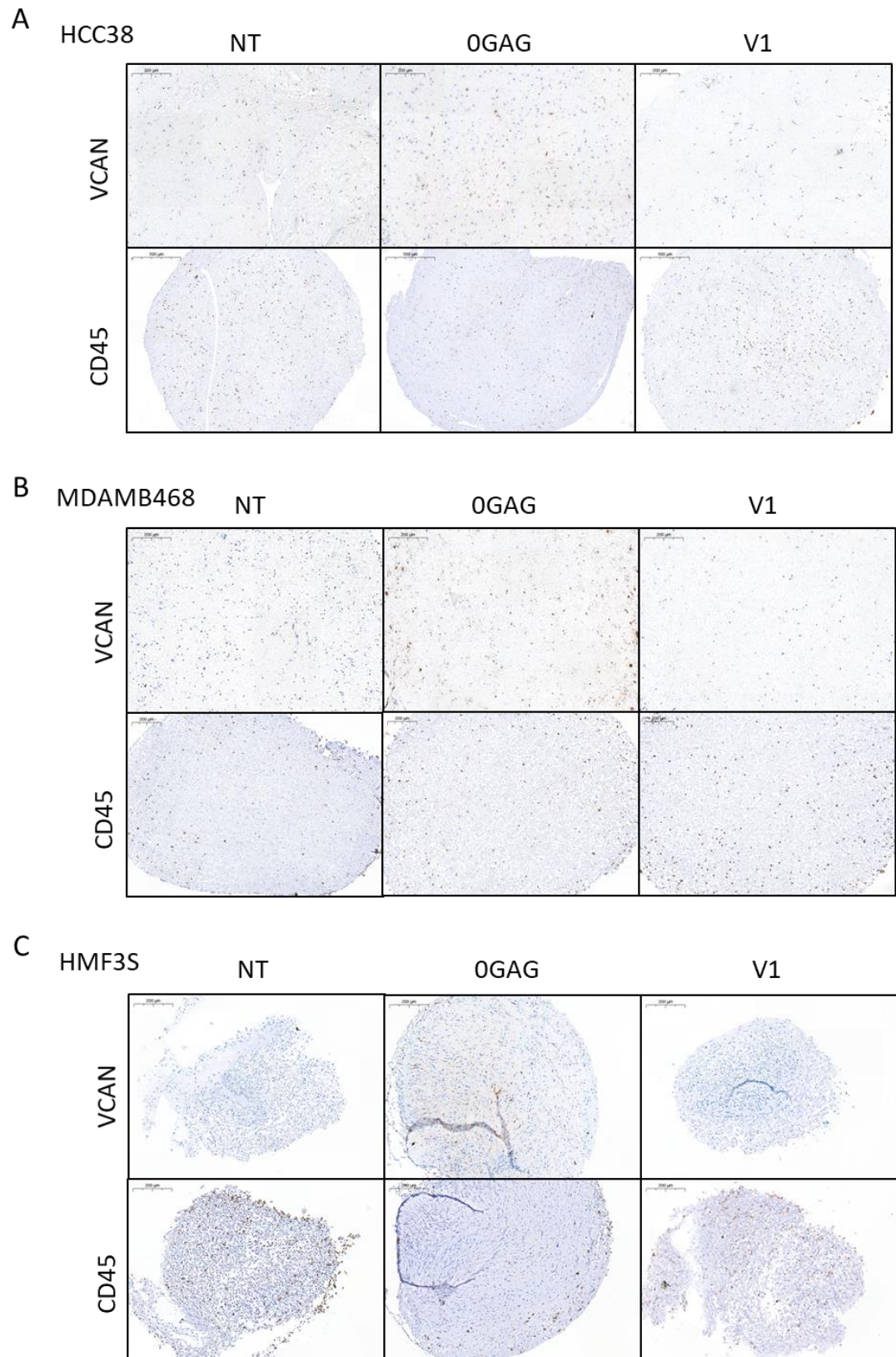
Based on the analysis it seemed that over expression of versikine was promoting T cell migration. This was then further compared to cells transfected with a plasmid expressing the V1 isoform where multiple CS chains will be present. The transfections were completed again on all three cell lines. The cells were made into collagen gels and then treated with T cells and sectioned for staining.

For HCC38 cells, very little VCAN staining was visible and the transfected cells showed the same level of VCAN as the NT (Figure 6.22A). This indicated that the transfection may have not been successful. The two colours represent different T cell donors. Interestingly the cells with the 0GAG transfection had the least T cell invasion (Figure 6.23A) however, little VCAN was being produced suggesting the transfection had not worked well.

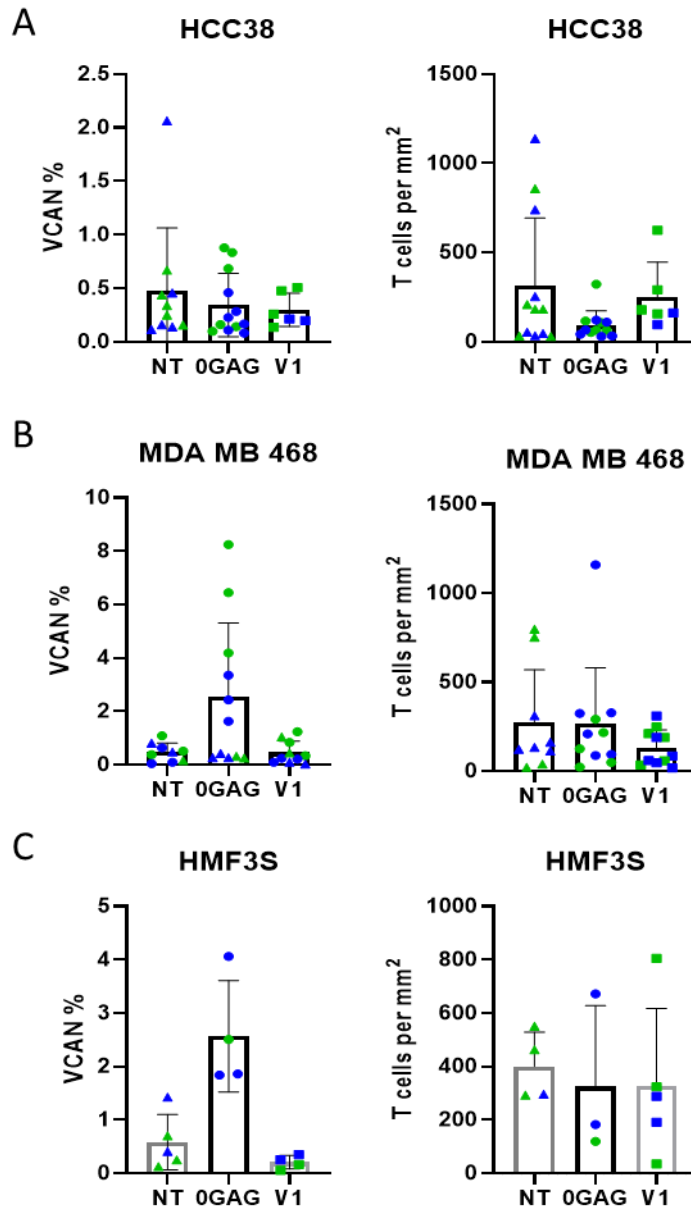
The transfection in MDAMB468 cells only seemed to work with the 0GAG plasmid where the VCAN levels in that plasmid were higher than the NT gels (Figure 6.22B). The levels of 0GAG did not seem to impact T cell invasion levels with a similar level seen in comparison to the NT gels. A lower level of invasion was seen in the gels with V1 transfected gels (Figure 6.23B), however as no V1 was detected it cannot be associated with the protein.

High levels of contraction were observed in all gels with HMF3S. The most VCAN was detected in the 0GAG gel (Figure 6.22C). Compared to the previous assay, the presence of 0GAG did not seem to increase T cell invasion (Figure 6.23C). The T cell donors used in this assay were different to the previous assay which may account for this difference.

Overall from the experiment, no direct comparisons could be made between the 0GAG and V1 constructs as no V1 was detected. However, it was observed that high levels of 0GAG in the gels did not seem to effect T cell invasion. Further repeats will need to be completed with the expression of V1 to understand whether the trends observed were due to T cell donors or due to the presence of CS.



**Figure 6.22. Effect of VCAN plasmid transfection on T cell invasion.** A) HCC38, B) MDAMB468 C) HMF3S cells were transfected with a 0GAG and V1 plasmid. The cells were grown in a gel for 7 days and treated with activated T cells for 3 days. Gels were paraffin embedded and sectioned. IHC was completed for VCAN and CD45. 2 biological repeats, 3 technical repeats.



**Figure 6.23. Effect of VCAN expression on T cell invasion.** A) HCC38, B) MDAMB468 C) HMF3S cells were transfected with a 0GAG and V1 plasmid. IHC was completed for VCAN and CD45 (T cell) Quantification of the VCAN area coverage and the number of invaded T cells was completed using QuPath. Colours represent the two T cell donors. 2 biological repeats, 3 technical repeats.

## 6.6. Discussion

To explore whether VCAN had a direct role on T cell invasion, several 3D in vitro models were designed and tested. The aim of the first model was to mimic the structure of the tissues and produce an epithelial and stromal zone for analysis. To try to achieve this, a collagen gel with tumour cells was embedded into a gel with fibroblasts. One of the limitations of analysing VCAN in a collagen gel was found to be the level of protein detected. The addition of VCAN to the gel or even the stimulation of cells to produce VCAN did not lead to high levels of VCAN and the presentation of VCAN was not representative of what was observed in tissues.

TGF $\beta$ 3 was used to stimulate the production of VCAN in cells as it has been previously shown to increase the expression of VCAN<sup>101,266</sup>. The limitation of using TGF $\beta$ 3 was that it can also stimulate the production of other ECM proteins<sup>266</sup> and therefore the effect observed may not be due to VCAN. TGF $\beta$ 3 treatment was found to limit T cell infiltration with all conditions showing lower infiltration compared to the untreated control. Surprisingly the least effect was seen when both the tumour cells and fibroblasts were exposed to TGF $\beta$ 3 as it was predicted that this condition would lead to the highest amount of protein production leading to a stiffer, dense gel. One possible reason for this is that TGF $\beta$ 3 can also stimulate the expression of MMPs<sup>270,271</sup> which could break down the ECM to assist T cell migration.

I then looked to use other invasion models to explore the interaction of VCAN and T cells. VCAN was enriched from HCC38 and HMF3S cells as described in the previous chapter. The enriched samples were not pure for VCAN, therefore to understand the effect of VCAN they were compared to the crude to show how more VCAN can effect migration. The impact of the structural components of VCAN to T cell invasion were tested by enzymatic treatment. Comparisons of the crude and enriched protein showed that the enriched protein increased migration. This was more notable in the enrichment from the protein of HMF3S cells. This showed that an increase in VCAN and possibly a loss of other proteins improves the invasive ability of T cells. To explore the role of CS, CSase treatment was applied to the samples. Contrasting results were seen with the CSase treatment where in a couple experiments the CSase treatment looked to reduce T cell infiltration while in the remaining trials it increased the infiltration. This could be due to the efficiency of the enzyme treatment or variations seen in the enrichment of VCAN. From the last experiments where two different samples of HCC38 were used with 3 different T cell donors, there was greater confidence to associate the removal of CS with the increased invasion of T cells.



The differences observed in sample A and B for HCC38 could be due to lower levels of VCAN in sample B as observed in the mass spec analysis (**Figure 5.13, 5.17**). The effect of the CSase treatment could be referred back to the Cs isomer analysis, where a higher CS-C level led to a greater effect of the CSase. This gives the impression that CS-C could affect T cell migration. This compliments what was observed in the tissues where the excluded tissues presented with more CS-C in the stroma in comparison to the inflamed tissues.

The effects could also be associated with the cell type where enrichment of VCAN from HMF3S cells compared to HCC38 cells having a greater improvement on invasion and the inverse seen with CS removal suggesting tumour associated VCAN may be linked to poor T cell infiltration as seen in excluded tissues with high levels of VCAN in the epithelial zone.

Using the transwell it was only possible to detect T cells that had infiltrated through the gel, therefore it was not possible to know if less cells were detected as the cells had not invaded or were immobilised within the collagen gel. Another experiment was completed in a 96 well plate where the T cells were imaged through a collagen gel. With CSase treatment of the protein added to the gel, a higher density of cells were observed at the bottom of the gel, confirming that the proteins in the gel were limiting cell movement. No cytokines were present within gels, indicating the cells could be interacting directly with the CS chains. It is predicted that the interaction of CD44 with the CS chains may be responsible for the slowing of T cell invasion<sup>254,255</sup>. The loss of CS could also affect the organisation and structure of the proteins in the gel with CS chains having the ability to interact with collagen fibrils<sup>272</sup>.

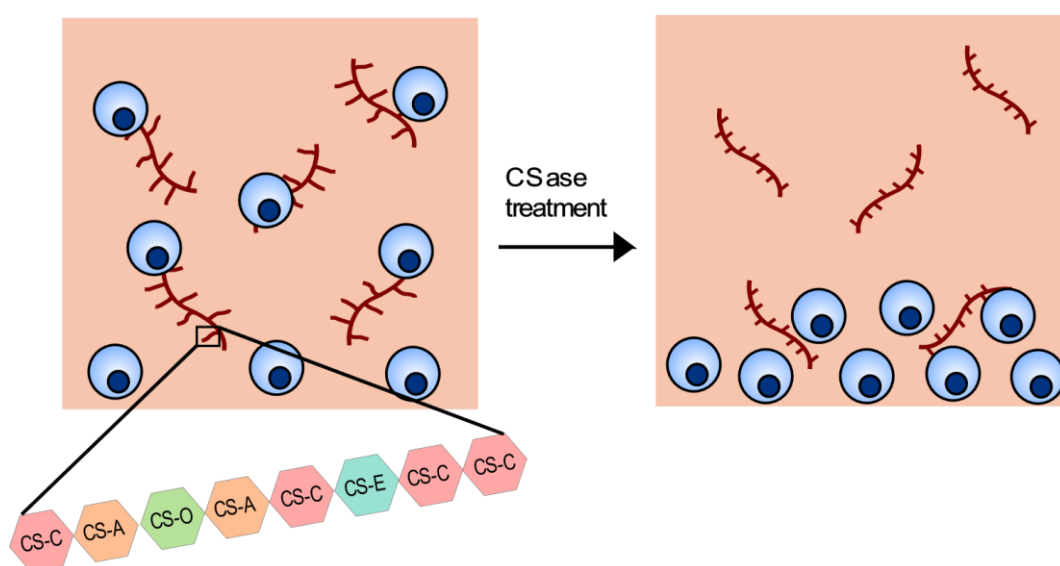
In tissues, the level of CS expressing VCAN can be reduced by degrading the protein to the peptide versikine. This has been seen in development where the breakdown of VCAN to versikine is important in the remodelling of cardiac tissue<sup>106</sup>. Within colorectal cancer, higher versikine levels have been associated with improved T cell infiltration to the tumour epithelium<sup>20</sup>. To see if this is also true for TNBC, I compared the different TIPs for their versikine expression levels. No associations were detected between versikine levels and phenotype but inflamed tissues were found to have a positive correlation between VCAN and versikine in the stroma, suggesting greater levels of VCAN degradation. To try to model the high level of degradation and quantify the level of T cell invasion, the transwell model was completed with ADAMTS4 treatment of the proteins. No T cell infiltration was seen after treatment. This may be due to other fibres also being degraded affecting



the architecture of the gel, which had been shown to be important for T cell migration <sup>273</sup>.

I then tried to look at the effect of versikine by expressing it using a plasmid and comparing it to full length VCAN isoform V1. The versikine plasmid was found to be expressed in the gel around the cells, however it was not possible to detect the V1 expressed by the plasmid. This might be associated with the large size of the plasmid (14,849bp) making it harder to express. To try to optimise the expression, different transfection reagents could be trialled out.

Overall, it could be concluded that the CS chains on VCAN play an important role in the invasion of T cells with the isomers affecting the impact of the effect (**Figure 6.24**). Investigating whether the CS isomers are directly interacting with the T cells or the surrounding matrix, would be important in understanding how the effect is occurring.



**Figure 6.24. Summary of T cell invasion and CS.** Removal of CS chains specifically with high CS-C isomer leads to an increase in T cell invasion.

## Chapter 7: Discussion

## 7.1. Exploring TIPs in TNBC.

Immune infiltration into tumours has been well known to be prognostic for response to immunotherapy <sup>274</sup>. The concept of immune infiltration had been initially split into two phenotypes of hot/inflamed (high immune infiltrate) and cold/desert (low immune infiltrate) tumours. Within the last 10 years, the concept of an immune excluded tumour was first described where tumours have a high immune infiltrate but the immune cells do not interact with the tumour cells <sup>41,275</sup>. This phenotype has been linked to response to checkpoint inhibition <sup>276</sup>. Therefore understanding the TIP of tissues can enable the prediction of response to therapy.

The assigning of TIPs to the library of TNBC tissues used here was completed by IHC analysis to locate where the immune cells were in respect to the tumour epithelium. Previous methods to define TIPs <sup>44,49,215,216</sup> were explored to determine the thresholds for the phenotypes. Different methods for determining TIPs tested here, varied in the areas selected and the definition of the thresholds for classifying tissues as excluded and inflamed. I found this to be important in order to correctly identify tumours as being excluded as opposed to less infiltrated in comparison to the inflamed.

From the analysis completed on the tissues I was able to identify around 50% of tissues to be excluded with only a quarter inflamed. The number of inflamed tissues was found to associate with immune checkpoint blockade responses in TNBC <sup>196</sup>. Intratumoural heterogeneity (ITH) has been studied in relation to the phenotypes of immune cells <sup>277</sup> however little is known on the ITH of the spatial localisation of immune cells. ITH of infiltrating lymphocyte counts has been linked to immunotherapy response <sup>227</sup>. The exploration of the TIPs in the different areas around the tissue identified multiple phenotypes within a tissue. The presence of more than one TIP within a tumour may affect the efficacy of immunotherapy with areas of an excluded/desert phenotype leading to residual disease.

## 7.2. VCAN expression around the epithelium associates with TIP

Having identified the TIPs of the tissues, the next step was to understand what leads to tumours having these phenotypes. Within literature it has been well established that the ECM can have a role in immune filtration <sup>24,25</sup> and immunotherapy response<sup>278</sup>.

The initial step to identifying a target within the ECM was to look at previously identified signature that associated with immune suppression and disease progression<sup>23</sup>. Within this signature the upregulated molecules were explored for how they correlated and co-localised to potentially identify key interactions between ECM molecules. VCAN and CTSB were found to be correlated but through dual RNAscope it was visible that the markers weren't specifically co-expressed. Comparisons of the dual RNAscope to CD8+ T cell staining indicated that areas of high VCAN expression were associated with CD8+ T cells being excluded within the stroma. This led to finding of VCAN as a potential key marker in the ECM linked with immune exclusion. VCAN has previously been found to be associated with inflammation<sup>279</sup> and it was also identified in a gene signature comparing responders and non-responders to immunotherapy<sup>228</sup>.

Within the ECM, multiple molecules such as HA, FN1 and collagen are present which VCAN can bind to through its multiple binding domains<sup>92</sup>. However when comparing these markers to CD8+ and CD68+ cells within the tumour epithelium and stroma a negative correlation was only observed with VCAN indicating the associations with immune infiltration are independent to other ECM molecules.

The link between VCAN and TIP's was explored where it was discovered that inflamed tissues predominantly expressed VCAN within the stroma whilst in excluded tissues VCAN was also present around the tumour epithelium. CD44 and HA form macromolecular structures with VCAN in the pericellular matrix around tumour cells<sup>91</sup>. This did not seem to be linked to the high level of VCAN within the tumour epithelium.

The stimulation of VCAN was also explored with VCAN+ fibroblasts identified around tumour cells. An increase in FAP expression has been associated with CAFs expressing VCAN<sup>136,280</sup>, but FAP detection was not linked to greater VCAN expression through RNAscope. Co-culture of a fibroblast cell line HMF3S with a TNBC cell line HCC38, led to an increase in VCAN expression. This suggested that the tumour cells may stimulate VCAN expression by the fibroblasts which may be leading to the formation of an excluded tumour.

### 7.3. CS proteoglycans and immune exclusion

Following the identification of VCAN as a target, I looked to explore why VCAN is responsible for immune exclusion. The structure of VCAN is important towards its

functions. VCAN is present as five isoforms which can be found to have different roles within cancer <sup>65,144,146,148,151</sup>. Comparisons of isoforms in tumour tissues and cell lines showed that the V0 and V1 isoforms were the highest expressed. Compared to the other isoforms, V0 and V1 present the highest amount of CS <sup>90,223</sup>. CS is a key structural feature of VCAN as it can impact the organisation of the matrix by binding to ECM molecules in addition to attracting chemokines and immune cells affecting their migration within the matrix<sup>242</sup>. Therefore CS expression on VCAN was thought to be linked to immune exclusion.

Comparing CS to TIPs, in inflamed tissues CS expression was similar to VCAN where it was greater in the stromal area however in the excluded tissues the CS levels were higher in the stromal area and lower in the tumour areas. The poor correlations of VCAN and CS in excluded tissues suggested that other CS proteoglycans may be present. DCN and BGN were found to be highly expressed within the ECM and could also be associated with immune exclusion. Both proteins were found to be primarily expressed within the stroma and showed no correlation to TIP, further indicating VCAN as the main CS proteoglycan associated with TIP.

#### 7.4. Isoforms of CS associate with TIP

CS can be present in different isoforms based on the sulphation site. The sulphation site can affect the binding affinity of the disaccharide to different chemokines and cell receptors <sup>254</sup>. This is due to changes in the conformation of the disaccharide following sulphation <sup>281</sup>. Through a collaboration with Dr Kim Alonge, I explored how the CS sulphation pattern varies in the excluded and inflamed tissues and the patterns in the stroma and tumour epithelial regions. From the analysis it was found that the CS pattern in the excluded tissues was very similar in both the tumour and stromal whilst in the inflamed tissues the ratio of CS-A and CS-C was greater in the stroma and lower in the tumour. Comparing both phenotypes showed no difference in the tumour regions but in the stromal area, the excluded tissues had a lower CS-A/CS-C ratio. A previous study looking at CS-C in colorectal cancer also identified that high levels of CS-C in the stroma surrounding tumour islands correlated with an excluded phenotype <sup>256</sup>. This strengthened the thought that a low CS-A/CS-C ratio in the stroma may be leading to T cell trapping, preventing immune cell migration towards the tumour epithelium.

One way in which T cell trapping may be occurring within the stroma is by the effect on the chemokine gradient. CS-C has been found to have a stronger binding

capacity to chemokines such as CCL5<sup>282</sup> and IL-8<sup>78</sup> in comparison to CS-A. Both CCL5 and IL-8 have been shown to be involved in the chemotaxis of T cells during inflammation<sup>283,284</sup>. With the level of CS-C increasing in the tumour area of inflamed tissues, T cells can be guided to migrate towards the tumour epithelium. In contrast within the excluded tissues there is no change in CS sulphation patterns and also level of CS, possibly leading to no gradient formation, and T cells remain in the stroma.

## 7.5. Removal of CS-C improves T cell invasion

To test the impact of VCAN and CS on T cell invasion, 3D gel models and transwell models were designed to include increasing amounts of VCAN and the use of CSase treatment to understand the impact of CS. From the gel models it was seen that the stimulation of ECM production by TGFβ3 led to a reduced T cell invasion. VCAN could not be singled out as a key ECM protein in this model. VCAN enriched protein was then trialled in a transwell model where the enrichment of the protein with VCAN increased T-cell invasion. CSase treatment was applied to the protein where it was observed that CSase treatment of protein from TNBC cell line HCC38 increased invasion whilst little effect was observed with protein from HMF3S cells. From the exploration of Cs isomers in these protein samples it was found that the HCC38 protein had a lower CS-A/CS-C ratio compared to HMF3S protein. The protein extracted from the HCC38 could be associated with the excluded tissues where higher VCAN levels were seen in the epithelial zone whilst the protein from HMF3S could be associated to inflamed tissues where the fibroblasts expressing VCAN were distanced from the tumour and possibly less stimulated.

It was noted that in the transwell model, there were no chemokines present in the gel to indicate that a chemokine gradient was causing the effect observed.

Therefore suggesting that there may be a direct interaction occurring between CS and T cells. T cells can interact with CS on VCAN via CD44 and selectins<sup>161</sup>.

Differences have been observed with the effect of CS structure on T cell activity<sup>285</sup>, however there has been no indication in literature to show a direct impact of Cs isomer on T cell migration.

Another reason for the differences observed with CSase treatment may be due to the effect the CS structure has towards the matrix organisation. The effect of CS structure towards the ECM can be seen in tissues where the loss of the expression of CS-C has been linked with Ehlers Danlos syndrome, where patients suffer from

loose joints, connective tissue and skin<sup>286</sup>. Higher levels of CS-C and loss of CS-A have been associated with tumour remodelling<sup>287</sup>. CS-C has been found to be able to bind to collagen fibrils and form interfibrillar bridges whereas no effect was observed with CS-A<sup>288</sup>. This could suggest a looser ECM where higher levels of CS-A are present in comparison to CS-C.

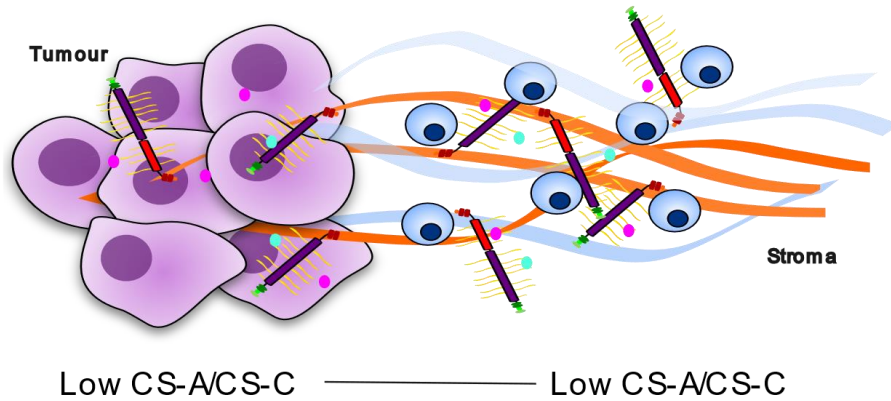
## 7.6. VCAN degradation affects T cell invasion

Another structural change to VCAN that affect its function is the degradation to versikine. Within inflamed tissues it was observed that in the stroma there was a strong correlation between VCAN and versikine indicating that a majority of the VCAN detected in the stroma may have been degraded to versikine. The presence of versikine has been associated with increased T cell infiltration<sup>20</sup>, therefore this was not surprising. VCAN degradation occurs by ADAMTS enzymes. However, the inclusion of ADAMTS4 treatment to the transwell was seen to stop T cell invasion. This indicates that degrading VCANs protein backbone to improve T cell infiltration may not be viable, in contrast to targeting the GAG isoforms on VCAN.

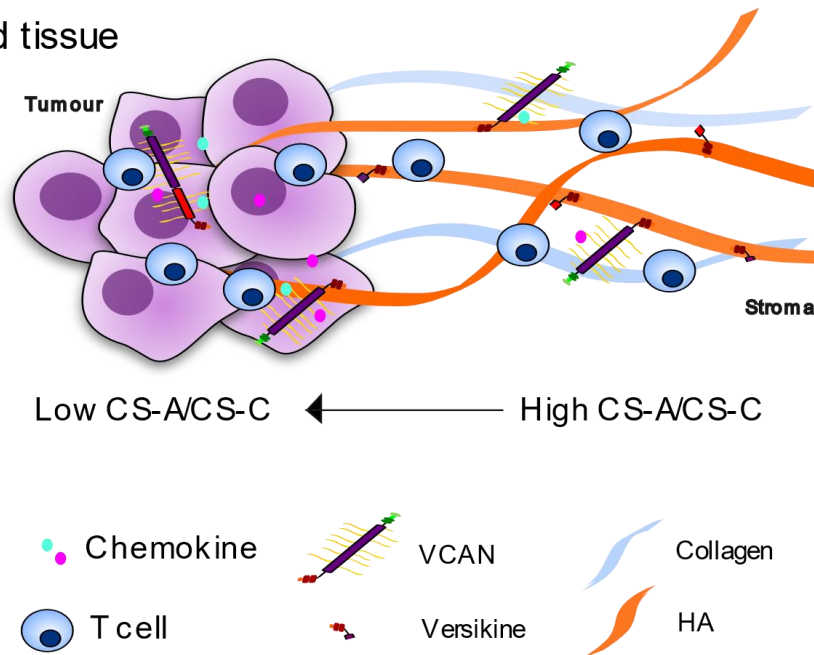
## 7.7. Conclusions

Phenotypes can be assigned to tissues based on the localisation of immune cells. Comparison of these phenotypes to key matrix molecules identified VCAN expression around the tumour epithelium as being correlated with an excluded phenotype. The functions of VCAN are dependent on its structure. CS levels in the tissue were found to be irregular in comparison to VCAN in excluded tissues. The analysis of sulphation patterns of CS identified higher levels C-6-S (CS-C) isoform to associate with T cell trapping and stromal restriction of T cells. Whereas a higher CS-A/CS-C ratio an increase in versikine levels in the stroma can promote an inflamed phenotype (Figure 7.1).

## Excluded tissue



## Inflamed tissue



**Figure 7.1. Effect of VCAN structure towards tumour immune phenotype.** Excluded tissues have a low CS-A/CS-C ratio in both the tumour and stroma leading to no chemokine gradient to guide T cells to the tumour epithelium. Inflamed tumours have a high CS-A/CS-C ratio in the stroma and a low ratio in the tumour leading to the formation of a chemokine gradient towards the tumour. A greater proportion of VCAN has been degraded to versikine in inflamed tissues further aiding T cell invasion.

## 7.8. Impact of study

Having identified a structural feature of VCAN that associates with its affect towards T cell infiltration potential treatment targets can be designed. From the study, CS-C has been identified to be an important molecule towards T cell invasion through its role in matrix organisation and the formation of chemokine gradients. Targeting CS-C could lead to an improvement in T cell infiltration to the tumour epithelium converting excluded tumours to inflamed tumours. The enzyme galactose-6-sulfatase is a lysosomal enzyme that hydrolyses the sulphation in CS-C. The enzyme has been shown to specifically affect the level of CS-C through siRNA



knockdown and overexpression <sup>289</sup>. Through the degradation of CS-C, it is predicted that T cells would be less restricted within the stroma due to breakdown of attraction towards CS-C in the stroma as well as opening up the matrix to provide more pores for T cell migration. However, treating tumours with the enzyme would lead to the removal of CS-C from both the stromal area and the tumour epithelium, removing the chemokine gradient during the process. Possible interventions would include an antibody-enzyme conjugated therapy where galactose-6-sulfatase could be directed towards stromal markers such as FAP. This would allow the sulfatase to degrade the CS-C isoforms in primarily the stroma, leading to less chemokine binding and a looser matrix.

## 7.9. Future work

To continue the research for the project, the proposed objectives are:

- TIP analysis of tissues

The TIP analysis would also be completed for other immune cell subtypes to identify whether different immune cells interact with the ECM in other ways and understand their abilities to invade into solid tumours. The analysis could also be extended to other tumour types to identify whether there are tissue specific features.

- VCAN expression analysis

At the end of Chapter 3 it was identified that tumour cells could impact VCAN expression in fibroblasts. Understanding the trigger for the increase in expression and the effect that it might have on fibroblast phenotypes could help to learn what leads to tumours turning into an excluded phenotype, and lead to potential interventions. This would be completed by using a range of tumour cells and comparing how they can activate fibroblasts from healthy donors. Different markers would be studied on fibroblasts as well as cytokines secreted by tumour cells. It would also be interesting to explore what is the difference in the VCAN secreted by the tumour cells and fibroblasts in the tumour and whether this may be associated with phenotype as well. This would require a range of different tumour cell lines and fibroblasts where secreted VCAN would be characterised for structure and tested in functional assays.

- CS analysis

Having studied how CS varies in sulphation in the tissues, it would also be important to investigate how CS can differ in chain length and also where the CS is presented on VCAN. The length of the CS disaccharide chains have been shown to be linked to inflammation <sup>290</sup>, while the location of CS expression on VCAN can affect the role of ADAMTS enzymes <sup>116</sup>. This would also provide more information as to why CS poorly correlated and co-localised with VCAN in excluded tissues.

- VCAN isolation

Further optimisation is required to achieve a better isolation of VCAN. To accomplish this other techniques could be trialled to further enrich VCAN in the sample, this includes size exclusion chromatography to isolate the large VCAN structures, as well as optimising current techniques of ion exchange possibly using a greater salt gradient range.

- 3D model work to study T cell invasion

Data shown in Chapter 6 included most studies completed as N=1 and therefore more repeats would be required. From the 3D models trailed, the limiting factor was found to be the level of VCAN in the gels. Optimisation of the overexpression of VCAN would overcome this.

- Studying CS-C and T cell invasion

Demonstration of CS-C as the cause of T cell trapping would be required. This could be achieved by knockdown of CS-C expression through the silencing of the enzyme CHST3 which is responsible for the sulphation, then repeating the experiments completed studying T cell invasion in gels and transwells. Recombinant galactose-6-sulfatase could also be used in comparison to CSase treatment to identify whether removal of CS-C is sufficient.

# Bibliography

- 1 Cox, T. R. The matrix in cancer. *Nature Reviews Cancer* **21**, 217-238, doi:doi:10.1038/s41568-020-00329-7 (2021).
- 2 Pickup, M. W., Mouw, J. K. & Weaver, V. M. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep* **15**, 1243-1253, doi:10.15252/embr.201439246 (2014).
- 3 Hynes, R. O. & Naba, A. Overview of the matrisome--an inventory of extracellular matrix constituents and functions. *Cold Spring Harb Perspect Biol* **4**, a004903, doi:10.1101/cshperspect.a004903 (2012).
- 4 Karamanos, N. K. *et al.* A guide to the composition and functions of the extracellular matrix. *Febs J* **288**, 6850-6912, doi:10.1111/febs.15776 (2021).
- 5 Yue, B. Biology of the extracellular matrix: an overview. *J Glaucoma* **23**, S20-23, doi:10.1097/IJG.000000000000108 (2014).
- 6 Naba, A. *et al.* The extracellular matrix: Tools and insights for the "omics" era. *Matrix Biol* **49**, 10-24, doi:10.1016/j.matbio.2015.06.003 (2016).
- 7 Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun* **11**, 5120, doi:10.1038/s41467-020-18794-x (2020).
- 8 Heneberg, P. Paracrine tumor signaling induces transdifferentiation of surrounding fibroblasts. *Crit Rev Oncol Hematol* **97**, 303-311, doi:10.1016/j.critrevonc.2015.09.008 (2016).
- 9 Fang, M., Yuan, J., Peng, C. & Li, Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol* **35**, 2871-2882, doi:10.1007/s13277-013-1511-7 (2014).
- 10 Liu, M., Tolg, C. & Turley, E. Dissecting the Dual Nature of Hyaluronan in the Tumor Microenvironment. *Front Immunol* **10**, 947, doi:10.3389/fimmu.2019.00947 (2019).
- 11 Wuest, T., Moosmayer, D. & Pfizenmaier, K. Construction of a bispecific single chain antibody for recruitment of cytotoxic T cells to the tumour stroma associated antigen fibroblast activation protein. *J Biotechnol* **92**, 159-168, doi:10.1016/s0168-1656(01)00355-8 (2001).
- 12 Setargew, Y. F. I., Wyllie, K., Grant, R. D., Chitty, J. L. & Cox, T. R. Targeting Lysyl Oxidase Family Mediated Matrix Cross-Linking as an Anti-Stromal Therapy in Solid Tumours. *Cancers (Basel)* **13**, doi:10.3390/cancers13030491 (2021).
- 13 Belli, C. *et al.* Targeting the microenvironment in solid tumors. *Cancer Treat Rev* **65**, 22-32, doi:10.1016/j.ctrv.2018.02.004 (2018).
- 14 Kaplan, G. In vitro differentiation of human monocytes. Monocytes cultured on glass are cytotoxic to tumor cells but monocytes cultured on collagen are not. *J Exp Med* **157**, 2061-2072, doi:10.1084/jem.157.6.2061 (1983).
- 15 Deligne, C. *et al.* Matrix-Targeting Immunotherapy Controls Tumor Growth and Spread by Switching Macrophage Phenotype. *Cancer Immunol Res* **8**, 368-382, doi:10.1158/2326-6066.Cir-19-0276 (2020).
- 16 Frey, H., Schroeder, N., Manon-Jensen, T., Iozzo, R. V. & Schaefer, L. Biological interplay between proteoglycans and their innate immune receptors in inflammation. *Febs J* **280**, 2165-2179, doi:10.1111/febs.12145 (2013).
- 17 Deligne, C. & Midwood, K. S. Macrophages and Extracellular Matrix in Breast Cancer: Partners in Crime or Protective Allies? *Front Oncol* **11**, 620773, doi:10.3389/fonc.2021.620773 (2021).
- 18 Huleihel, L. *et al.* Macrophage phenotype in response to ECM bioscaffolds. *Semin Immunol* **29**, 2-13, doi:10.1016/j.smim.2017.04.004 (2017).
- 19 Garcia-Nieto, S. *et al.* Laminin and fibronectin treatment leads to generation of dendritic cells with superior endocytic capacity. *PLoS One* **5**, e10123, doi:10.1371/journal.pone.0010123 (2010).

- 20 Hope, C. *et al.* Versican-Derived Matrikines Regulate Batf3-Dendritic Cell Differentiation and Promote T Cell Infiltration in Colorectal Cancer. *J Immunol* **199**, 1933-1941, doi:10.4049/jimmunol.1700529 (2017).
- 21 Gao, H. *et al.* 3D Extracellular Matrix Regulates the Activity of T Cells and Cancer Associated Fibroblasts in Breast Cancer. *Front Oncol* **11**, 764204, doi:10.3389/fonc.2021.764204 (2021).
- 22 Chirivi, M. *et al.* Tumor Extracellular Matrix Stiffness Promptly Modulates the Phenotype and Gene Expression of Infiltrating T Lymphocytes. *Int J Mol Sci* **22**, doi:10.3390/ijms22115862 (2021).
- 23 Pearce, O. M. T. *et al.* Deconstruction of a Metastatic Tumor Microenvironment Reveals a Common Matrix Response in Human Cancers. *Cancer Discov* **8**, 304-319, doi:10.1158/2159-8290.CD-17-0284 (2018).
- 24 Hallmann, R. *et al.* The regulation of immune cell trafficking by the extracellular matrix. *Curr Opin Cell Biol* **36**, 54-61, doi:10.1016/j.ceb.2015.06.006 (2015).
- 25 Gordon-Weeks, A. & Yuzhalin, A. E. Cancer Extracellular Matrix Proteins Regulate Tumour Immunity. *Cancers (Basel)* **12**, doi:10.3390/cancers12113331 (2020).
- 26 Ivanoff, J., Talme, T. & Sundqvist, K. G. The role of chemokines and extracellular matrix components in the migration of T lymphocytes into three-dimensional substrata. *Immunology* **114**, 53-62, doi:10.1111/j.1365-2567.2004.02005.x (2005).
- 27 Huang, J. Y. *et al.* Extracellular matrix of glioblastoma inhibits polarization and transmigration of T cells: the role of tenascin-C in immune suppression. *J Immunol* **185**, 1450-1459, doi:10.4049/jimmunol.0901352 (2010).
- 28 da Silva, P. H. R. *et al.* Chemokines and the extracellular matrix: Set of targets for tumor development and treatment. *Cytokine* **144**, 155548, doi:10.1016/j.cyto.2021.155548 (2021).
- 29 Gray, A. L., Pun, N., Ridley, A. J. L. & Dyer, D. P. Role of extracellular matrix proteoglycans in immune cell recruitment. *Int J Exp Pathol* **103**, 34-43, doi:10.1111/iep.12428 (2022).
- 30 Gray, A. L. *et al.* Chemokine CXCL4 interactions with extracellular matrix proteoglycans mediate widespread immune cell recruitment independent of chemokine receptors. *Cell Rep* **42**, 111930, doi:10.1016/j.celrep.2022.111930 (2023).
- 31 Cripps, J. G., Crespo, F. A., Romanovskis, P., Spatola, A. F. & Fernandez-Botran, R. Modulation of acute inflammation by targeting glycosaminoglycan-cytokine interactions. *Int Immunopharmacol* **5**, 1622-1632, doi:10.1016/j.intimp.2005.04.010 (2005).
- 32 Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**, 1-10, doi:10.1016/j.immuni.2013.07.012 (2013).
- 33 Bonnotte, B. *et al.* Role of tumor cell apoptosis in tumor antigen migration to the draining lymph nodes. *J Immunol* **164**, 1995-2000, doi:10.4049/jimmunol.164.4.1995 (2000).
- 34 Franciszkiewicz, K., Boissonnas, A., Boutet, M., Combadiere, C. & Mami-Chouaib, F. Role of chemokines and chemokine receptors in shaping the effector phase of the antitumor immune response. *Cancer Res* **72**, 6325-6332, doi:10.1158/0008-5472.Can-12-2027 (2012).
- 35 Franciszkiewicz, K. *et al.* CD103 or LFA-1 engagement at the immune synapse between cytotoxic T cells and tumor cells promotes maturation and regulates T-cell effector functions. *Cancer Res* **73**, 617-628, doi:10.1158/0008-5472.Can-12-2569 (2013).
- 36 Spranger, S. *et al.* Density of immunogenic antigens does not explain the presence or absence of the T-cell-inflamed tumor microenvironment in

- melanoma. *Proc Natl Acad Sci U S A* **113**, E7759-E7768, doi:10.1073/pnas.1609376113 (2016).
- 37 De Guillebon, E. *et al.* Beyond the concept of cold and hot tumors for the development of novel predictive biomarkers and the rational design of immunotherapy combination. *Int J Cancer* **147**, 1509-1518, doi:10.1002/ijc.32889 (2020).
- 38 Harlin, H. *et al.* Chemokine expression in melanoma metastases associated with CD8+ T-cell recruitment. *Cancer Res* **69**, 3077-3085, doi:10.1158/0008-5472.Can-08-2281 (2009).
- 39 Ullman, N. A., Burchard, P. R., Dunne, R. F. & Linehan, D. C. Immunologic Strategies in Pancreatic Cancer: Making Cold Tumors Hot. *J Clin Oncol* **40**, 2789-2805, doi:10.1200/JCO.21.02616 (2022).
- 40 Liu, Y. T. & Sun, Z. J. Turning cold tumors into hot tumors by improving T-cell infiltration. *Theranostics* **11**, 5365-5386, doi:10.7150/thno.58390 (2021).
- 41 Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321-330, doi:10.1038/nature21349 (2017).
- 42 Galon, J. *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313**, 1960-1964, doi:10.1126/science.1129139 (2006).
- 43 Pages, F. *et al.* International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* **391**, 2128-2139, doi:10.1016/S0140-6736(18)30789-X (2018).
- 44 Kather, J. N. *et al.* Topography of cancer-associated immune cells in human solid tumors. *Elife* **7**, doi:10.7554/eLife.36967 (2018).
- 45 Gruosso, T. *et al.* Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. *J Clin Invest* **129**, 1785-1800, doi:10.1172/JCI96313 (2019).
- 46 Xu, Q., Chen, S., Hu, Y. & Huang, W. Landscape of Immune Microenvironment Under Immune Cell Infiltration Pattern in Breast Cancer. *Front Immunol* **12**, 711433, doi:10.3389/fimmu.2021.711433 (2021).
- 47 Van Den Eeckhout, B., Tavernier, J. & Gerlo, S. Interleukin-1 as Innate Mediator of T Cell Immunity. *Front Immunol* **11**, 621931, doi:10.3389/fimmu.2020.621931 (2020).
- 48 Shi, L., Chen, S., Yang, L. & Li, Y. The role of PD-1 and PD-L1 in T-cell immune suppression in patients with hematological malignancies. *J Hematol Oncol* **6**, 74, doi:10.1186/1756-8722-6-74 (2013).
- 49 Hammerl, D. *et al.* Spatial immunophenotypes predict response to anti-PD1 treatment and capture distinct paths of T cell evasion in triple negative breast cancer. *Nat Commun* **12**, 5668, doi:10.1038/s41467-021-25962-0 (2021).
- 50 Kolesnikoff, N., Chen, C. H. & Samuel, M. S. Interrelationships between the extracellular matrix and the immune microenvironment that govern epithelial tumour progression. *Clin Sci (Lond)* **136**, 361-377, doi:10.1042/CS20210679 (2022).
- 51 Blackmon, R. L. *et al.* Imaging Extracellular Matrix Remodeling In Vitro by Diffusion-Sensitive Optical Coherence Tomography. *Biophys J* **110**, 1858-1868, doi:10.1016/j.bpj.2016.03.014 (2016).
- 52 McKee, T. D. *et al.* Degradation of fibrillar collagen in a human melanoma xenograft improves the efficacy of an oncolytic herpes simplex virus vector. *Cancer Res* **66**, 2509-2513, doi:10.1158/0008-5472.CAN-05-2242 (2006).
- 53 Abyaneh, H. S., Regenold, M., McKee, T. D., Allen, C. & Gauthier, M. A. Towards extracellular matrix normalization for improved treatment of solid tumors. *Theranostics* **10**, 1960-1980, doi:10.7150/thno.39995 (2020).
- 54 Hingorani, S. R. *et al.* HALO 202: Randomized Phase II Study of PEGPH20 Plus Nab-Paclitaxel/Gemcitabine Versus Nab-Paclitaxel/Gemcitabine in

- Patients With Untreated, Metastatic Pancreatic Ductal Adenocarcinoma. *J Clin Oncol* **36**, 359-366, doi:10.1200/JCO.2017.74.9564 (2018).
- 55 Clift, R., Souratha, J., Garroville, S. A., Zimmerman, S. & Blouw, B. Remodeling the Tumor Microenvironment Sensitizes Breast Tumors to Anti-Programmed Death-Ligand 1 Immunotherapy. *Cancer Res* **79**, 4149-4159, doi:10.1158/0008-5472.CAN-18-3060 (2019).
- 56 Wang, J. *et al.* Collagenase-loaded pH-sensitive nanocarriers efficiently remodeled tumor stroma matrixes and improved the enrichment of nanomedicines. *Nanoscale* **13**, 9402-9414, doi:10.1039/d1nr00950h (2021).
- 57 Ebel, N. D., Zuniga, E., Passi, K. B., Sobocinski, L. J. & Manuel, E. R. Hyaluronidase-Expressing Salmonella Effectively Targets Tumor-Associated Hyaluronic Acid in Pancreatic Ductal Adenocarcinoma. *Mol Cancer Ther* **19**, 706-716, doi:10.1158/1535-7163.MCT-19-0556 (2020).
- 58 Hirani, P., Gauthier, V., Allen, C. E., Wight, T. N. & Pearce, O. M. T. Targeting Versican as a Potential Immunotherapeutic Strategy in the Treatment of Cancer. *Front Oncol* **11**, 712807, doi:10.3389/fonc.2021.712807 (2021).
- 59 Sotoodehnejadnematalahi, F. & Burke, B. Structure, function and regulation of versican: the most abundant type of proteoglycan in the extracellular matrix. *Acta Med Iran* **51**, 740-750 (2013).
- 60 Seyfried, N. T. *et al.* Expression and purification of functionally active hyaluronan-binding domains from human cartilage link protein, aggrecan and versican: formation of ternary complexes with defined hyaluronan oligosaccharides. *J Biol Chem* **280**, 5435-5448, doi:10.1074/jbc.M411297200 (2005).
- 61 Zimmermann, D. R. & Ruoslahti, E. Multiple domains of the large fibroblast proteoglycan, versican. *Embo J* **8**, 2975-2981, doi:10.1002/j.1460-2075.1989.tb08447.x (1989).
- 62 Dours-Zimmermann, M. T. & Zimmermann, D. R. A novel glycosaminoglycan attachment domain identified in two alternative splice variants of human versican. *J Biol Chem* **269**, 32992-32998 (1994).
- 63 Zako, M., Shinomura, T., Ujita, M., Ito, K. & Kimata, K. Expression of PG-M(V3), an alternatively spliced form of PG-M without a chondroitin sulfate attachment in region in mouse and human tissues. *J Biol Chem* **270**, 3914-3918, doi:10.1074/jbc.270.8.3914 (1995).
- 64 Lemire, J. M. *et al.* Versican/Pg-M isoforms in vascular smooth muscle cells. *Arterioscler Thromb Vasc Biol* **19**, 1630-1639, doi:10.1161/01.atv.19.7.1630 (1999).
- 65 Kischel, P. *et al.* Versican overexpression in human breast cancer lesions: known and new isoforms for stromal tumor targeting. *Int J Cancer* **126**, 640-650, doi:10.1002/ijc.24812 (2010).
- 66 Wight, T. N. Versican: a versatile extracellular matrix proteoglycan in cell biology. *Curr Opin Cell Biol* **14**, 617-623, doi:10.1016/s0955-0674(02)00375-7 (2002).
- 67 Sammon, D. *et al.* Molecular mechanism of decision-making in glycosaminoglycan biosynthesis. *Nat Commun* **14**, 6425, doi:10.1038/s41467-023-42236-z (2023).
- 68 Wen, J. *et al.* Xylose phosphorylation functions as a molecular switch to regulate proteoglycan biosynthesis. *Proc Natl Acad Sci U S A* **111**, 15723-15728, doi:10.1073/pnas.1417993111 (2014).
- 69 Prydz, K. Determinants of Glycosaminoglycan (GAG) Structure. *Biomolecules* **5**, 2003-2022, doi:10.3390/biom5032003 (2015).
- 70 Mizumoto, S. & Yamada, S. An Overview of in vivo Functions of Chondroitin Sulfate and Dermatan Sulfate Revealed by Their Deficient Mice. *Front Cell Dev Biol* **9**, 764781, doi:10.3389/fcell.2021.764781 (2021).

- 71 Levene, P. & La Forge, F. ON CHONDROITIN SULPHURIC ACID: FOURTH PAPER. *Journal of Biological Chemistry* **20**, 433-444 (1915).
- 72 Levene, P. & La Forge, F. On chondroitin sulphuric acid: second paper. *Journal of Biological Chemistry* **15**, 155-160 (1913).
- 73 Gulberti, S. *et al.* Chondroitin sulfate N-acetylgalactosaminyltransferase-1 (CSGαNACT-1) involved in chondroitin sulfate initiation: Impact of sulfation on activity and specificity. *Glycobiology* **22**, 561-571, doi:10.1093/glycob/cwr172 (2012).
- 74 Varki, A. *et al.* in *Essentials of glycobiology* (Cold Spring Harbor Laboratory Press %! Proteoglycans and glycosaminoglycans, 1999).
- 75 Cooney, C. A. *et al.* Chondroitin sulfates play a major role in breast cancer metastasis: a role for CSPG4 and CHST11 gene expression in forming surface P-selectin ligands in aggressive breast cancer cells. *Breast Cancer Res* **13**, R58, doi:10.1186/bcr2895 (2011).
- 76 Han, J. *et al.* Altered expression of chondroitin sulfate structure modifying sulfotransferases in the articular cartilage from adult osteoarthritis and Kashin-Beck disease. *Osteoarthritis Cartilage* **25**, 1372-1375, doi:10.1016/j.joca.2017.02.803 (2017).
- 77 Sugiura, N. *et al.* Construction of a chondroitin sulfate library with defined structures and analysis of molecular interactions. *J Biol Chem* **287**, 43390-43400, doi:10.1074/jbc.M112.412676 (2012).
- 78 Pichert, A. *et al.* Characterization of the interaction of interleukin-8 with hyaluronan, chondroitin sulfate, dermatan sulfate and their sulfated derivatives by spectroscopy and molecular modeling. *Glycobiology* **22**, 134-145, doi:10.1093/glycob/cwr120 (2012).
- 79 Saito, A. & Munakata, H. Detection of chondroitin sulfate-binding proteins on the membrane. *Electrophoresis* **25**, 2452-2460, doi:10.1002/elps.200305997 (2004).
- 80 Resende, M. *et al.* Identification of glycosaminoglycan binding regions in the Plasmodium falciparum encoded placental sequestration ligand, VAR2CSA. *Malar J* **7**, 104, doi:10.1186/1475-2875-7-104 (2008).
- 81 Wakao, M. *et al.* Synthesis of a chondroitin sulfate disaccharide library and a GAG-binding protein interaction analysis. *Bioorg Med Chem Lett* **25**, 1407-1411, doi:10.1016/j.bmcl.2015.02.054 (2015).
- 82 Sato, Y. *et al.* Nephronectin binds to heparan sulfate proteoglycans via its MAM domain. *Matrix Biol* **32**, 188-195, doi:10.1016/j.matbio.2013.01.005 (2013).
- 83 Koike, T., Izumikawa, T., Tamura, J. & Kitagawa, H. Chondroitin sulfate-E fine-tunes osteoblast differentiation via ERK1/2, Smad3 and Smad1/5/8 signaling by binding to N-cadherin and cadherin-11. *Biochem Biophys Res Commun* **420**, 523-529, doi:10.1016/j.bbrc.2012.03.024 (2012).
- 84 Izumikawa, T., Sato, B. & Kitagawa, H. Chondroitin sulfate is indispensable for pluripotency and differentiation of mouse embryonic stem cells. *Sci Rep* **4**, 3701, doi:10.1038/srep03701 (2014).
- 85 Kastana, P. *et al.* Insight into the role of chondroitin sulfate E in angiogenesis. *Febs J* **286**, 2921-2936, doi:10.1111/febs.14830 (2019).
- 86 Matsumoto, K. *et al.* Distinct interaction of versican/PG-M with hyaluronan and link protein. *J Biol Chem* **278**, 41205-41212, doi:10.1074/jbc.M305060200 (2003).
- 87 Spicer, A. P., Joo, A. & Bowling, R. A., Jr. A hyaluronan binding link protein gene family whose members are physically linked adjacent to chondroitin sulfate proteoglycan core protein genes: the missing links. *J Biol Chem* **278**, 21083-21091, doi:10.1074/jbc.M213100200 (2003).
- 88 Camenisch, T. D. *et al.* Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of



- epithelium to mesenchyme. *J Clin Invest* **106**, 349-360, doi:10.1172/JCI10272 (2000).
- 89 Mjaatvedt, C. H., Yamamura, H., Capehart, A. A., Turner, D. & Markwald, R. R. The Cspg2 gene, disrupted in the hdf mutant, is required for right cardiac chamber and endocardial cushion formation. *Dev Biol* **202**, 56-66, doi:10.1006/dbio.1998.9001 (1998).
- 90 Nandadasa, S. *et al.* The versican-hyaluronan complex provides an essential extracellular matrix niche for Flk1(+) hematoendothelial progenitors. *Matrix Biol* **97**, 40-57, doi:10.1016/j.matbio.2021.01.002 (2021).
- 91 Ween, M. P., Oehler, M. K. & Ricciardelli, C. Role of versican, hyaluronan and CD44 in ovarian cancer metastasis. *Int J Mol Sci* **12**, 1009-1029, doi:10.3390/ijms12021009 (2011).
- 92 Wu, Y. J., La Pierre, D. P., Wu, J., Yee, A. J. & Yang, B. B. The interaction of versican with its binding partners. *Cell Res* **15**, 483-494, doi:10.1038/sj.cr.7290318 (2005).
- 93 Camejo, G., Lopez, A., Lopez, F. & Quinones, J. Interaction of low density lipoproteins with arterial proteoglycans. The role of charge and sialic acid content. *Atherosclerosis* **55**, 93-105, doi:10.1016/0021-9150(85)90169-8 (1985).
- 94 Kang, H., Lu, J., Yang, J., Fan, Y. & Deng, X. Interaction of arterial proteoglycans with low density lipoproteins (LDLs): From theory to promising therapeutic approaches. *Medicine in Novel Technology and Devices* **3**, 100016, doi:<https://doi.org/10.1016/j.medntd.2019.100016> (2019).
- 95 Olin, A. I. *et al.* The proteoglycans aggrecan and Versican form networks with fibulin-2 through their lectin domain binding. *J Biol Chem* **276**, 1253-1261, doi:10.1074/jbc.M006783200 (2001).
- 96 Lundell, A. *et al.* Structural basis for interactions between tenascins and lectican C-type lectin domains: evidence for a crosslinking role for tenascins. *Structure* **12**, 1495-1506, doi:10.1016/j.str.2004.05.021 (2004).
- 97 Aspberg, A., Binkert, C. & Ruoslahti, E. The versican C-type lectin domain recognizes the adhesion protein tenascin-R. *Proc Natl Acad Sci U S A* **92**, 10590-10594, doi:10.1073/pnas.92.23.10590 (1995).
- 98 Weber, P. *et al.* Mice deficient for tenascin-R display alterations of the extracellular matrix and decreased axonal conduction velocities in the CNS. *J Neurosci* **19**, 4245-4262, doi:10.1523/JNEUROSCI.19-11-04245.1999 (1999).
- 99 Yamagata, M., Yamada, K. M., Yoneda, M., Suzuki, S. & Kimata, K. Chondroitin sulfate proteoglycan (PG-M-like proteoglycan) is involved in the binding of hyaluronic acid to cellular fibronectin. *J Biol Chem* **261**, 13526-13535 (1986).
- 100 Chen, D. *et al.* Versican binds collagen via its G3 domain and regulates the organization and mechanics of collagenous matrices. *bioRxiv*, 2022.2003.2027.485990, doi:10.1101/2022.03.27.485990 %J bioRxiv (2022).
- 101 Norian, J. M. *et al.* Transforming growth factor beta3 regulates the versican variants in the extracellular matrix-rich uterine leiomyomas. *Reprod Sci* **16**, 1153-1164, doi:10.1177/1933719109343310 (2009).
- 102 Berdiaki, A. *et al.* Regulation of hyaluronan and versican deposition by growth factors in fibrosarcoma cell lines. *Biochim Biophys Acta* **1780**, 194-202, doi:10.1016/j.bbagen.2007.10.005 (2008).
- 103 Schonherr, E., Jarvelainen, H. T., Sandell, L. J. & Wight, T. N. Effects of platelet-derived growth factor and transforming growth factor-beta 1 on the synthesis of a large versican-like chondroitin sulfate proteoglycan by arterial smooth muscle cells. *J Biol Chem* **266**, 17640-17647 (1991).

- 104 Domenzain-Reyna, C. *et al.* Structure and regulation of the versican promoter: the versican promoter is regulated by AP-1 and TCF transcription factors in invasive human melanoma cells. *J Biol Chem* **284**, 12306-12317, doi:10.1074/jbc.M807108200 (2009).
- 105 Rahmani, M. *et al.* Regulation of the versican promoter by the beta-catenin-T-cell factor complex in vascular smooth muscle cells. *J Biol Chem* **280**, 13019-13028, doi:10.1074/jbc.M411766200 (2005).
- 106 Kern, C. B. *et al.* Proteolytic cleavage of versican during cardiac cushion morphogenesis. *Dev Dyn* **235**, 2238-2247, doi:10.1002/dvdy.20838 (2006).
- 107 Faggian, J., Fosang, A. J., Zieba, M., Wallace, M. J. & Hooper, S. B. Changes in versican and chondroitin sulfate proteoglycans during structural development of the lung. *Am J Physiol Regul Integr Comp Physiol* **293**, R784-792, doi:10.1152/ajpregu.00801.2006 (2007).
- 108 Bode-Lesniewska, B. *et al.* Distribution of the large aggregating proteoglycan versican in adult human tissues. *J Histochem Cytochem* **44**, 303-312, doi:10.1177/44.4.8601689 (1996).
- 109 Cattaruzza, S. *et al.* Distribution of PG-M/versican variants in human tissues and de novo expression of isoform V3 upon endothelial cell activation, migration, and neoangiogenesis in vitro. *J Biol Chem* **277**, 47626-47635, doi:10.1074/jbc.M206521200 (2002).
- 110 Dours-Zimmermann, M. T. *et al.* Versican V2 assembles the extracellular matrix surrounding the nodes of ranvier in the CNS. *J Neurosci* **29**, 7731-7742, doi:10.1523/JNEUROSCI.4158-08.2009 (2009).
- 111 Theocharis, A. D., Tsolakis, I., Hjerpe, A. & Karamanos, N. K. Human abdominal aortic aneurysm is characterized by decreased versican concentration and specific downregulation of versican isoform V(0). *Atherosclerosis* **154**, 367-376, doi:10.1016/s0021-9150(00)00504-9 (2001).
- 112 Merrilees, M. J. *et al.* Retrovirally mediated overexpression of versican v3 by arterial smooth muscle cells induces tropoelastin synthesis and elastic fiber formation in vitro and in neointima after vascular injury. *Circ Res* **90**, 481-487, doi:10.1161/hh0402.105791 (2002).
- 113 Passi, A., Negrini, D., Albertini, R., Miserocchi, G. & De Luca, G. The sensitivity of versican from rabbit lung to gelatinase A (MMP-2) and B (MMP-9) and its involvement in the development of hydraulic lung edema. *FEBS Lett* **456**, 93-96, doi:10.1016/s0014-5793(99)00929-1 (1999).
- 114 Kenagy, R. D. *et al.* Increased plasmin and serine proteinase activity during flow-induced intimal atrophy in baboon PTFE grafts. *Arterioscler Thromb Vasc Biol* **22**, 400-404, doi:10.1161/hq0302.105376 (2002).
- 115 Nandadasa, S., Foulcer, S. & Apte, S. S. The multiple, complex roles of versican and its proteolytic turnover by ADAMTS proteases during embryogenesis. *Matrix Biol* **35**, 34-41, doi:10.1016/j.matbio.2014.01.005 (2014).
- 116 Foulcer, S. J. *et al.* Determinants of versican-V1 proteoglycan processing by the metalloproteinase ADAMTS5. *J Biol Chem* **289**, 27859-27873, doi:10.1074/jbc.M114.573287 (2014).
- 117 Westling, J. *et al.* ADAMTS4 (aggrecanase-1) cleaves human brain versican V2 at Glu405-Gln406 to generate glial hyaluronate binding protein. *Biochem J* **377**, 787-795, doi:10.1042/BJ20030896 (2004).
- 118 Martin, D. R., Santamaria, S., Koch, C. D., Ahnstrom, J. & Apte, S. S. Identification of novel ADAMTS1, ADAMTS4 and ADAMTS5 cleavage sites in versican using a label-free quantitative proteomics approach. *J Proteomics* **249**, 104358, doi:10.1016/j.jprot.2021.104358 (2021).
- 119 Tian, C. *et al.* Proteomic analyses of ECM during pancreatic ductal adenocarcinoma progression reveal different contributions by tumor and

- stromal cells. *Proc Natl Acad Sci U S A* **116**, 19609-19618, doi:10.1073/pnas.1908626116 (2019).
- 120 Li, D. *et al.* Tumor-produced versican V1 enhances hCAP18/LL-37 expression in macrophages through activation of TLR2 and vitamin D3 signaling to promote ovarian cancer progression in vitro. *PLoS One* **8**, e56616, doi:10.1371/journal.pone.0056616 (2013).
- 121 Xia, L. *et al.* Forkhead box Q1 promotes hepatocellular carcinoma metastasis by transactivating ZEB2 and VersicanV1 expression. *Hepatology* **59**, 958-973, doi:10.1002/hep.26735 (2014).
- 122 Bogels, M. *et al.* Carcinoma origin dictates differential skewing of monocyte function. *Oncoimmunology* **1**, 798-809, doi:10.4161/onci.20427 (2012).
- 123 Hu, F. *et al.* Glioma-derived versican promotes tumor expansion via glioma-associated microglial/macrophages Toll-like receptor 2 signaling. *Neuro Oncol* **17**, 200-210, doi:10.1093/neuonc/nou324 (2015).
- 124 Ricciardelli, C. *et al.* Regulation of stromal versican expression by breast cancer cells and importance to relapse-free survival in patients with node-negative primary breast cancer. *Clin Cancer Res* **8**, 1054-1060 (2002).
- 125 Kodama, J. *et al.* Versican expression in human cervical cancer. *Eur J Cancer* **43**, 1460-1466, doi:10.1016/j.ejca.2007.02.007 (2007).
- 126 Yeung, T. L. *et al.* TGF-beta modulates ovarian cancer invasion by upregulating CAF-derived versican in the tumor microenvironment. *Cancer Res* **73**, 5016-5028, doi:10.1158/0008-5472.CAN-13-0023 (2013).
- 127 Pukkila, M. *et al.* High stromal versican expression predicts unfavourable outcome in oral squamous cell carcinoma. *J Clin Pathol* **60**, 267-272, doi:10.1136/jcp.2005.034181 (2007).
- 128 Sakko, A. J. *et al.* Versican accumulation in human prostatic fibroblast cultures is enhanced by prostate cancer cell-derived transforming growth factor beta1. *Cancer Res* **61**, 926-930 (2001).
- 129 Grout, J. A. *et al.* Spatial Positioning and Matrix Programs of Cancer-Associated Fibroblasts Promote T-cell Exclusion in Human Lung Tumors. *Cancer Discov* **12**, 2606-2625, doi:10.1158/2159-8290.CD-21-1714 (2022).
- 130 Gao, D., Vahdat, L. T., Wong, S., Chang, J. C. & Mittal, V. Microenvironmental regulation of epithelial-mesenchymal transitions in cancer. *Cancer Res* **72**, 4883-4889, doi:10.1158/0008-5472.CAN-12-1223 (2012).
- 131 Cheng, S. *et al.* A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* **184**, 792-809 e723, doi:10.1016/j.cell.2021.01.010 (2021).
- 132 Tsidulko, A. Y., Matskova, L., Astakhova, L. A., Ernberg, I. & Grigorieva, E. V. Proteoglycan expression correlates with the phenotype of malignant and non-malignant EBV-positive B-cell lines. *Oncotarget* **6**, 43529-43539, doi:10.18632/oncotarget.5984 (2015).
- 133 Zhang, Q. *et al.* Upregulation of Versican Associated with Tumor Progression, Metastasis, and Poor Prognosis in Bladder Carcinoma. *Biomed Res Int* **2021**, 6949864, doi:10.1155/2021/6949864 (2021).
- 134 Kodama, J. *et al.* Prognostic significance of stromal versican expression in human endometrial cancer. *Ann Oncol* **18**, 269-274, doi:10.1093/annonc/mdl370 (2007).
- 135 Mitsui, Y. *et al.* Versican Promotes Tumor Progression, Metastasis and Predicts Poor Prognosis in Renal Carcinoma. *Mol Cancer Res* **15**, 884-895, doi:10.1158/1541-7786.MCR-16-0444 (2017).
- 136 Song, J., Wei, R., Huo, S., Liu, C. & Liu, X. Versican enrichment predicts poor prognosis and response to adjuvant therapy and immunotherapy in gastric cancer. *Front Immunol* **13**, 960570, doi:10.3389/fimmu.2022.960570 (2022).

- 137 Kim, N. S., Lee, H. H., Jung, C. K. & Jeon, H. M. Versican expression in tumor epithelial cells is correlated with a good prognosis in gastric cancer. *Anticancer Res* **34**, 5613-5619 (2014).
- 138 de Wit, M. *et al.* Lumican and versican are associated with good outcome in stage II and III colon cancer. *Ann Surg Oncol* **20 Suppl 3**, S348-359, doi:10.1245/s10434-012-2441-0 (2013).
- 139 Gao, H. *et al.* The expression of versican and its role in pancreatic neuroendocrine tumors. *Pancreatology* **20**, 142-147, doi:10.1016/j.pan.2019.11.009 (2020).
- 140 Asimakopoulos, F. *et al.* Versican proteolysis is associated with robust CD8+ T-cell infiltration in human mismatch repair-proficient and-deficient colorectal cancers. *The Journal of Immunology* **198**, 76.28-76.28 (2017).
- 141 Dhakal, B. *et al.* Versican proteolysis predicts immune effector infiltration and post-transplant survival in myeloma. *Leuk Lymphoma* **60**, 2558-2562, doi:10.1080/10428194.2019.1585836 (2019).
- 142 Arslan, F. *et al.* The role of versican isoforms V0/V1 in glioma migration mediated by transforming growth factor-beta2. *Br J Cancer* **96**, 1560-1568, doi:10.1038/sj.bjc.6603766 (2007).
- 143 Touab, M., Villena, J., Barranco, C., Arumi-Uria, M. & Bassols, A. Versican is differentially expressed in human melanoma and may play a role in tumor development. *Am J Pathol* **160**, 549-557, doi:10.1016/S0002-9440(10)64874-2 (2002).
- 144 LaPierre, D. P. *et al.* The ability of versican to simultaneously cause apoptotic resistance and sensitivity. *Cancer Res* **67**, 4742-4750, doi:10.1158/0008-5472.CAN-06-3610 (2007).
- 145 Wasa, J. *et al.* Versican V1 isoform regulates cell-associated matrix formation and cell behavior differentially from aggrecan in Swarm rat chondrosarcoma cells. *Int J Cancer* **130**, 2271-2281, doi:10.1002/ijc.26230 (2012).
- 146 Sheng, W. *et al.* The roles of versican V1 and V2 isoforms in cell proliferation and apoptosis. *Mol Biol Cell* **16**, 1330-1340, doi:10.1091/mbc.e04-04-0295 (2005).
- 147 Domenzain, C., Docampo, M. J., Serra, M., Miquel, L. & Bassols, A. Differential expression of versican isoforms is a component of the human melanoma cell differentiation process. *Biochim Biophys Acta* **1642**, 107-114, doi:10.1016/s0167-4889(03)00104-6 (2003).
- 148 Yang, W. & Yee, A. J. Versican V2 isoform enhances angiogenesis by regulating endothelial cell activities and fibronectin expression. *FEBS Lett* **587**, 185-192, doi:10.1016/j.febslet.2012.11.023 (2013).
- 149 Lee, H. C. *et al.* Cancer metastasis and EGFR signaling is suppressed by amiodarone-induced versican V2. *Oncotarget* **6**, 42976-42987, doi:10.18632/oncotarget.5621 (2015).
- 150 Hernandez, D. *et al.* V3 versican isoform alters the behavior of human melanoma cells by interfering with CD44/ErbB-dependent signaling. *J Biol Chem* **286**, 1475-1485, doi:10.1074/jbc.M110.127522 (2011).
- 151 Miquel-Serra, L. *et al.* V3 versican isoform expression has a dual role in human melanoma tumor growth and metastasis. *Lab Invest* **86**, 889-901, doi:10.1038/labinvest.3700449 (2006).
- 152 Proudfoot, A. E. *et al.* Glycosaminoglycan binding and oligomerization are essential for the in vivo activity of certain chemokines. *Proc Natl Acad Sci U S A* **100**, 1885-1890, doi:10.1073/pnas.0334864100 (2003).
- 153 Crijns, H., Vanheule, V. & Proost, P. Targeting Chemokine-Glycosaminoglycan Interactions to Inhibit Inflammation. *Front Immunol* **11**, 483, doi:10.3389/fimmu.2020.00483 (2020).

- 154 Hirose, J., Kawashima, H., Yoshie, O., Tashiro, K. & Miyasaka, M. Versican interacts with chemokines and modulates cellular responses. *J Biol Chem* **276**, 5228-5234, doi:10.1074/jbc.M007542200 (2001).
- 155 Masuda, A. *et al.* Versican is upregulated in circulating monocytes in patients with systemic sclerosis and amplifies a CCL2-mediated pathogenic loop. *Arthritis Res Ther* **15**, R74, doi:10.1186/ar4251 (2013).
- 156 Zhou, Z. H. *et al.* Oversulfated chondroitin sulfate binds to chemokines and inhibits stromal cell-derived factor-1 mediated signaling in activated T cells. *PLoS One* **9**, e94402, doi:10.1371/journal.pone.0094402 (2014).
- 157 Kim, S. *et al.* Carcinoma-produced factors activate myeloid cells through TLR2 to stimulate metastasis. *Nature* **457**, 102-106, doi:10.1038/nature07623 (2009).
- 158 Tang, M. *et al.* Toll-like Receptor 2 Activation Promotes Tumor Dendritic Cell Dysfunction by Regulating IL-6 and IL-10 Receptor Signaling. *Cell Rep* **13**, 2851-2864, doi:10.1016/j.celrep.2015.11.053 (2015).
- 159 Tang, F. *et al.* Defining the versican interactome in lung health and disease. *Am J Physiol Cell Physiol* **323**, C249-C276, doi:10.1152/ajpcell.00162.2022 (2022).
- 160 Papadas, A. *et al.* Stromal remodeling regulates dendritic cell abundance and activity in the tumor microenvironment. *Cell Rep* **40**, 111201, doi:10.1016/j.celrep.2022.111201 (2022).
- 161 Kawashima, H. *et al.* Binding of a large chondroitin sulfate/dermatan sulfate proteoglycan, versican, to L-selectin, P-selectin, and CD44. *J Biol Chem* **275**, 35448-35456, doi:10.1074/jbc.M003387200 (2000).
- 162 Evanko, S. P., Potter-Perigo, S., Bollyky, P. L., Nepom, G. T. & Wight, T. N. Hyaluronan and versican in the control of human T-lymphocyte adhesion and migration. *Matrix Biol* **31**, 90-100, doi:10.1016/j.matbio.2011.10.004 (2012).
- 163 Zheng, P. S. *et al.* PG-M/versican binds to P-selectin glycoprotein ligand-1 and mediates leukocyte aggregation. *J Cell Sci* **117**, 5887-5895, doi:10.1242/jcs.01516 (2004).
- 164 Papadas, A., Arauz, G., Cicala, A., Wiesner, J. & Asimakopoulos, F. Versican and Versican-matrikines in Cancer Progression, Inflammation, and Immunity. *J Histochem Cytochem* **68**, 871-885, doi:10.1369/0022155420937098 (2020).
- 165 Nolan, E., Lindeman, G. J. & Visvader, J. E. Deciphering breast cancer: from biology to the clinic. *Cell* **186**, 1708-1728, doi:10.1016/j.cell.2023.01.040 (2023).
- 166 Harbeck, N. *et al.* Breast cancer. *Nat Rev Dis Primers* **5**, 66, doi:10.1038/s41572-019-0111-2 (2019).
- 167 Lukasiewicz, S. *et al.* Breast Cancer-Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies-An Updated Review. *Cancers (Basel)* **13**, doi:10.3390/cancers13174287 (2021).
- 168 Robson, M. *et al.* Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med* **377**, 523-533, doi:10.1056/NEJMoa1706450 (2017).
- 169 Angeli, D., Salvi, S. & Tedaldi, G. Genetic Predisposition to Breast and Ovarian Cancers: How Many and Which Genes to Test? *Int J Mol Sci* **21**, doi:10.3390/ijms21031128 (2020).
- 170 McGuire, A., Brown, J. A., Malone, C., McLaughlin, R. & Kerin, M. J. Effects of age on the detection and management of breast cancer. *Cancers (Basel)* **7**, 908-929, doi:10.3390/cancers7020815 (2015).
- 171 Hill, D. A., Prossnitz, E. R., Royce, M. & Nibbe, A. Temporal trends in breast cancer survival by race and ethnicity: A population-based cohort study. *PLoS One* **14**, e0224064, doi:10.1371/journal.pone.0224064 (2019).

- 172 Sun, L., Zhu, Y., Qian, Q. & Tang, L. Body mass index and prognosis of breast cancer: An analysis by menstruation status when breast cancer diagnosis. *Medicine (Baltimore)* **97**, e11220, doi:10.1097/MD.00000000000011220 (2018).
- 173 Lynch, B. M., Neilson, H. K. & Friedenreich, C. M. Physical activity and breast cancer prevention. *Recent Results Cancer Res* **186**, 13-42, doi:10.1007/978-3-642-04231-7\_2 (2011).
- 174 Zeinomar, N. *et al.* Alcohol consumption, cigarette smoking, and familial breast cancer risk: findings from the Prospective Family Study Cohort (ProF-SC). *Breast Cancer Res* **21**, 128, doi:10.1186/s13058-019-1213-1 (2019).
- 175 Albrektsen, G., Heuch, I., Hansen, S. & Kvale, G. Breast cancer risk by age at birth, time since birth and time intervals between births: exploring interaction effects. *Br J Cancer* **92**, 167-175, doi:10.1038/sj.bjc.6602302 (2005).
- 176 Beral, V. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer. *Lancet* **350**, 1047-1059 (1997).
- 177 Giordano, S. H. Breast Cancer in Men. *N Engl J Med* **378**, 2311-2320, doi:10.1056/NEJMra1707939 (2018).
- 178 Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752, doi:10.1038/35021093 (2000).
- 179 Orrantia-Borunda, E., Anchondo-Nuñez, P., Acuña-Aguilar, L. E., Gómez-Valles, F. O. & Ramírez-Valdespino, C. A. J. B. C. Subtypes of breast cancer. (2022).
- 180 Baranova, A. *et al.* Triple-negative breast cancer: current treatment strategies and factors of negative prognosis. *J Med Life* **15**, 153-161, doi:10.25122/jml-2021-0108 (2022).
- 181 Dent, R. *et al.* Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res* **13**, 4429-4434, doi:10.1158/1078-0432.CCR-06-3045 (2007).
- 182 Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* **121**, 2750-2767, doi:10.1172/JCI45014 (2011).
- 183 Yin, L., Duan, J. J., Bian, X. W. & Yu, S. C. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res* **22**, 61, doi:10.1186/s13058-020-01296-5 (2020).
- 184 Lehmann, B. D. *et al.* Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One* **11**, e0157368, doi:10.1371/journal.pone.0157368 (2016).
- 185 Burstein, M. D. *et al.* Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin Cancer Res* **21**, 1688-1698, doi:10.1158/1078-0432.CCR-14-0432 (2015).
- 186 Elsayaf, Z. *et al.* Biological subtypes of triple-negative breast cancer are associated with distinct morphological changes and clinical behaviour. *Breast* **22**, 986-992, doi:10.1016/j.breast.2013.05.012 (2013).
- 187 Tutt, A. *et al.* Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT Trial. *Nat Med* **24**, 628-637, doi:10.1038/s41591-018-0009-7 (2018).
- 188 Algan, O., Zhao, Y. D. & Herman, T. Radiotherapy in Patients 70 Years and Older With Triple-Negative Breast Cancer. *Clin Breast Cancer* **16**, e99-e106, doi:10.1016/j.clbc.2016.05.011 (2016).

- 189 Zhai, Z. *et al.* Evaluation of Adjuvant Treatments for T1 N0 M0 Triple-Negative Breast Cancer. *JAMA Netw Open* **3**, e2021881, doi:10.1001/jamanetworkopen.2020.21881 (2020).
- 190 de Faria Bessa, J. & Marta, G. N. Triple-negative breast cancer and radiation therapy. *Rep Pract Oncol Radiother* **27**, 545-551, doi:10.5603/RPOR.a2022.0025 (2022).
- 191 Tramm, T. *et al.* Relationship between the prognostic and predictive value of the intrinsic subtypes and a validated gene profile predictive of loco-regional control and benefit from post-mastectomy radiotherapy in patients with high-risk breast cancer. *Acta Oncol* **53**, 1337-1346, doi:10.3109/0284186X.2014.925580 (2014).
- 192 Barchiesi, G. *et al.* Emerging Role of PARP Inhibitors in Metastatic Triple Negative Breast Cancer. Current Scenario and Future Perspectives. *Front Oncol* **11**, 769280, doi:10.3389/fonc.2021.769280 (2021).
- 193 Robson, M. E. *et al.* OlympiAD final overall survival and tolerability results: Olaparib versus chemotherapy treatment of physician's choice in patients with a germline BRCA mutation and HER2-negative metastatic breast cancer. *Ann Oncol* **30**, 558-566, doi:10.1093/annonc/mdz012 (2019).
- 194 Khadela, A. *et al.* Anti-Androgenic Therapies Targeting the Luminal Androgen Receptor of a Typical Triple-Negative Breast Cancer. *Cancers (Basel)* **15**, doi:10.3390/cancers15010233 (2022).
- 195 Nanda, R. *et al.* Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. *J Clin Oncol* **34**, 2460-2467, doi:10.1200/JCO.2015.64.8931 (2016).
- 196 Adams, S. *et al.* Pembrolizumab monotherapy for previously untreated, PD-L1-positive, metastatic triple-negative breast cancer: cohort B of the phase II KEYNOTE-086 study. *Ann Oncol* **30**, 405-411, doi:10.1093/annonc/mdy518 (2019).
- 197 Winer, E. P. *et al.* Pembrolizumab versus investigator-choice chemotherapy for metastatic triple-negative breast cancer (KEYNOTE-119): a randomised, open-label, phase 3 trial. *Lancet Oncol* **22**, 499-511, doi:10.1016/S1470-2045(20)30754-3 (2021).
- 198 Dirix, L. Y. *et al.* Avelumab, an anti-PD-L1 antibody, in patients with locally advanced or metastatic breast cancer: a phase 1b JAVELIN Solid Tumor study. *Breast Cancer Res Treat* **167**, 671-686, doi:10.1007/s10549-017-4537-5 (2018).
- 199 Emens, L. A. *et al.* Long-term Clinical Outcomes and Biomarker Analyses of Atezolizumab Therapy for Patients With Metastatic Triple-Negative Breast Cancer: A Phase 1 Study. *JAMA Oncol* **5**, 74-82, doi:10.1001/jamaoncol.2018.4224 (2019).
- 200 Ahlmann, M. & Hempel, G. The effect of cyclophosphamide on the immune system: implications for clinical cancer therapy. *Cancer Chemother Pharmacol* **78**, 661-671, doi:10.1007/s00280-016-3152-1 (2016).
- 201 Zitvogel, L., Tesniere, A. & Kroemer, G. Cancer despite immunosurveillance: immunoselection and immunosubversion. *Nat Rev Immunol* **6**, 715-727, doi:10.1038/nri1936 (2006).
- 202 Cortes, J. *et al.* Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for previously untreated locally recurrent inoperable or metastatic triple-negative breast cancer (KEYNOTE-355): a randomised, placebo-controlled, double-blind, phase 3 clinical trial. *Lancet* **396**, 1817-1828, doi:10.1016/S0140-6736(20)32531-9 (2020).
- 203 Schmid, P. *et al.* Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. *N Engl J Med* **379**, 2108-2121, doi:10.1056/NEJMoa1809615 (2018).

- 204 Miles, D. *et al.* Primary results from IMpassion131, a double-blind, placebo-controlled, randomised phase III trial of first-line paclitaxel with or without atezolizumab for unresectable locally advanced/metastatic triple-negative breast cancer. *Ann Oncol* **32**, 994-1004, doi:10.1016/j.annonc.2021.05.801 (2021).
- 205 Schmid, P. *et al.* Pembrolizumab for Early Triple-Negative Breast Cancer. *N Engl J Med* **382**, 810-821, doi:10.1056/NEJMoa1910549 (2020).
- 206 Li, L., Zhang, F., Liu, Z. & Fan, Z. Immunotherapy for Triple-Negative Breast Cancer: Combination Strategies to Improve Outcome. *Cancers (Basel)* **15**, doi:10.3390/cancers15010321 (2023).
- 207 Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci Rep* **7**, 16878, doi:10.1038/s41598-017-17204-5 (2017).
- 208 Ehlicke, F., Koster, N., Salzig, D. & Czermak, P. Non-invasive Raman Spectroscopy and Quantitative Real-Time PCR Distinguish Among Undifferentiated Human Mesenchymal Stem Cells and Redifferentiated Nucleus Pulposus Cells and Chondrocytes In Vitro. *Open Biomed Eng J* **11**, 72-84, doi:10.2174/1874120701711010072 (2017).
- 209 Song, R. H. *et al.* Aggrecan degradation in human articular cartilage explants is mediated by both ADAMTS-4 and ADAMTS-5. *Arthritis Rheum* **56**, 575-585, doi:10.1002/art.22334 (2007).
- 210 Yang, P., Bu, P. & Li, C. miR-124 inhibits proliferation, migration and invasion of malignant melanoma cells via targeting versican. *Exp Ther Med* **14**, 3555-3562, doi:10.3892/etm.2017.4998 (2017).
- 211 Osuna de la Pena, D. *et al.* Bioengineered 3D models of human pancreatic cancer recapitulate in vivo tumour biology. *Nat Commun* **12**, 5623, doi:10.1038/s41467-021-25921-9 (2021).
- 212 Alonge, K. M. *et al.* Changes in Brain Matrix Glycan Sulfation Associate With Reactive Gliosis and Motor Coordination in Mice With Head Trauma. *Front Behav Neurosci* **15**, 745288, doi:10.3389/fnbeh.2021.745288 (2021).
- 213 Alonge, K. M. *et al.* Hypothalamic perineuronal net assembly is required for sustained diabetes remission induced by fibroblast growth factor 1 in rats. *Nat Metab* **2**, 1025-1033, doi:10.1038/s42255-020-00275-6 (2020).
- 214 Alonge, K. M. *et al.* Quantitative analysis of chondroitin sulfate disaccharides from human and rodent fixed brain tissue by electrospray ionization-tandem mass spectrometry. *Glycobiology* **29**, 847-860, doi:10.1093/glycob/cwz060 (2019).
- 215 Li, X. *et al.* Infiltration of CD8(+) T cells into tumor cell clusters in triple-negative breast cancer. *Proc Natl Acad Sci U S A* **116**, 3678-3687, doi:10.1073/pnas.1817652116 (2019).
- 216 Derks, S. *et al.* Characterizing diversity in the tumor-immune microenvironment of distinct subclasses of gastroesophageal adenocarcinomas. *Ann Oncol* **31**, 1011-1020, doi:10.1016/j.annonc.2020.04.011 (2020).
- 217 Kalecky, K., Modisette, R., Pena, S., Cho, Y.-R. & Taube, J. - Integrative analysis of breast cancer profiles in TCGA by TNBC subgrouping reveals novel microRNA-specific clusters, including miR-17-92a, distinguishing basal-like 1 and basal-like 2 TNBC subtypes. - **20** (2020).
- 218 Wu, S. Z. *et al.* Stromal cell diversity associated with immune evasion in human triple-negative breast cancer. *The EMBO Journal* **39**, e104063, doi:<https://doi.org/10.15252/embj.2019104063> (2020).
- 219 Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol* **1711**, 243-259, doi:10.1007/978-1-4939-7493-1\_12 (2018).
- 220 Ween, M. P., Oehler, M. K. & Ricciardelli, C. J. I. j. o. m. s. Role of versican, hyaluronan and CD44 in ovarian cancer metastasis. **12**, 1009-1029 (2011).



- 221 Wu, Y. J., La Pierre, D. P., Jin, W., Albert, J. Y. & Burton, B. Y. The interaction of versican with its binding partners. *Cell research* **15**, 483-494 (2005).
- 222 Koppensteiner, L., Mathieson, L., O'Connor, R. A. & Akram, A. R. Cancer Associated Fibroblasts - An Impediment to Effective Anti-Cancer T Cell Immunity. *Front Immunol* **13**, 887380, doi:10.3389/fimmu.2022.887380 (2022).
- 223 Wight, T. N., Kinsella, M. G., Evanko, S. P., Potter-Perigo, S. & Merrilees, M. J. Versican and the regulation of cell phenotype in disease. *Biochim Biophys Acta* **1840**, 2441-2451, doi:10.1016/j.bbagen.2013.12.028 (2014).
- 224 Egelston, C. A. *et al.* Resident memory CD8+ T cells within cancer islands mediate survival in breast cancer patients. *JCI Insight* (2019).
- 225 Hegde, P. S., Karanikas, V. & Evers, S. The Where, the When, and the How of Immune Monitoring for Cancer Immunotherapies in the Era of Checkpoint Inhibition. *Clinical Cancer Research* **22**, 1865-1874, doi:10.1158/1078-0432.Ccr-15-1507 (2016).
- 226 Adams, S. *et al.* Pembrolizumab monotherapy for previously untreated, PD-L1-positive, metastatic triple-negative breast cancer: cohort B of the phase II KEYNOTE-086 study. *Annals of Oncology* **30**, 405-411, doi:<https://doi.org/10.1093/annonc/mdy518> (2019).
- 227 Wu, W., Liu, Y., Zeng, S., Han, Y. & Shen, H. Intratumor heterogeneity: the hidden barrier to immunotherapy against MSI tumors from the perspective of IFN-gamma signaling and tumor-infiltrating lymphocytes. *J Hematol Oncol* **14**, 160, doi:10.1186/s13045-021-01166-3 (2021).
- 228 Chakravarthy, A., Khan, L., Bensler, N. P., Bose, P. & De Carvalho, D. D. TGF-beta-associated extracellular matrix genes link cancer-associated fibroblasts to immune evasion and immunotherapy failure. *Nat Commun* **9**, 4692, doi:10.1038/s41467-018-06654-8 (2018).
- 229 Chen, C. C. *et al.* The endothelin-integrin axis is involved in macrophage-induced breast cancer cell chemotactic interactions with endothelial cells. (2014).
- 230 Hirose, J., Kawashima, H., Yoshie, O., Tashiro, K. & Miyasaka, M. Versican interacts with chemokines and modulates cellular responses. *Journal of Biological Chemistry* **276**, 5228-5234 (2001).
- 231 Masuda, A. *et al.* Versican is upregulated in circulating monocytes in patients with systemic sclerosis and amplifies a CCL2-mediated pathogenic loop. *Arthritis research & therapy* **15**, 1-15 (2013).
- 232 Deligne, C. & Midwood, K. S. Macrophages and Extracellular Matrix in Breast Cancer: Partners in Crime or Protective Allies? *Frontiers in Oncology* **11**, doi:10.3389/fonc.2021.620773 (2021).
- 233 Gondi, C. S. & Rao, J. S. Cathepsin B as a cancer target.
- 234 Gorter, A. *et al.* Versican expression is associated with tumor-infiltrating CD8-positive T cells and infiltration depth in cervical cancer. *Modern Pathology* **23**, 1605-1615 (2010).
- 235 Hope, C. *et al.* Versican-derived matrikines regulate Batf3-dendritic cell differentiation and promote T cell infiltration in colorectal cancer. *The Journal of Immunology* **199**, 1933-1941 (2017).
- 236 Salmon, H. *et al.* Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *Journal of clinical investigation* **122**, 899-910 (2012).
- 237 Ghasempour, S. & Freeman, S. A. The glycocalyx and immune evasion in cancer. *The FEBS Journal* **290**, 55-65, doi:<https://doi.org/10.1111/febs.16236> (2023).

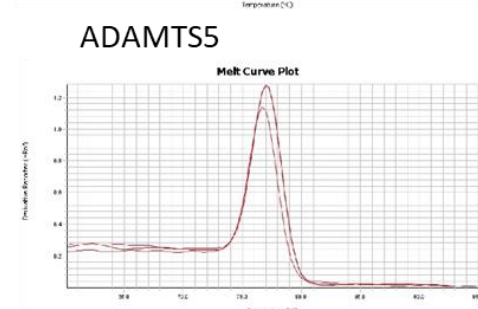
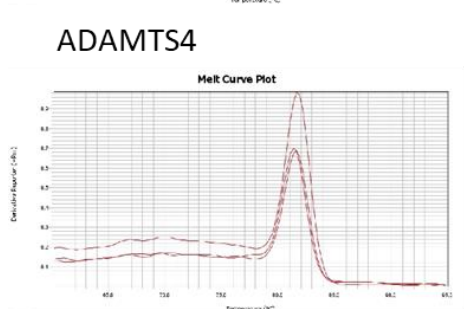
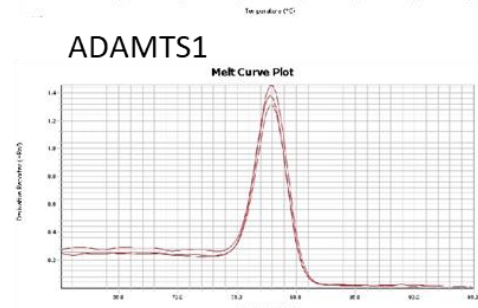
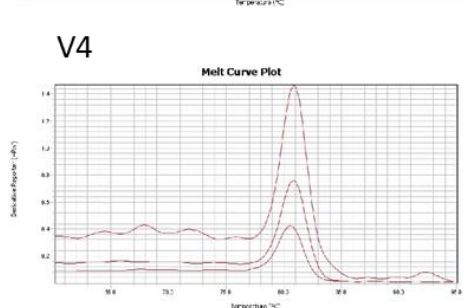
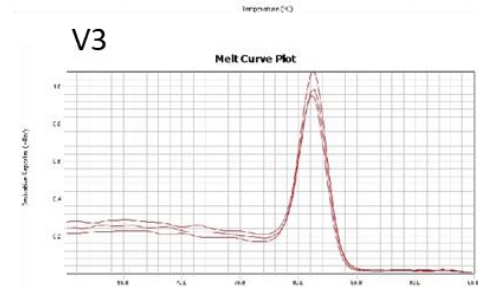
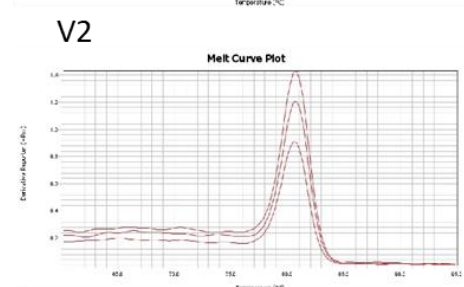
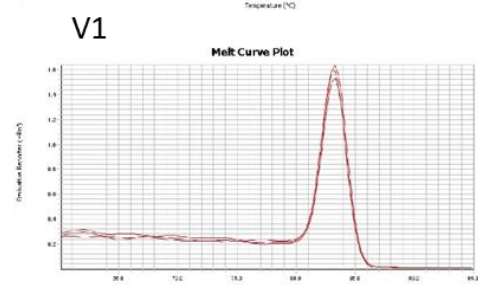
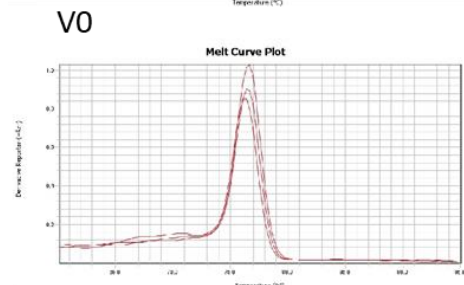
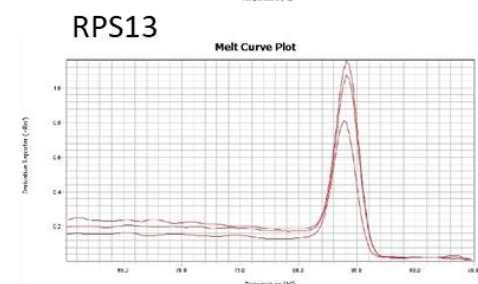
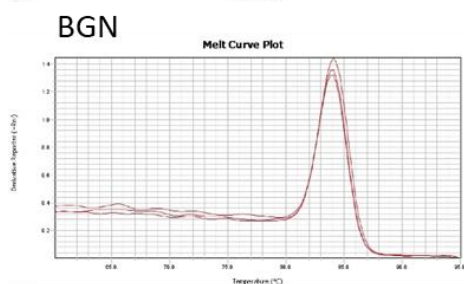
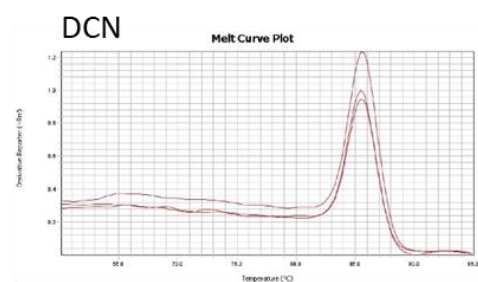
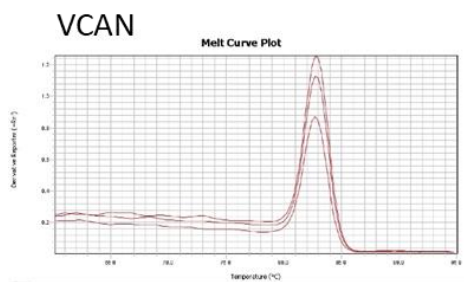
- 238 Sutherland, T. E., Dyer, D. P. & Allen, J. E. The extracellular matrix and the immune system: A mutually dependent relationship. *Science* **379**, eabp8964, doi:10.1126/science.abp8964 (2023).
- 239 Scherzer, M. T. *et al.* Fibroblast-Derived Extracellular Matrices: An Alternative Cell Culture System That Increases Metastatic Cellular Properties. *PLoS One* **10**, e0138065, doi:10.1371/journal.pone.0138065 (2015).
- 240 Ramarajan, M. G. *et al.* Mass spectrometric analysis of chondroitin sulfate-linked peptides. *J Proteins Proteom* **13**, 187-203, doi:10.1007/s42485-022-00092-3 (2022).
- 241 Noborn, F., Nikpour, M., Persson, A., Nilsson, J. & Larson, G. Expanding the Chondroitin Sulfate Glycoproteome - But How Far? *Front Cell Dev Biol* **9**, 695970, doi:10.3389/fcell.2021.695970 (2021).
- 242 Djerbal, L., Lortat-Jacob, H. & Kwok, J. Chondroitin sulfates and their binding molecules in the central nervous system. *Glycoconj J* **34**, 363-376, doi:10.1007/s10719-017-9761-z (2017).
- 243 Wu, S. Z. *et al.* Stromal cell diversity associated with immune evasion in human triple-negative breast cancer. *Embo J* **39**, e104063, doi:10.15252/embj.2019104063 (2020).
- 244 Lauder, R. M., Huckerby, T. N., Brown, G. M., Bayliss, M. T. & Nieduszynski, I. A. Age-related changes in the sulphation of the chondroitin sulphate linkage region from human articular cartilage aggrecan. *Biochem J* **358**, 523-528, doi:10.1042/0264-6021:3580523 (2001).
- 245 Wei, J., Hu, M., Huang, K., Lin, S. & Du, H. Roles of Proteoglycans and Glycosaminoglycans in Cancer Development and Progression. *Int J Mol Sci* **21**, doi:10.3390/ijms21175983 (2020).
- 246 Martinez, J. R., Dhawan, A. & Farach-Carson, M. C. Modular Proteoglycan Perlecan/HSPG2: Mutations, Phenotypes, and Functions. *Genes (Basel)* **9**, doi:10.3390/genes9110556 (2018).
- 247 Zhang, W. *et al.* Decorin is a pivotal effector in the extracellular matrix and tumour microenvironment. *Oncotarget* **9**, 5480-5491, doi:10.18632/oncotarget.23869 (2018).
- 248 Diehl, V. *et al.* The Role of Decorin and Biglycan Signaling in Tumorigenesis. *Front Oncol* **11**, 801801, doi:10.3389/fonc.2021.801801 (2021).
- 249 Kahari, V. M., Larjava, H. & Uitto, J. Differential regulation of extracellular matrix proteoglycan (PG) gene expression. Transforming growth factor-beta 1 up-regulates biglycan (PGI), and versican (large fibroblast PG) but down-regulates decorin (PGII) mRNA levels in human fibroblasts in culture. *J Biol Chem* **266**, 10608-10615 (1991).
- 250 Van Bockstal, M. *et al.* Differential regulation of extracellular matrix protein expression in carcinoma-associated fibroblasts by TGF-beta1 regulates cancer cell spreading but not adhesion. *Oncoscience* **1**, 634-648, doi:10.18632/oncoscience.87 (2014).
- 251 Tan, E. M. *et al.* Decorin, versican, and biglycan gene expression by keloid and normal dermal fibroblasts: differential regulation by basic fibroblast growth factor. *Exp Cell Res* **209**, 200-207, doi:10.1006/excr.1993.1302 (1993).
- 252 Yu, Y. *et al.* Sequencing the Dermatan Sulfate Chain of Decorin. *J Am Chem Soc* **139**, 16986-16995, doi:10.1021/jacs.7b10164 (2017).
- 253 Bianco, P., Fisher, L. W., Young, M. F., Termine, J. D. & Robey, P. G. Expression and localization of the two small proteoglycans biglycan and decorin in developing human skeletal and non-skeletal tissues. *J Histochem Cytochem* **38**, 1549-1563, doi:10.1177/38.11.2212616 (1990).
- 254 Pudelko, A., Wisowski, G., Olczyk, K. & Kozma, E. M. The dual role of the glycosaminoglycan chondroitin-6-sulfate in the development, progression

- and metastasis of cancer. *Febs J* **286**, 1815-1837, doi:10.1111/febs.14748 (2019).
- 255 Murai, T., Sougawa, N., Kawashima, H., Yamaguchi, K. & Miyasaka, M. CD44-chondroitin sulfate interactions mediate leukocyte rolling under physiological flow conditions. *Immunol Lett* **93**, 163-170, doi:10.1016/j.imlet.2004.03.013 (2004).
- 256 Wu, Q. *et al.* Remodeling Chondroitin-6-Sulfate-Mediated Immune Exclusion Enhances Anti-PD-1 Response in Colorectal Cancer with Microsatellite Stability. *Cancer Immunol Res* **10**, 182-199, doi:10.1158/2326-6066.CIR-21-0124 (2022).
- 257 Foulcer, S. J., Day, A. J. & Apte, S. S. Isolation and purification of versican and analysis of versican proteolysis. *Methods Mol Biol* **1229**, 587-604, doi:10.1007/978-1-4939-1714-3\_46 (2015).
- 258 Schmalfeldt, M., Dours-Zimmermann, M. T., Winterhalter, K. H. & Zimmermann, D. R. Versican V2 is a major extracellular matrix component of the mature bovine brain. *J Biol Chem* **273**, 15758-15764, doi:10.1074/jbc.273.25.15758 (1998).
- 259 Evanko, S. P., Chan, C. K., Johnson, P. Y., Frevert, C. W. & Wight, T. N. The biochemistry and immunohistochemistry of versican. *Methods Cell Biol* **143**, 261-279, doi:10.1016/bs.mcb.2017.08.015 (2018).
- 260 Ohno-Jinno, A. *et al.* Versican and fibrillin-1 form a major hyaluronan-binding complex in the ciliary body. *Invest Ophthalmol Vis Sci* **49**, 2870-2877, doi:10.1167/iovs.07-1488 (2008).
- 261 Ishihama, Y. *et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* **4**, 1265-1272, doi:10.1074/mcp.M500061-MCP200 (2005).
- 262 Iozzo, R. V. Matrix proteoglycans: from molecular design to cellular function. *Annu Rev Biochem* **67**, 609-652, doi:10.1146/annurev.biochem.67.1.609 (1998).
- 263 Crossman, M. V. & Mason, R. M. Purification and characterization of a hyaluronan-binding protein from rat chondrosarcoma. *Biochem J* **266**, 399-406, doi:10.1042/bj2660399 (1990).
- 264 Bono, P., Rubin, K., Higgins, J. M. & Hynes, R. O. Layilin, a novel integral membrane protein, is a hyaluronan receptor. *Mol Biol Cell* **12**, 891-900, doi:10.1091/mbc.12.4.891 (2001).
- 265 Vuento, M. & Vaheri, A. Purification of fibronectin from human plasma by affinity chromatography under non-denaturing conditions. *Biochem J* **183**, 331-337, doi:10.1042/bj1830331 (1979).
- 266 Delaine-Smith, R. M. *et al.* Modelling TGFbetaR and Hh pathway regulation of prognostic matrix molecules in ovarian cancer. *iScience* **24**, 102674, doi:10.1016/j.isci.2021.102674 (2021).
- 267 Noborn, F. *et al.* Identification of chondroitin sulfate linkage region glycopeptides reveals prohormones as a novel class of proteoglycans. *Mol Cell Proteomics* **14**, 41-49, doi:10.1074/mcp.M114.043703 (2015).
- 268 Sandy, J. D. *et al.* Versican V1 proteolysis in human aorta in vivo occurs at the Glu441-Ala442 bond, a site that is cleaved by recombinant ADAMTS-1 and ADAMTS-4. *J Biol Chem* **276**, 13372-13378, doi:10.1074/jbc.M009737200 (2001).
- 269 Russell, D. L., Doyle, K. M., Ochsner, S. A., Sandy, J. D. & Richards, J. S. Processing and localization of ADAMTS-1 and proteolytic cleavage of versican during cumulus matrix expansion and ovulation. *J Biol Chem* **278**, 42330-42339, doi:10.1074/jbc.M300519200 (2003).
- 270 Petrella, B. L., Armstrong, D. A. & Vincenti, M. P. Interleukin-1 beta and transforming growth factor-beta 3 cooperate to activate matrix

- metalloproteinase expression and invasiveness in A549 lung adenocarcinoma cells. *Cancer Lett* **325**, 220-226, doi:10.1016/j.canlet.2012.07.009 (2012).
- 271 Gibson, S. V. *et al.* TGF $\beta$ -mediated MMP13 secretion drives myoepithelial cell dependent breast cancer progression. *NPJ Breast Cancer* **9**, 9, doi:10.1038/s41523-023-00513-6 (2023).
- 272 Douglas, T. *et al.* Interactions of collagen types I and II with chondroitin sulfates A-C and their effect on osteoblast adhesion. *Biomacromolecules* **8**, 1085-1092, doi:10.1021/bm0609644 (2007).
- 273 Fowell, D. J. & Kim, M. The spatio-temporal control of effector T cell migration. *Nat Rev Immunol* **21**, 582-596, doi:10.1038/s41577-021-00507-0 (2021).
- 274 Sharma, P. & Allison, J. P. The future of immune checkpoint therapy. *Science* **348**, 56-61, doi:10.1126/science.aaa8172 (2015).
- 275 Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor microenvironment. *Science* **348**, 74-80, doi:10.1126/science.aaa6204 (2015).
- 276 Hegde, P. S., Karanikas, V. & Evers, S. The Where, the When, and the How of Immune Monitoring for Cancer Immunotherapies in the Era of Checkpoint Inhibition. *Clin Cancer Res* **22**, 1865-1874, doi:10.1158/1078-0432.Ccr-15-1507 (2016).
- 277 Nguyen, P. H. D. *et al.* Intratumoural immune heterogeneity as a hallmark of tumour evolution and progression in hepatocellular carcinoma. *Nat Commun* **12**, 227, doi:10.1038/s41467-020-20171-7 (2021).
- 278 Henke, E., Nandigama, R. & Ergun, S. Extracellular Matrix in the Tumor Microenvironment and Its Impact on Cancer Therapy. *Front Mol Biosci* **6**, 160, doi:10.3389/fmolb.2019.00160 (2019).
- 279 Wight, T. N., Kang, I. & Merrilees, M. J. Versican and the control of inflammation. *Matrix Biol* **35**, 152-161, doi:10.1016/j.matbio.2014.01.015 (2014).
- 280 Ping, Q. *et al.* TGF-beta1 dominates stromal fibroblast-mediated EMT via the FAP/VCAN axis in bladder cancer cells. *J Transl Med* **21**, 475, doi:10.1186/s12967-023-04303-3 (2023).
- 281 Faller, C. E. & Guvench, O. Sulfation and cation effects on the conformational properties of the glycan backbone of chondroitin sulfate disaccharides. *J Phys Chem B* **119**, 6063-6073, doi:10.1021/jp511431q (2015).
- 282 Deshauer, C. *et al.* Interactions of the Chemokine CCL5/RANTES with Medium-Sized Chondroitin Sulfate Ligands. *Structure* **23**, 1066-1077, doi:10.1016/j.str.2015.03.024 (2015).
- 283 Murooka, T. T., Rahbar, R., Platanias, L. C. & Fish, E. N. CCL5-mediated T-cell chemotaxis involves the initiation of mRNA translation through mTOR/4E-BP1. *Blood* **111**, 4892-4901, doi:10.1182/blood-2007-11-125039 (2008).
- 284 Taub, D. D., Anver, M., Oppenheim, J. J., Longo, D. L. & Murphy, W. J. T lymphocyte recruitment by interleukin-8 (IL-8). IL-8-induced degranulation of neutrophils releases potent chemoattractants for human T lymphocytes both in vitro and in vivo. *J Clin Invest* **97**, 1931-1941, doi:10.1172/JCI118625 (1996).
- 285 Akiyama, H. *et al.* Chondroitin sulphate structure affects its immunological activities on murine splenocytes sensitized with ovalbumin. *Biochem J* **382**, 269-278, doi:10.1042/bj20031851 (2004).
- 286 Soares da Costa, D., Reis, R. L. & Pashkuleva, I. Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders.

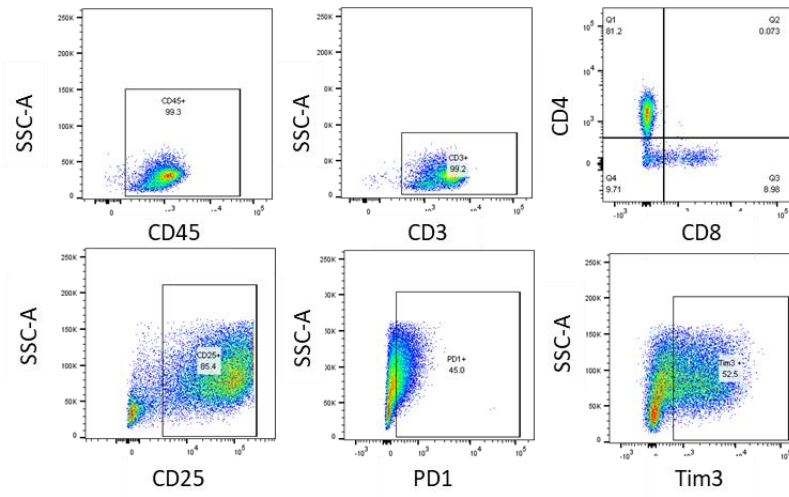
- Annu Rev Biomed Eng* **19**, 1-26, doi:10.1146/annurev-bioeng-071516-044610 (2017).
- 287 Theocharis, A. D., Tsara, M. E., Papageorgacopoulou, N., Karavias, D. D. & Theocharis, D. A. Pancreatic carcinoma is characterized by elevated content of hyaluronan and chondroitin sulfate with altered disaccharide composition. *Biochim Biophys Acta* **1502**, 201-206, doi:10.1016/s0925-4439(00)00051-x (2000).
- 288 Raspanti, M. *et al.* Glycosaminoglycans show a specific periodic interaction with type I collagen fibrils. *J Struct Biol* **164**, 134-139, doi:10.1016/j.jsb.2008.07.001 (2008).
- 289 Bhattacharyya, S., Kotlo, K., Shukla, S., Danziger, R. S. & Tobacman, J. K. Distinct effects of N-acetylgalactosamine-4-sulfatase and galactose-6-sulfatase expression on chondroitin sulfates. *J Biol Chem* **283**, 9523-9530, doi:10.1074/jbc.M707967200 (2008).
- 290 Hatano, S. & Watanabe, H. Regulation of Macrophage and Dendritic Cell Function by Chondroitin Sulfate in Innate to Antigen-Specific Adaptive Immunity. *Front Immunol* **11**, 232, doi:10.3389/fimmu.2020.00232 (2020).

# Appendices



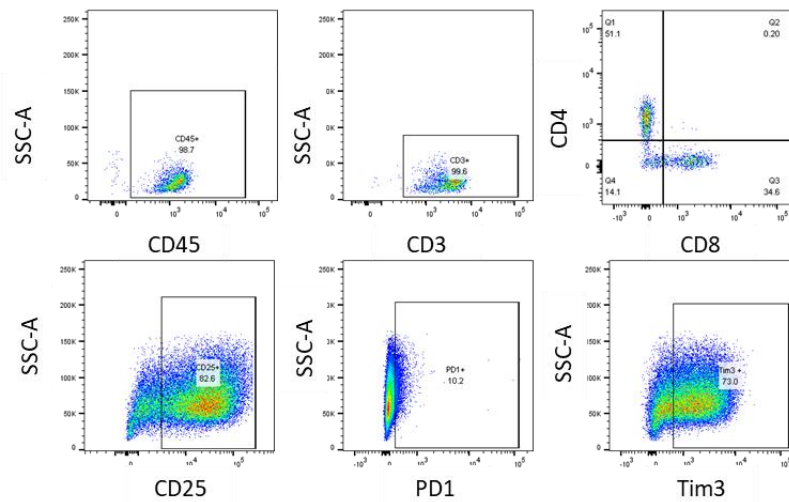
Appendix 1. Melt curve of each primer.

### Donor 1:



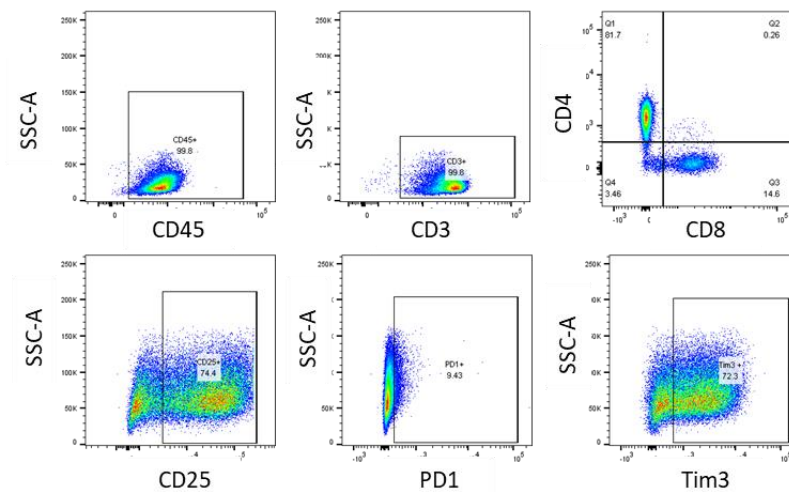
Marker	%
CD3	99.2
CD8	9.7
CD4	81.2
CD25	85.4
PD1	45
Tim3	52.5

### Donor 2:



Marker	%
CD3	99.6
CD8	34.6
CD4	51.1
CD25	82.6
PD1	10.2
Tim3	73

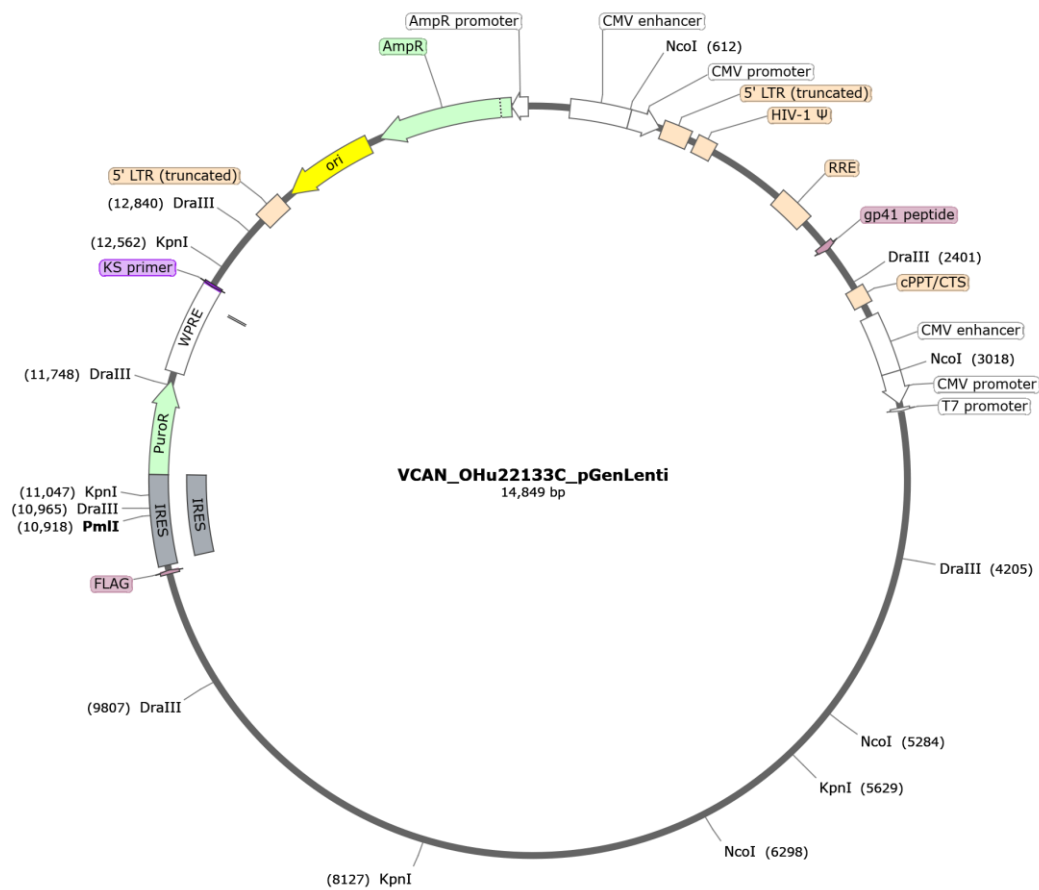
### Donor 3:



Marker	%
CD3	99.8
CD8	14.6
CD4	81.7
CD25	74.4
PD1	9.43
Tim3	72.3

Appendix 2. Flow cytometry gating to confirm T cell isolation and T cell activation.





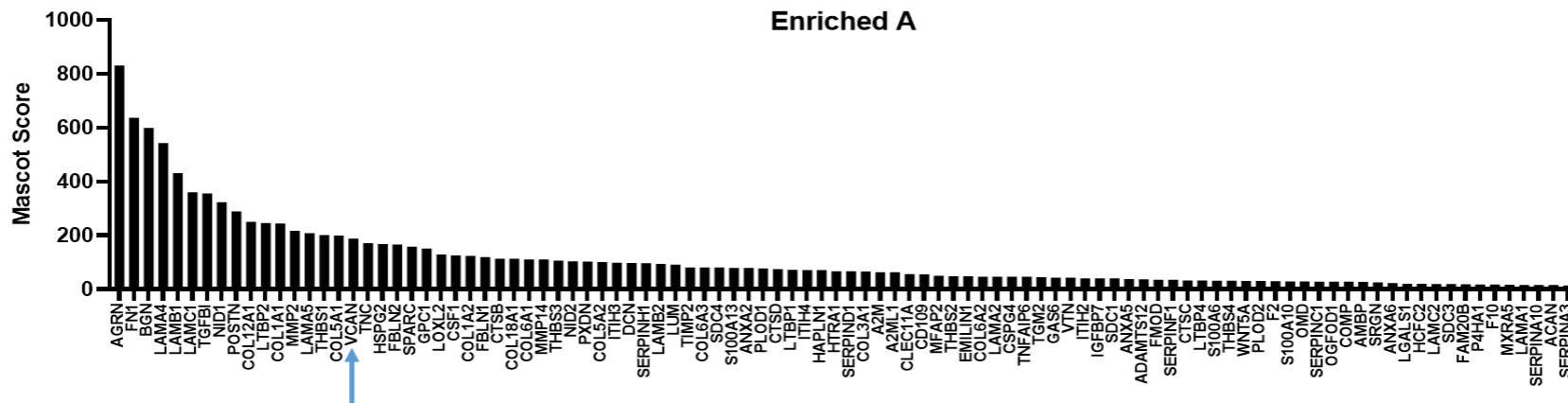
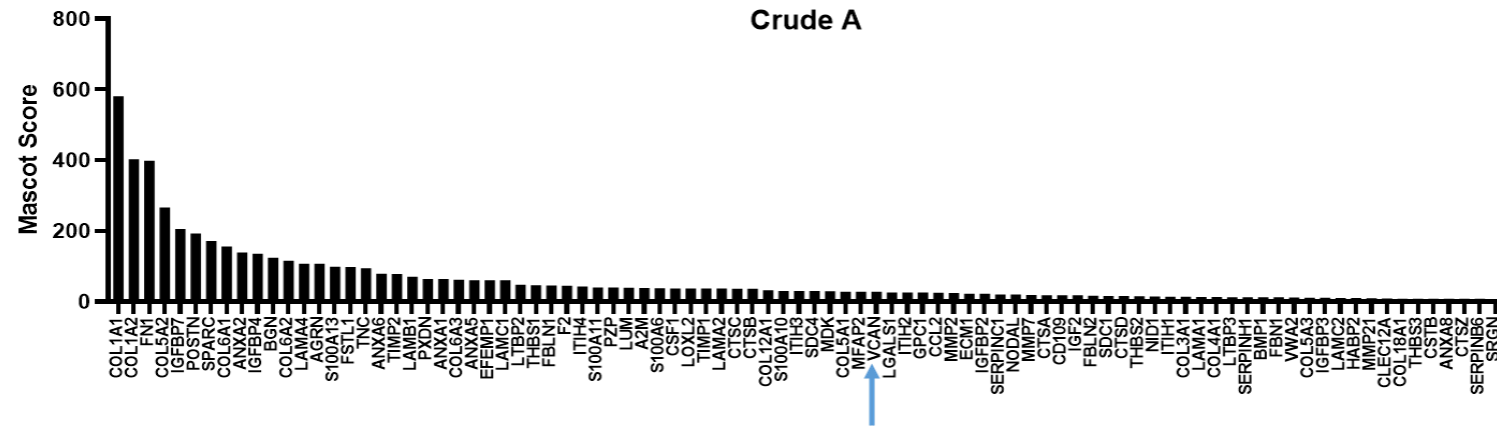
Appendix 3. Plasmid structure of the V1 plasmid.

mappedCaseId	Sample Type	Subtype	TCGA-A2-A04P	Primary Tumor	LAR
TCGA-A1-A0SP	Primary Tumor	BL1	TCGA-A2-A0EQ	Primary Tumor	LAR
TCGA-A2-A04T	Primary Tumor	BL1	TCGA-A2-A1G6	Primary Tumor	LAR
TCGA-A2-A0CM	Primary Tumor	BL1	TCGA-A7-A0CE	Primary Tumor	LAR
TCGA-A2-A0D0	Primary Tumor	BL1	TCGA-AO-A0J2	Primary Tumor	LAR
TCGA-A2-A0D2	Primary Tumor	BL1	TCGA-AO-A1KR	Primary Tumor	LAR
TCGA-A2-A0YM	Primary Tumor	BL1	TCGA-AR-A2LH	Primary Tumor	LAR
TCGA-A2-A3XS	Primary Tumor	BL1	TCGA-B6-A0IE	Primary Tumor	LAR
TCGA-A2-A3XT	Primary Tumor	BL1	TCGA-B6-A0RS	Primary Tumor	LAR
TCGA-A2-A3XX	Primary Tumor	BL1	TCGA-BH-A0B3	Primary Tumor	LAR
TCGA-A8-A07C	Primary Tumor	BL1	TCGA-BH-A0E0	Primary Tumor	LAR
TCGA-A8-A07O	Primary Tumor	BL1	TCGA-BH-A18Q	Primary Tumor	LAR
TCGA-A8-A07R	Primary Tumor	BL1	TCGA-BH-A18V	Primary Tumor	LAR
TCGA-A8-A08R	Primary Tumor	BL1	TCGA-BH-A1EW	Primary Tumor	LAR
TCGA-AC-A6IW	Primary Tumor	BL1	TCGA-BH-A1F6	Primary Tumor	LAR
TCGA-AC-A8OQ	Primary Tumor	BL1	TCGA-BH-A1FC	Primary Tumor	LAR
TCGA-AN-A04D	Primary Tumor	BL1	TCGA-BH-A42U	Primary Tumor	LAR
TCGA-AN-A0AT	Primary Tumor	BL1	TCGA-C8-A26X	Primary Tumor	LAR
TCGA-AO-A0J4	Primary Tumor	BL1	TCGA-C8-A26Y	Primary Tumor	LAR
TCGA-AO-A0J6	Primary Tumor	BL1	TCGA-C8-A3M7	Primary Tumor	LAR
TCGA-AO-A124	Primary Tumor	BL1	TCGA-D8-A143	Primary Tumor	LAR
TCGA-AO-A128	Primary Tumor	BL1	TCGA-D8-A1JF	Primary Tumor	LAR
TCGA-AO-A129	Primary Tumor	BL1	TCGA-D8-A1JG	Primary Tumor	LAR
TCGA-AQ-A04J	Primary Tumor	BL1	TCGA-E2-A1L7	Primary Tumor	LAR
TCGA-AR-A0TS	Primary Tumor	BL1	TCGA-E2-A1LH	Primary Tumor	LAR
TCGA-AR-A0U4	Primary Tumor	BL1	TCGA-A1-A0SK	Primary Tumor	M
TCGA-A2-A0SX	Primary Tumor	BL2	TCGA-A1-A0SO	Primary Tumor	M
TCGA-A7-A26F	Primary Tumor	BL2	TCGA-A2-A04U	Primary Tumor	M
TCGA-A7-A26G	Primary Tumor	BL2	TCGA-A2-A0T0	Primary Tumor	M
TCGA-A7-A26I	Primary Tumor	BL2	TCGA-A2-A0T2	Primary Tumor	M
TCGA-A7-A4SD	Primary Tumor	BL2	TCGA-A2-A3XU	Primary Tumor	M
TCGA-A7-A4SE	Primary Tumor	BL2	TCGA-A2-A3XY	Primary Tumor	M
TCGA-A7-A6VW	Primary Tumor	BL2	TCGA-A7-A0DA	Primary Tumor	M
TCGA-A7-A6VY	Primary Tumor	BL2	TCGA-A7-A5ZV	Primary Tumor	M
TCGA-AC-A2QJ	Primary Tumor	BL2	TCGA-A7-A6VV	Primary Tumor	M
TCGA-AN-A0AL	Primary Tumor	BL2	TCGA-AC-A2BK	Primary Tumor	M
TCGA-AN-A0XU	Primary Tumor	BL2	TCGA-AC-A2QH	Primary Tumor	M
TCGA-AO-A03U	Primary Tumor	BL2	TCGA-AN-A0AR	Primary Tumor	M
TCGA-AR-A1AJ	Primary Tumor	BL2	TCGA-AN-A0FJ	Primary Tumor	M
TCGA-AR-A1AO	Primary Tumor	BL2	TCGA-AN-A0FL	Primary Tumor	M
TCGA-AR-A1AR	Primary Tumor	BL2	TCGA-AN-A0FX	Primary Tumor	M
TCGA-AR-A5QQ	Primary Tumor	BL2	TCGA-AN-A0G0	Primary Tumor	M
TCGA-BH-A0BL	Primary Tumor	BL2	TCGA-AO-A12F	Primary Tumor	M
TCGA-C8-A131	Primary Tumor	BL2	TCGA-AQ-A54N	Primary Tumor	M
TCGA-D8-A13Z	Primary Tumor	BL2	TCGA-AR-A0TP	Primary Tumor	M
TCGA-D8-A1JL	Primary Tumor	BL2	TCGA-AR-A1AH	Primary Tumor	M
TCGA-D8-A27M	Primary Tumor	BL2	TCGA-AR-A1AY	Primary Tumor	M
TCGA-E2-A150	Primary Tumor	BL2	TCGA-AR-A2LR	Primary Tumor	M
TCGA-E2-A159	Primary Tumor	BL2	TCGA-B6-A0I6	Primary Tumor	M
TCGA-E2-A1AZ	Primary Tumor	BL2	TCGA-B6-A0IQ	Primary Tumor	M
TCGA-E2-A1LG	Primary Tumor	BL2			

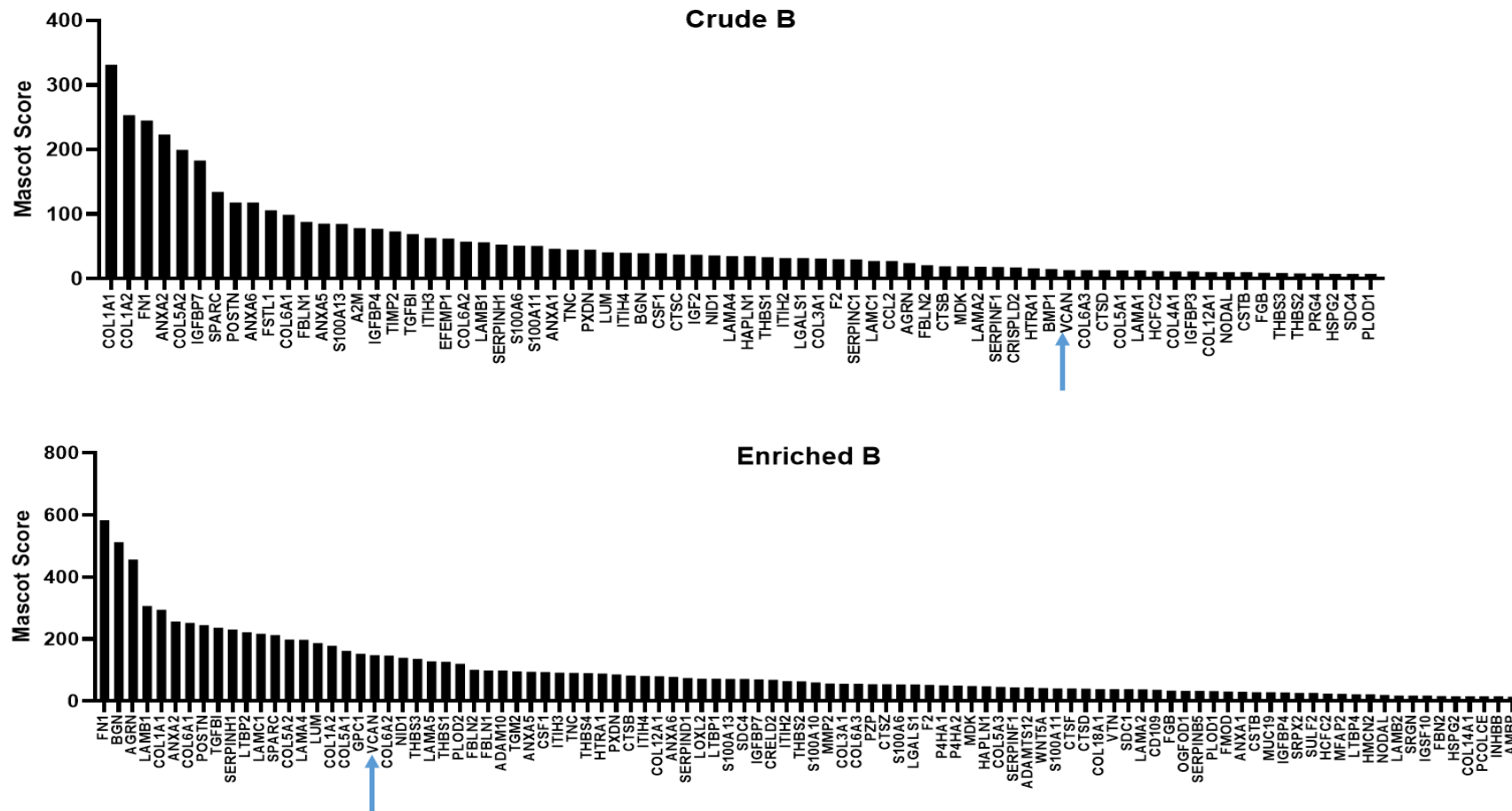
Appendix 4. TNBC TCGA Case IDs with TNBC subtype

mappedCaseId	Sample Type	Subtype
TCGA-A7-A0CE	Primary Tumor	LAR
TCGA-A7-A0CE	Solid Tissue Normal	LAR
TCGA-BH-A0B3	Primary Tumor	LAR
TCGA-BH-A0B3	Solid Tissue Normal	LAR
TCGA-BH-A0E0	Primary Tumor	LAR
TCGA-BH-A0E0	Solid Tissue Normal	LAR
TCGA-BH-A18Q	Primary Tumor	LAR
TCGA-BH-A18Q	Solid Tissue Normal	LAR
TCGA-BH-A18V	Primary Tumor	LAR
TCGA-BH-A18V	Solid Tissue Normal	LAR
TCGA-BH-A1EW	Primary Tumor	LAR
TCGA-BH-A1EW	Solid Tissue Normal	LAR
TCGA-BH-A1F6	Primary Tumor	LAR
TCGA-BH-A1F6	Solid Tissue Normal	LAR
TCGA-BH-A1FC	Primary Tumor	LAR
TCGA-BH-A1FC	Solid Tissue Normal	LAR
TCGA-E2-A1L7	Primary Tumor	LAR
TCGA-E2-A1L7	Solid Tissue Normal	LAR
TCGA-E2-A1LH	Primary Tumor	LAR
TCGA-E2-A1LH	Solid Tissue Normal	LAR

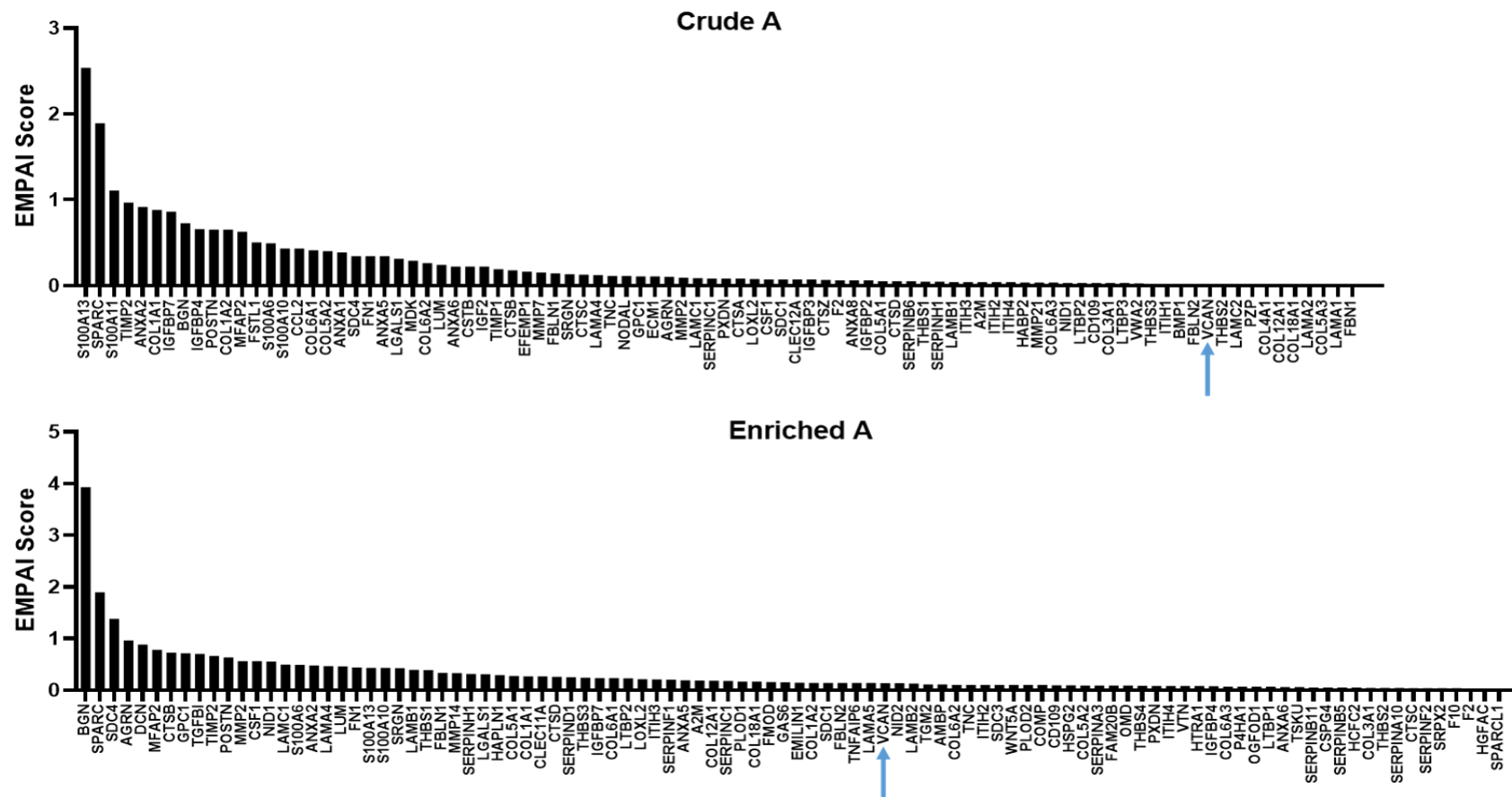
Appendix 5. TCGA Case IDs for Normal and tumour matched samples.



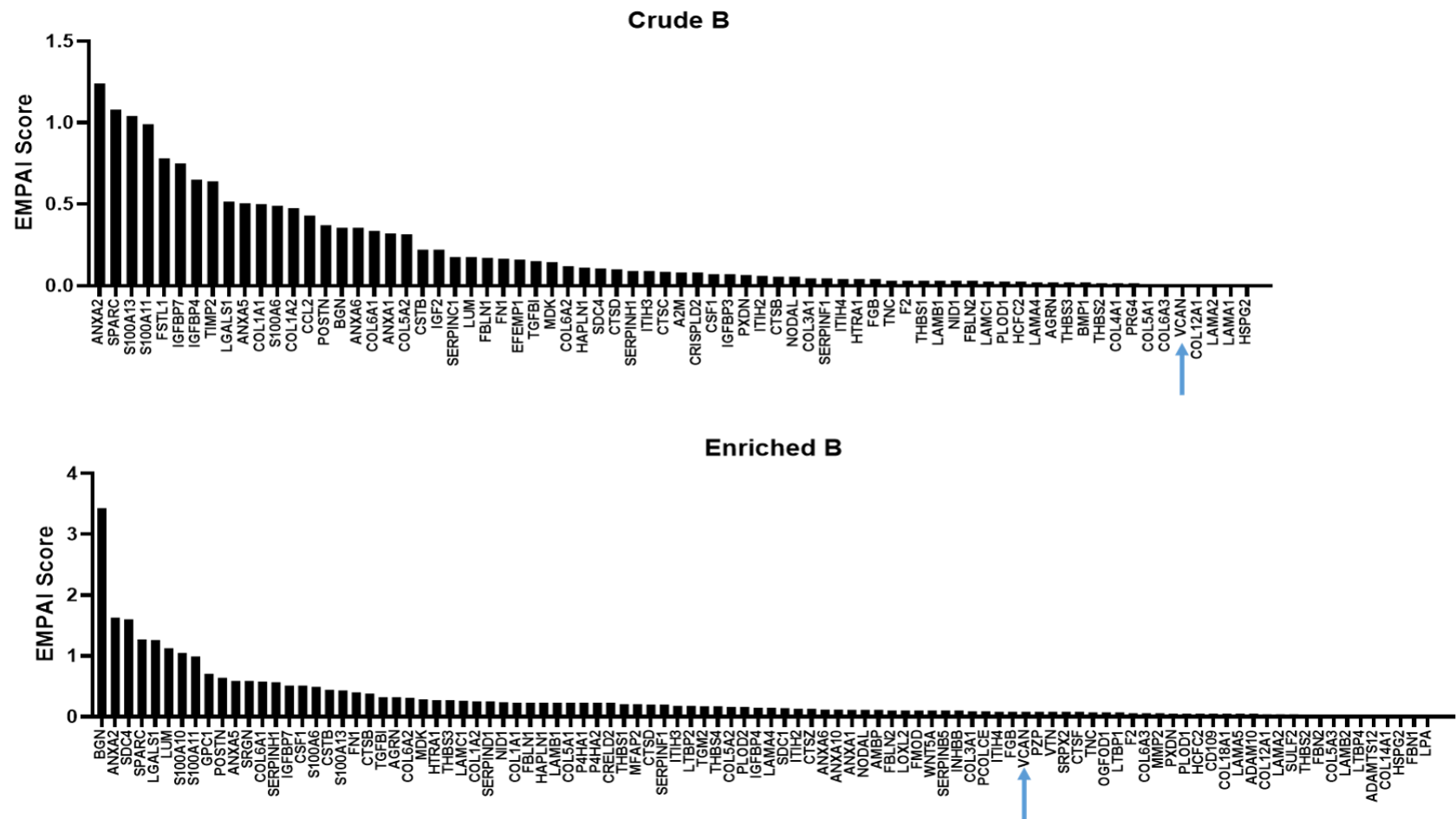
Appendix 6. Mascot scores of the crude and enriched samples from HCC38 sample A. Arrows indicating VCAN.



Appendix 7. Mascot scores of the crude and enriched samples from HCC38 sample B. Arrows indicating VCAN.



Appendix 8. emPAI scores of the crude and enriched samples from HCC38 sample A. Arrows indicating VCAN .



Appendix 9. emPAI scores of the crude and enriched samples from HCC38 sample. Arrows indicating VCAN.