



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Investigating the role of the tumour microenvironment in Invasive Lobular Carcinoma

Esme Bullock



Doctorate of Philosophy
University of Edinburgh
2023

Declaration

This thesis has been entirely composed by myself between September 2019 and June 2023 under the supervision of Professor Valerie Brunton at the Edinburgh Cancer Research Centre. This work has not been submitted for any other degrees or qualifications. The work presented is entirely my own, except where acknowledgement has been made.

Parts of this work have been included in the following published papers:

Gómez-Cuadrado L, **Bullock E**, Mabruk Z, et al. Characterisation of the Stromal Microenvironment in Lobular Breast Cancer. *Cancers (Basel)*. 2022;14(4):904. doi:10.3390/cancers14040904

Houthuijzen J, de Bruijn R, van der Burg E, Drenth AP, Wientjens E, Filipovic T, **Bullock E**, Brambillasca CS, Pulver EM, Nieuwland M, de Rink I, van Diepen F, Klarenbeek S, Kerkhoven R, Brunton VG, Scheele C L.G.J., Boelens MC, Jonkers J; CD26-negative and CD26-positive tissue-resident fibroblasts contribute to functionally distinct CAF subpopulations in breast cancer. *Nat. Commun.* **14**, 183 (2023). doi: 10.1038/s41467-023-35793-w

Esme Bullock

June 2023

Acknowledgments

First of all, I'd like to say a massive thank you to Val Brunton for being a fantastic supervisor. Thank you for your guidance and for finding so many amazing opportunities for me, for supporting my ideas and lifting me out of rabbit holes I'd dug for myself and gently pointing me in the right direction. I've thoroughly enjoyed my PhD (despite the pandemic in the middle of it) and a lot of that is due to you! I couldn't have asked for a better, more supportive supervisor. I'd also like to thank my thesis committee Professor Liz Patton and Dr Arran Turnbull and my second supervisor Dr Luke Bolter for their guidance and advice.

I'd also like to thank my second supervisor Dr Andy Sims. Andy was an incredibly enthusiastic and supportive researcher. He was the first person that ever made bioinformatics seem interesting or even vaguely understandable to me. At the beginning of my PhD, there was nothing I liked less than R, but bioinformatics has now become one of my favourite parts of this project – that is entirely down to Andy. He contributed an enormous amount to science and was a lovely person. I am so grateful I had the opportunity to work with him.

I am grateful to Cancer Research UK for funding my PhD and for the Royal Society of Edinburgh for funding my Early Career Fellowship research at the University of Turin. Thank you to Professor Valeria Poli for hosting me in her lab during this fellowship and for suggesting the initial Zebrafish experiments. Thank you to the whole Poli group – Lidia, Daniele, Somayeh, Andrea, Michele, Shamanth, Natalie, Lorenzo – for their support, helping me practice Italian and the lovely aperitivo trips.

Much of this work has been made possible due to incredible support and collaboration across the University of Edinburgh. I'd like to thank Kenneth MacLeod for his help with the RPPA and FPPA experiments and Alison Munro at the HTPU. Thanks to the Wellcome Trust Clinical Research Facility for carrying out library preparation and sequencing and to Henry Beetham for assistance in analysis. Thank you to John Dawson for his help with the incucyte and CAF-derived matrix experiments. Thanks to Martin Lee, Amy Davies and the whole AIR team for help with microscopy and analysis. A massive thank you to the Zebrafish facility and Professor Liz Patton's group for help with the zebrafish work. In particular I'd like to say thank you to Adelaide Young

for her great advice, collaboration on the zebrafish experiments and the number of weekends she came in to set up experiments – you're amazing! Thank you to the wonderful Jayne Culley for getting me set up and making me feel so welcome when I first started. Also, thank you to Jan Irvine for sorting out all our meetings and solving all my admin and conference problems.

A massive thank you to all the lobular group – Laura Gómez-Cuadrado, Zeanap Mabruk, Aleksandra Rozyczko and Sana Shabbir – and previous members of the lab Katy Teo and Kenneth MacLeod for collecting the patient-derived CAFs. Thank you to Laura in particular for helping me to find my feet, for her endless knowledge, ideas, detailed labelling, kindness and quick response to panicked WhatsApp messages! I am particularly grateful to Aleks Rozczyko, a brilliant Masters student and now research assistant. I thoroughly enjoyed teaching her and working with her and I am very thankful for the hours she spent on the microscope and for all the ways she helped me out.

I'm very grateful to have had the opportunity to meet and talk to so many amazing lobular patient advocates from LBCUK. In particular I'd like to thank Clare Turner and Helen Coulthard for all their incredible work for lobular patients and for taking time to talk to me about what they want from lobular research and inspiring me. I'd like to thank all the Lobular breast cancer patients and their families for kindly providing samples that made this research possible.

An enormous thank you, for their help with experiments, finding things, life advice, lovely chats, technical support and generally making my PhD fun and the lab a lovely place to be, to all past and present members of Level 3 and the Frame/Brunton group: Dr Giovana Carrasco Gonzalez, Dr Emily Webb, Dr Alex Loftus, Dr Roza Masalmeh, Dr Martin Lee, Dr Fonso Bolado Carrancio, Zeanap Mabruk, Aleks Rozyczko, Dr Laura Gomez-Cuadrado, Dr Jayne Culley, Sana Shabbir, Marianna Romano, Molly Danks, Annabel Black, Dr Piotr Manasterski, Dr Henry Beetham, Morwenna Muir, Craig Steven, Ander Maguregui Villalobos, Dr John Dawson, Dr Robb Hollis, Lorna Stillie, Rosie Matthews, Nathalie Feeley, Kitty Sherwood, Chloe Henry, Athanasia Yiapanas, Dr Alain Kemp, Dr Joanne Simpson, Najib Abualetam and Dr Agata Makar.

To the Lunch At Twelve gang – Beth Bartlett, Brad Harris, Lucy Scott, Cavan Bennet-Ness, Heather Love, Fraser McPhie, Ross Hannah, Nathan Constantine-Cooke, Nathalie Feeley, Matilda Bui, Mia Raab, Nabid Bhuiyan, Sarah Taylor and the lovely Chloe Brook. I love you all, thank you for the much needed breaks, chats, laughs, memes, karaoke, pub trips, sardines and endless fun. I couldn't have asked for a lovelier or kinder group of pals to get me through a PhD and a pandemic. You're all wonderful people and brilliant scientists and I'm endlessly grateful to have you as friends.

To my Arden Street pals – Beth, Bailey and Cob: thanks for a wonderful and chaotic home for our first 13 months in Edinburgh under Audrey's loving gaze. Here's to surviving hallway cricket and a pandemic together and to the Earl! A particular thank you to Brilliant Beef Bartlett, one of the smartest and funniest people I know, for being a great friend through our undergrad, Masters' and PhDs together. I'm immensely grateful for my incredibly kind, funny, lovely friends Issie Campbell, Susy Harrison and Eliette Harris for being the best pals I could ask for for the last 14 fantastic years. Thanks to Darran and Stacie Brookes for all the lovely trips to Sennen and incredible roast dinners over the years.

An immense, huge, massive thank you to my fantastic family. To my sisters Josie and Poppy – thanks for all the poorly timed video calls that made me laugh and for just being brilliant and funny. To my wonderful mum and dad – thank you for listening to me waffle on about my project, for reading lay summaries, for checking in with me and for being all around wonderful people. Love you!

Lastly, thank you to Bailey Brookes for being so funny, kind and supportive and making me happy every day. Thank you for listening to me, for looking after me and putting up with so many biologists! I'm sorry that you know what a western blot is now! I love you, thank you for being you.

Contents

Declaration.....	ii
Acknowledgments.....	iii
Contents.....	vi
Abstract.....	xiv
Lay Summary.....	xvi
Abbreviations	xviii
List of Figures	xxii
List of Tables	xxiii
Chapter 1 Introduction	1
Chapter 1.1 Invasive Lobular Carcinoma.....	1
1.1.1 Clinical features of ILC.....	3
1.1.1.1 ILC and IDC cohort characteristics and risk factors	3
1.1.1.2 Histopathological and Molecular Subtypes	3
1.1.1.3 Diagnosis and detection.....	4
1.1.1.4 Metastasis.....	4
1.1.1.5 Outcomes.....	5
1.1.2 Genomic and molecular characteristics of ILC	6
1.1.2.1 Mutations	6
1.1.2.2 Phenotypic consequences of E-cadherin loss in ILC	6
1.1.2.3 IGF-1 and PI3K signalling	7
1.1.2.4 Response to Estrogen and Endocrine-targeted therapies.....	7
1.1.2.5 Origin of ILC tumours	8
1.1.3 Therapies.....	9
1.1.3.1 Endocrine targeted therapies	9
1.1.3.2 CDK4/6 inhibitors	10
1.1.3.3 Chemotherapy	11

1.1.3.4	Immunotherapy	11
1.1.3.5	ILC specific clinical trials	11
1.1.4	Models of ILC	12
1.1.4.1	In vitro models.....	12
1.1.4.2	Genetically engineered mouse models	13
1.1.4.3	Xenografts.....	13
1.1.4.4	Publicly available datasets – TCGA and METABRIC.....	14
Chapter 1.2	The Tumour Microenvironment	16
1.2.1	Cancer Associated Fibroblasts in Breast Cancer	17
1.2.1.1	Origins of breast CAFs.....	19
1.2.1.2	CAF Heterogeneity in breast cancer	19
1.2.1.3	CAF functions in breast cancer	21
1.2.1.4	Cancer associated fibroblasts in ILC.....	22
Chapter 1.3	Interleukin-6 (IL-6) and Signal Transducer and Activator of Transcription 3 (STAT3) signalling in Breast Cancer	24
1.3.1	IL-6/STAT3 pathway	24
1.3.2	IL-6/STAT3 in breast cancer.....	26
1.3.2.1	Breast cancer progression	26
1.3.2.2	Influence on the TME	27
1.3.2.3	Breast cancer therapy resistance.....	28
1.3.3	IL-6/STAT3 targeted therapies and outcomes in breast cancer.....	30
Chapter 1.4	Hypothesis and aims.....	32
Chapter 2	Materials and Methods.....	34
Chapter 2.1	<i>In vitro</i> methods.....	34
2.1.1	Cell culture.....	34
2.1.1.1	CAF isolation.....	35
2.1.1.2	Collagen coating plates.....	36

2.1.2	Conditioned media.....	36
2.1.3	Stimulation of cell lines	36
2.1.4	siSTAT3.....	37
2.1.5	Western blotting.....	37
2.1.6	Immunofluorescence	39
2.1.7	RT-qPCR.....	42
2.1.8	Ethynyl-2'-deoxyuridine (EdU) proliferation assay	43
2.1.9	ELISA	44
2.1.10	CAF derived matrices	44
2.1.11	Migration assay	46
2.1.12	ALDEFLUOR assay.....	46
2.1.13	FPPA	46
2.1.14	RPPA.....	47
Chapter 2.2	RNAseq experiments and analysis	49
2.2.1	Primary ILC and IDC CAFs sample preparation	49
2.2.2	CAF CM stimulation of SUM44PE cells sample preparation	49
2.2.3	Lexogen 3`-mRNA QuantSeq.....	49
2.2.4	Gene set enrichment analysis and ssGSEA.....	50
2.2.5	IL6-dependent gene signature and score	50
2.2.6	Network analysis	51
2.2.7	Heatmaps	51
2.2.8	Data availability	51
Chapter 2.3	Publicly available datasets	52
Chapter 2.4	Zebrafish experiments.....	53
2.4.1	Zebrafish husbandry	53
2.4.2	Zebrafish embryo xenograft assay	53
2.4.2.1	SUM44PE cell labelling.....	53
2.4.2.2	Embryo injections.....	53

2.4.2.3	4 hours post injection screen	54
2.4.2.4	1 dpi – tumour size screen	54
2.4.2.5	2 dpi – screen for embryos with disseminated cells	54
2.4.3	Whole mount embryo processing and microscopy.	55
2.4.4	EdU staining of embryos.....	55
2.4.5	Whole mount embryo image analysis	56
2.4.5.1	Disseminated cell quantification	56
2.4.5.2	EdU quantification	56
2.4.5.3	Angiogenesis quantification	56
Chapter 2.5	General	58
2.5.1	Statistics	58
2.5.2	Schematics and Diagrams	58
Chapter 3	Characterising Invasive Lobular and Invasive Ductal Carcinoma	
	Patient-derived Cancer Associated Fibroblasts	59
Chapter 3.1	Introduction	60
Chapter 3.2	Results	62
3.2.1	RNA-seq analysis of 11 primary ILC and 5 primary IDC CAF lines	62
3.2.2	Differential gene expression in ILC and IDC CAFs	64
3.2.3	Estrogen receptor expression in ILC and IDC CAFs	68
3.2.4	ILC and IDC CAFs do not differ in expression of CAF subtype signatures	69
3.2.5	Integrating ILC and IDC CAF RNAseq data with laser-capture microdissection datasets	72
3.2.6	ILC and IDC CAF secretome	75
Chapter 3.3	Discussion.....	80
3.3.1	The prognostic role of the stroma in ILC.....	80
3.3.2	ILC stroma and CAF enriched genes	81
3.3.3	<i>ESR1</i> and <i>GPER1</i> expression in CAFs	83

3.3.4	CAF secreted proteins	84
3.3.5	CAF subtypes	86
Chapter 3.4	Conclusions	88
Chapter 4	Paracrine IL-6/STAT3 signalling in Invasive Lobular Carcinoma	89
Chapter 4.1	Introduction	90
Chapter 4.2	Results	92
4.2.1	CAF conditioned media induces activation of a number of pro-tumorigenic pathways in ILC tumour cells	92
4.2.2	Identifying CAF-expressed STAT3 activating factors	97
4.2.3	IL-6 in CAF conditioned media drives STAT3 activation	99
4.2.4	IL-6/STAT3 signalling is enriched in ILC tumours and correlates with stromal content	102
4.2.5	IL-6 drives gene expression changes in ILC tumour cells.....	104
4.2.6	An IL-6 gene signature is enriched in ILC tumours.....	112
Chapter 4.3	Discussion.....	117
4.3.1	ILC CAF cell lines used for experiments.....	117
4.3.2	CAF CM driven signalling pathway protein changes	118
4.3.2.1	IGF-1/PI3K pathway activation.....	118
4.3.2.2	STAT3 activation.....	118
4.3.3	IL-6 dependent gene expression	120
4.3.3.1	IL-6 driven suppression of pro-proliferative gene sets.....	121
4.3.3.2	IL-6 regulation of metabolism related genes	123
4.3.3.3	IL-6 and therapy resistance.....	123
4.3.3.4	Epithelial to mesenchymal transition and migration	125
4.3.3.5	Inflammation and immune signalling	126
4.3.3.6	Angiogenesis	128
Chapter 4.4	Conclusions	130

Chapter 5	Investigating the relationship between Estrogen Receptor signalling and IL-6/STAT3 in Invasive Lobular Carcinoma	131
Chapter 5.1	Introduction	131
Chapter 5.2	Results	133
5.2.1	IL-6 stimulation of SUM44PE cells leads to downregulation of <i>ESR1</i> and ER α signalling.....	133
5.2.2	IL-6 and <i>ESR1</i> negatively correlate with each other in ER+ breast cancer tumours and cell lines	135
5.2.3	CAF CM suppresses estrogen signalling and upregulates endocrine therapy resistance signatures.....	137
5.2.4	IL-6 does not drive ALDH+ cancer stem cell formation in SUM44PE cells	142
Chapter 5.3	Discussion.....	144
5.3.1	IL-6/STAT3 and ER α signalling in breast cancer	144
5.3.2	CAF CM induced endocrine therapy resistance genes.....	146
5.3.2.1	IL-6 driven upregulation of letrozole resistance related genes	146
5.3.2.2	CAF CM driven SUM44PE Tamoxifen resistance genes	147
5.3.2.3	CDK4/6 inhibitor resistance.....	148
5.3.3	Breast cancer stem cells.....	149
Chapter 5.4	Conclusions	151
Chapter 6	Functional consequences of IL-6/STAT3 signalling in Invasive Lobular Carcinoma	152
Chapter 6.1	Introduction	153
Chapter 6.2	Results	155
6.2.1	CAF conditioned media drives ILC cell proliferation	155
6.2.2	IL-6 drives epithelial-to-mesenchymal transition gene signatures and morphological changes in SUM44PE cells	158
6.2.3	IL-6 drives random 2D migration of SUM44PE cells.....	164

6.2.4	IL-6 drives SUM44PE cell dissemination in a Zebrafish embryo assay	166
6.2.5	IL-6 alters vasculature in the ZFEX model.....	175
Chapter 6.3	Discussion.....	178
6.3.1	Proliferation	178
6.3.2	EMT and migration <i>in vitro</i>	180
6.3.3	Zebrafish xenograft model	182
6.3.3.1	SUM44PE cell dissemination in ZFEX model	182
6.3.3.2	SUM44PE driven angiogenesis in ZFEX model	184
Chapter 6.4	Conclusions	186
Chapter 7	Discussion and Future Work.....	187
Chapter 7.1	ILC Cancer Associated Fibroblasts	187
Chapter 7.2	IL-6/STAT3 pathway in ILC	188
7.2.1	IL-6/STAT3 impact on breast cancer prognosis.....	188
7.2.2	IL-6/STAT3 pathway targeted therapy in breast cancer	189
7.2.3	STAT3 activity in CAFs.....	190
7.2.4	IL-6 driven migration	191
Chapter 7.3	The IGF-1/PI3K pathway.....	192
Chapter 7.4	Future <i>in vivo</i> work.....	193
7.4.1	Intraductal co-injection studies	193
7.4.2	ILC GEMMs	194
Chapter 7.5	The immune microenvironment of ILC	194
Chapter 7.6	Conclusions	197
References		199
Appendix 1	Image analysis macros	226
Appendix 2	FPPA and RPPA antibody lists	228
Appendix 2.1	FPPA antibodies.....	228

Appendix 2.2 RPPA antibodies	230
Appendix 3 Gene sets used for Gene Set Enrichment Analysis	234
Appendix 4 IL-6 dependent gene signature literature review	240
Appendices References.....	249

Abstract

Invasive lobular carcinoma (ILC) is the second most common histological subtype of breast cancer, after invasive ductal carcinoma (IDC), and is defined by loss of the adherens junction protein E-cadherin. ILC tumours are very slow growing and uniquely metastasise to the gastrointestinal tract, peritoneum and ovaries. Although prognosis for ILC and IDC patients initially is good, with roughly 90% of ILC tumours being ER+ and so ideal candidates for endocrine therapies, ILC patients have a poorer long term outcome in multivariate analyses and a poorer response to neoadjuvant chemotherapy than IDC patients. Despite the myriad molecular, histological and clinical differences that distinguish ILC tumours, ILC and IDC patients are treated in the same way in the clinic. Understanding the differences between these two subtypes and identifying therapeutic targets that are enriched in ILC tumours is necessary to improve patient care and outcome for this under-studied but common subtype of breast cancer.

ILC tumours contain a particularly high stromal content and in particular, have a very high density of cancer-associated fibroblasts (CAFs). Analysis of primary ILC and IDC patient-derived CAFs by RNA-sequencing identified 153 differentially expressed genes between ILC and IDC CAFs and a significant enrichment for the pro-tumorigenic mTORC1 pathway in ILC CAFs. A forward-phase protein array (FPPA) identified high expression of insulin-like growth factor-1 (IGF-1) and IGF-1 pathway proteins by CAFs, a pro-proliferative pathway that is enriched in ILC.

To determine the consequences of paracrine signalling from CAFs to ILC tumour cells, the ILC cell line SUM44PE was stimulated with conditioned media collected from the primary ILC CAFs. Reverse-phase protein array analysis identified a significant activation of Mitogen-Activated Protein Kinase pathway proteins and a significant and sustained phosphorylation of Signal Transducer and Activator of Transcription 3 (STAT3) at Tyr705. This phosphorylation leads to STAT3 translocation to the nucleus where it acts as a transcription factor.

IL-6 was identified as the CAF-secreted factor driving STAT3 activation and RNA-sequencing was used to generate a 42-gene IL-6 dependent gene signature. In the TCGA dataset, this signature, as well as IL-6 and pSTAT3 are all significantly more highly expressed in ER+ ILC tumours compared to ER+ IDC tumours and their

expression correlates with stromal content, supporting the paracrine nature of this signalling pathway in ILC.

CAF conditioned media stimulation of SUM44PE cells downregulated expression of the Estrogen receptor (ER α) at the RNA and protein level and suppressed expression of generic and ILC-specific ER α -dependent gene sets as well as upregulating endocrine therapy resistance signatures, dependent on IL-6. In ER⁺ ILC and IDC tumours in the TCGA dataset, there is a significant negative correlation between *ESR1* expression and expression of *IL6* and the IL6 dependent gene signature. This suggests that paracrine IL-6 signalling from CAFs suppresses Estrogen signalling and may contribute to endocrine-targeted therapy resistance in ILC.

IL-6 stimulation also led to increased expression of EMT gene sets and caused SUM44PE cells to become more elongated and mesenchymal-like when adhered to CAF-derived extracellular matrices, dependent on STAT3 expression. IL-6 stimulation promoted increased cell motility in 2D and in a zebrafish embryo xenograft model of metastasis, IL-6 stimulation significantly increased SUM44PE cell dissemination throughout the embryos and promoted angiogenesis, suggesting that IL-6/STAT3 signalling may contribute to the metastatic spread of ILC tumours.

Overall, this work identifies ILC-enriched CAF expressed genes and secreted proteins, as well as highlighting the IL-6/STAT3 pathway as a potential therapeutic target in ILC.

Lay Summary

There are multiple different types of breast cancer and 10-15% of people diagnosed with breast cancer are diagnosed with Lobular Breast Cancer. Lobular breast cancer is the second most common type of breast cancer after ductal breast cancer. Lobular tumours grow in a single-file line instead of forming a lump, like most breast cancers, because the cancer cells lose expression of E-cadherin, a protein that helps cells stick together. Lobular breast cancer is very different from ductal breast cancer but is treated in the exact same way even though lobular tumours are driven by different mutations, grow differently and respond to treatments differently.

Because of the single-file growth of lobular tumours, lobular cancer cells come into contact with lots of non-cancerous 'healthy' cells in the breast, more so than other breast cancer types. The most abundant type of these 'healthy' cells in lobular tumours are cancer-associated fibroblasts or CAFs. These CAFs support the growth, invasion and spread of cancer cells throughout the body by secreting proteins that change cancer cell behaviour and by also laying down the mesh of extracellular proteins that the cancer cells attach to and use to move along. This work focuses on understanding the differences between CAFs isolated from lobular and ductal patients, as well as the effect CAFs have on the behaviour of lobular cancer cells, in order to identify potential therapeutic targets for this understudied disease.

In this work, I've investigated the differences between CAFs isolated from lobular patients and those isolated from ductal patients. By comparing the differences in the genes that are switched on in CAFs, we identified 153 genes that were different between the lobular and ductal CAFs, and found that lobular CAFs have higher activity of the mTORC1 protein that promotes cancer growth.

One of the main methods of communication between CAFs and cancer cells is through CAF-secreted proteins, which cancer cells recognise and interact with, changing which genes the cancer cells switch on and how the cells behave. To understand how CAF-secreted proteins impact lobular cancer cell behaviour, lobular cancer cells were stimulated with factors secreted by the CAFs. This activated a protein called STAT3 in the cancer cells, which can switch on a wide range of genes. IL-6 was identified as the CAF-secreted protein driving these changes. IL-6/STAT3 signalling led to

activation of genes involved in the formation of new blood vessels and cell migration, as well as decreasing levels of genes involved in cell growth and estrogen signalling. Levels of IL-6, activated STAT3 and the genes they switch on are higher in lobular tumours than ductal tumours, suggesting targeting this pathway might specifically help lobular patients.

IL-6 stimulation of lobular cancer cells caused them to elongate and move faster in a dish, features that cause cancer cells to spread throughout patients' bodies. When lobular cancer cells were treated with IL-6 and injected into zebrafish embryos, the IL-6 made them more likely to migrate away from the injection site and more cells were found at distant regions of the zebrafish than when cancer cells were not treated with IL-6. This acts as a model for how the cells would behave in a patient, with IL-6 secreted by CAFs making the cancer cells move faster and spread around the body, forming new tumours at distant sites known as metastases, which is the main cause of cancer death.

Overall, this work adds to our understanding of the under-studied lobular subtype of breast cancer and has identified a potential target that could be used to slow the spread of lobular breast cancer.

Abbreviations

ALDH	Aldehyde dehydrogenase
AR	Aspect ratio
BCSS	Breast cancer specific survival
BH	Benjamini-Hochberg correction
BKY	Benjamini, Krieger and Yekutieli two-stage step-up method of correction
BSA	Bovine serum albumin
CAF	Cancer associated fibroblast
CDH1	E-cadherin gene
CDK4/6i	Cyclin dependent kinase 4 and 6 inhibitor
CDM	CAF-derived matrix
CIRC	Zebrafish embryos with tumour cells injected into circulation
CK	Cytokeratin
CM	Conditioned media
CPE	Consensus measurement of purity estimate
CPM	Counts per million
CSC	cancer stem cell
DAPI	4',6-diamidino-2-phenylindole
DEGs	Differentially expressed genes
DFS	Disease free survival
DMEM	Dulbecco's modified Eagle's media
DMSO	Dimethyl sulfoxide
DoC	Duct of Cuvier
dpf	days post fertilisation
dpi	days post injection
E2	17 β -estradiol
ECL	Enhanced chemiluminescence
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EdU	5-ethynyl-2'-deoxyuridine

ELISA	enzyme-linked immunosorbent assay
EMT	epithelial to mesenchymal transition
ERK	Extracellular signal-related kinase
ER α	Estrogen Receptor
ESR1	Estrogen receptor gene
ET	Endocrine-targeted therapy
FBS	Fetal bovine serum
FDR	False discovery rate
FGF	Fibroblast growth factor
FGFR1	Fibroblast growth factor receptor 1
FGFRi	Infgratinib, Fibroblast growth factor receptor 1 inhibitor
FoV	Fields of view
FPPA	forward-phase protein array
GEMM	Genetically engineered mouse model
GFP	Green fluorescent protein
GO	Gene ontology
GO BP	Gene ontology biological processes
gp130/IL6ST	glycoprotein 130/IL6 co-receptor
GSEA	Gene set enrichment analysis
HER2	Human epidermal growth factor receptor 2
iCAF	Inflammatory cancer associated fibroblast
ICI	Immune checkpoint inhibitor
IDC	Invasive Ductal Carcinoma
IF	Immunofluorescence
IGF-1	Insulin-like growth factor
IGF1R	Insulin-like growth factor receptor
IGFBP	Insulin-like growth factor binding protein
IL6	Interleukin-6
IL6GS	Interleukin-6 dependent gene signature
IL6R	Interleukin-6 receptor
ILC	Invasive Lobular Carcinoma
JAK	Janus Kinase

KEP	<i>K14Cre;E-cadherin^{fl/fl};Tp53^{fl/fl}</i> model of ILC
LCM	Laser capture microdissection
LTED	Long-term estrogen-deprived
MAPK	Mitogen activated protein kinase
MDSC	myeloid derived suppressor cells
MET	Zebrafish embryos with disseminated tumour cells
METABRIC	Molecular Taxonomy of Breast Cancer International Consortium
MMP	Matrix metalloproteinase
MSC	Mesenchymal stem cell
myCAF	Myofibroblastic cancer associated fibroblast
NA	not available
NES	Normalised enrichment score
NF	normal fibroblast
ns	not significant
OS	Overall survival
P/S	Penicillin/Streptomycin
PAM50	intrinsic molecular subtypes of breast cancer
PAPPA	Pregnancy-associated plasma protein A/Pappalysin-1
pCR	Pathologic complete response
PDAC	Pancreatic Ductal Adenocarcinoma
PDGFR	platelet-derived growth factor receptors
PDX	Patient derived xenograft
PFS	Progression free survival
PI3K	phosphatidylinositol-3-kinase
PR	Progesterone receptor
pSTAT3	STAT3 phosphorylated at Tyr705/activated STAT3
PTEN	Phosphatase and tensin homolog
PVS	perivitelline space
pY705	phosphorylated Tyrosine 705
RPPA	reverse-phase protein array
RSE	Royal Society of Edinburgh
RT-qPCR	Quantitative reverse transcriptase polymerase chain reaction

scRNAseq	Single cell RNAseq
SIV	Subintestinal vein
ssGSEA	Single sample gene set enrichment analysis
STAT1	Signal Transducer and activator of transcription 1
STAT3	Signal Transducer and activator of transcription 3
TBS	Tris-buffered saline
TBST	Tris-buffered saline + Tween
TBS-Tr	Tris-buffered saline + Triton X100
TCGA	The Cancer Genome Atlas
TE	Tumour epithelium
Tg(Fli:eGFP)	Transparent zebrafish lineage with GFP-labelled vasculature
TGF β	Transforming growth factor beta
TIMP	Tissue inhibitor of metalloproteinases
TMA	Tumour microarray
TME	Tumour microenvironment
TNBC	Triple negative breast cancer
TN-ILC	Triple negative invasive lobular carcinoma
TS	Tumour stroma
WB	Western blot
WEPTn	<i>WapCre;E-cadherin^{ff};Pten^{ff}</i> model of ILC
WT	Wild type
ZFEX	Zebrafish embryo xenograft assay
α SMA	α -smooth muscle actin

List of Figures

Figure 1-1. Key differences between ILC and IDC tumours.	2
Figure 1-2. Structure of healthy mammary tissue.....	9
Figure 1-3. The tumour microenvironment of IDC (left, blue cells) and ILC (right, purple cells) tumours.....	17
Figure 1-4. Summary of CAF origins, heterogeneity and functions in breast cancer.	18
Figure 1-5. STAT3 activation.....	25
Figure 1-6. IL-6/STAT3 pathway targeted drugs and monoclonal antibodies.....	31
Figure 1-7. Summary of the aims of this thesis.	33
Figure 3-1. ILC and IDC CAFs are fairly similar at the transcriptional level.	63
Figure 3-2. Differential gene expression in ILC and IDC CAFs.	65
Figure 3-3. ESR1 and GPER1 expression in ILC and IDC CAFs.....	68
Figure 3-4. CAF subtype signatures in ILC and IDC CAFs.	70
Figure 3-5. Expression of ILC stroma enriched genes identified by LCM analysis in the ILC and IDC CAF RNAseq dataset.	72
Figure 3-6. Analysis of genes encoding secreted proteins in the CAF RNAseq dataset.	75
Figure 3-7. ILC and IDC CAF secretome analysed by FPPA.....	77
Figure 4-1 . Heatmap of proteins and phospho-proteins significantly altered by ED38 ILC CAF CM stimulation of SUM44PE cells.	92
Figure 4-2. CAF conditioned media drives MAPK and STAT3 pathway activation...	93
Figure 4-3. pSTAT3 and MAPK pathway activation is enriched in ER+ ILC tumours.	95
Figure 4-4. mRNA expression of STAT3 activating factors in ILC CAFs.....	97
Table 4-1. Expression of STAT3 activating factors in ILC CAFs and in bulk tumours.	97
Figure 4-5. IL-6 drives STAT3 activation.	99
Figure 4-6. IL-6 expression in primary ILC and IDC CAFs.	100
Figure 4-7. IL-6/STAT3 signalling is enriched in ILC tumours and correlates with stromal content.....	102

Figure 4-8. CAF secreted IL-6 drives gene expression changes in ILC tumour cells	104
Figure 4-9. IL-6 in CAF CM drives inflammatory and migratory gene set expression.	106
Figure 4-10. Conditioned media treated cells are enriched for immune signalling, migration and angiogenesis gene sets.	109
Figure 4-12. An IL-6 dependent gene signature is enriched in ILC tumours in TCGA dataset.	112
Figure 4-13. Schematic of the paracrine IL-6/STAT3 pathway in ILC.	128
Figure 5-1. IL-6 stimulation of ILC cell lines leads to downregulation of the Estrogen Receptor at the RNA and protein level.	132
Figure 5-2. IL6 and the IL6 gene signature negatively correlate with ESR1 and FOXA1 in ER+ tumours.	134
Figure 5-3. CAF CM suppresses generic and ILC-specific estrogen signalling and upregulates gene sets involved in endocrine therapy resistance.	136
Figure 5-4. CAF CM induces expression of genes implicated in Tamoxifen resistance.	139
Figure 5-5. IL-6 does not drive the formation of ALDH+ breast cancer stem cells.	141
Figure 6-1. An IGF1R/InsR inhibitor reduces SUM44PE cell proliferation.	154
Figure 6-2. IL-6 treated cells have decreased rates of proliferation.	155
Figure 6-3. IL-6 in CAF CM drives EMT genes in SUM44PE cells.	156
Figure 6-4. IL-6 drives a morphological change in SUM44PE cells on CAF-derived matrices.	159
Figure 6-5. IL-6 driven changes in morphology are dependent on STAT3 expression.	161
Figure 6-6. IL-6 increases SUM44PE 2D migration.	163
Figure 6-7. Zebrafish embryo xenograft model.	164
Figure 6-8. Screening embryos immediately after injection of SUM44PE cells.	166
Figure 6-9. Representative images of 1dpi/3dpf zebrafish embryos.	168
Figure 6-10. IL-6 pre-treated SUM44PE cells more readily disseminate in ZFEX at 2dpi.	169
Figure 6-11. IL-6 pre-treatment increases the number of disseminated SUM44PE cells in the ZFEX model	171

Figure 6-12. IL-6 pre-treated SUM44PE cells trend towards being less proliferative in the ZFEX model at 3dpi/5dpf.....	172
Figure 6-13. SUM44PE cells promote angiogenesis in ZFEX model.	174
Figure 6-14. IL-6 pre-treated cells promote increased blood vessel density in the ZFEX assay.....	175
Figure 6-15. Schematic of the functional consequences of paracrine IL-6/STAT3 signalling.	184
Figure 7-1. Summary of the findings of this thesis	196
Figure Ap 1-1. Nuclear intensity of STAT3 and ER α macro.	224
Figure Ap 1-2. EdU+ cell quantification macro	224
Figure Ap 1-3. Cell shape analysis macro.....	225
Figure Ap 1-4. Representative images of SUM44PE cell shape analysis on CDMs.	225

List of Tables

Table 2-1. Cell culture media.	35
Table 2-2. Cell culture media reagent	35
Table 2-3. Recombinant proteins, antibodies and drugs used for in cell culture.	37
Table 2-4. C12-200 lipoparticle siRNA sequences.....	37
Table 2-5. Western blot buffers	38
Table 2-6. Immunofluorescence buffers	40
Table 2-7. Antibodies used for Western blotting (WB) and immunofluorescence (IF).	41
Table 2-8. Primers used for RT-qPCR.	42
Table 2-9. EdU labelling solution.....	43
Table 2-10. MDA Anderson lysis buffer	48
Table 2-11. Zebrafish experiment solutions	57
Table 3-1. Clinical details of ILC and IDC CAFs.....	62
Table 3-2. 36 genes upregulated in ILC CAFs and their expression in LCM datasets and TCGA.	67
Table 3-3. 27 ILC stroma-enriched genes differentially expressed between ILC and IDC CAFs	73
Table 3-4. Expression of genes significantly upregulated in ILC CAFs encoding for secreted proteins.....	74
Table 3-5. Expression of the 20 most highly expressed proteins in ILC CAF conditioned media.	78
Table 4-2. Expression of IL6 gene signature genes in SUM44PE RNAseq data, ER+ tumours in the TCGA, ILC CAF RNAseq13 and ILC laser capture microdissection dataset13.	114
Table 5-1. Descriptions of gene sets used in Figure 5-3.	137
Table Ap 2-1. Antibodies used for the FPPA experiment in Chapter 3.....	227
Table Ap 2-2. RPPA antibodies.....	231
Table Ap 3-1. Details of gene sets used for gene set enrichment analysis.	237
Table Ap 4-1. IL-6 dependent gene signature literature review.	246

Chapter 1 Introduction

Breast cancer is the most commonly diagnosed cancer in the UK with an average of 55 900 diagnoses per year¹. Breast cancer can be divided into subtypes based on expression of the Estrogen (ER), progesterone (PR) and Human Epidermal Growth Factor (HER2) receptors, intrinsic molecular subtypes (PAM50) and histological subtypes.

Invasive Lobular Carcinoma (ILC) accounts for 10-15% of breast cancer diagnoses, with 6765 patients diagnosed in England in 2016², and is the second most common histological subtype of breast cancer, after Invasive Ductal Carcinoma (IDC). Despite the high prevalence and distinct features of ILC, this disease has been historically under researched and no specific treatments are available for ILC patients.

Chapter 1.1 Invasive Lobular Carcinoma

ILC is defined by E-cadherin loss, with 90-100% of ILC tumours showing inactivating or truncating mutations in the *CDH1* gene^{3,4}. Due to this, ILC tumour cells invade in single-file line into the breast stroma, making detection difficult and leading to an increased stromal content in ILC tumours^{5,6}. The vast majority (90-95%) of ILC tumours are ER+ and Luminal A (80%), whereas only roughly 70% of IDC patients are ER+ and 40% are Luminal A^{3,7-9}, and ILC tumours are generally slow growing and low grade^{10,11}. ILC and IDC tumours metastasise to the brain, bone, lungs and liver, but uniquely ILC shows a predilection for metastases in the peritoneum, ovaries and gastrointestinal tract too^{7,12}. Although most research has focused on tumour cell intrinsic phenotypes, recent work has also identified differences in the stromal and immune tumour microenvironment (TME) of ILC and IDC^{5,13-15}.

The high number of ER+ ILC patients makes them ideal candidates for endocrine-targeted therapy but, due to the slow growing nature of ILC tumours, chemotherapy is significantly less effective in ILC than IDC¹⁶⁻¹⁹. Although ILC and IDC have been shown to have divergent characteristics, clinically ILC and IDC patients are treated the same.

Until recently, the vast majority of breast cancer research focused on IDC, due to the ease of use of *in vitro* and *in vivo* IDC models and the higher number of patient tumours available for research. ILC patients have historically either been excluded from trials

and retrospective analyses or combined with IDC patients in a luminal A or ER+ group, despite the myriad of histopathological, clinical and molecular differences between the two. Increasingly research is recognising ILC as a distinct disease but more work is needed to fully understand ILC and identify therapeutic options specifically tailored for ILC patients.

The key differences between ILC and IDC are summarised in Figure 1-1.

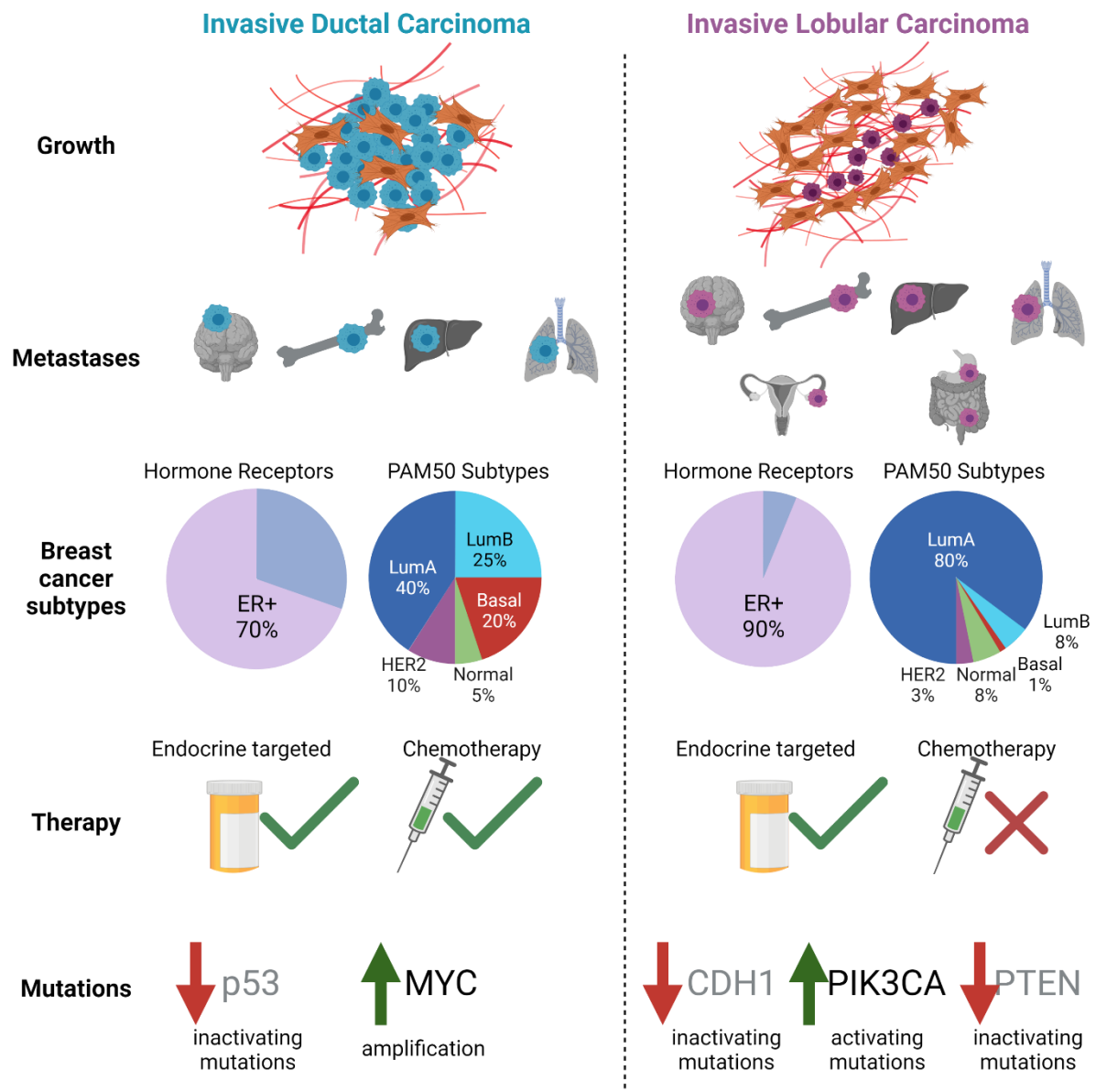


Figure 1-1. Key differences between ILC and IDC tumours. Made with Biorender.

1.1.1 Clinical features of ILC

1.1.1.1 ILC and IDC cohort characteristics and risk factors

ILC patients tend to be older at diagnosis than IDC patients¹⁹ but with later stage disease, potentially due to difficulties with detection of ILC (see 1.1.1.3). Use of menopausal hormonal therapies have been associated with increased risk of breast cancer, with a particular association with ILC compared to IDC²⁰. One meta-study found that current use of combined estrogen and progesterone menopausal hormonal therapies had a relative risk of 1.7 for IDC patients but 3.1 for ILC patients²¹. Early menarche and late age at first pregnancy are also risk factors for both ILC and IDC but are more strongly associated with ILC development.

Germline mutations in *BRCA1* and *TP53* are more strongly associated with risk of IDC than ILC development whereas *BRCA2* mutations are associated with both ILC and IDC²⁰. Germline *CDH1* mutations are rare in ILC patients, representing less than 1% of all ILC tumours. However, germline *CDH1* mutations are solely a risk factor for ILC and not IDC, though more commonly associated with diffuse gastric cancer^{20,22}.

1.1.1.2 Histopathological and Molecular Subtypes

The loss of E-cadherin in ILC leads to diffuse and discohesive single-file invasive growth and cytoplasmic re-localisation of β - and p120 catenin. There are a number of different morphological subtypes of ILC, defined by their growth pattern and morphology. The most common is classical ILC, which displays the characteristic single-file invasive pattern with a large amount of stroma and has the best outcomes²³, with other subtypes including solid, alveolar, pleomorphic and mixed-non-classic type. The histology and morphological characteristics of ILC have been extensively reviewed by Christgen *et al*²⁴.

ILC tumours have also been further categorised into three transcriptomic subtypes based on gene expression by Ciriello *et al*³. The proliferative subtype is defined by high expression of cell cycle genes and proteins and have the poorest rates of overall survival of the three. 'Immune-related' tumours have high expression of immune modulatory genes and macrophage related signatures. The 'reactive-like' subtype is defined by high levels of stromal-related signalling and has significantly lower rates of proliferation and improved overall survival compared to the proliferative subtype of ILC.

The vast majority of ILC tumours are Luminal A, with low proliferation, low grade and high expression of ER^{3,11,23,25}. However, non-classical ILC morphological subtypes are more frequently classified as Luminal B, particularly the more aggressive pleomorphic subtype²³. Very few ILC patients are classed as HER2+ (2-5%)^{3,9,10,26}, and most HER2+ tumours are non-classical subtypes. HER2+ ILC has been shown to have the poorest outcomes relative to Luminal and Triple Negative ILC (TN-ILC), though these analyses included HER2 patients from before the routine use of HER2-targeted therapies was introduced^{23,27}. TN-ILC patients represent roughly 5% of all ILC patients and 1% of all Triple Negative breast cancer (TNBC) patients in total^{10,28,29}. However, TN-ILC tumours are often found to be luminal-like rather than basal^{30,31} and to have a poorer response to chemotherapy than TN-IDC tumours^{16,28}.

The work in this thesis focuses primarily on ER+ classical ILC, which represent the majority of ILC patients.

1.1.1.3 Diagnosis and detection

Because of the single-file invasive pattern of ILC, tumours more frequently present as dimpled or 'orange-peel' like skin on the breast rather than a lump. This also makes ILC tumours difficult to detect by palpation and as well as by mammography, with one study showing a 30% false negative rate for ILC tumours assessed by mammography³². The use of ultrasound and magnetic resonance imaging (MRI) is more effective for ILC than mammography, with an accuracy of up to 100% for MRI (imaging of ILC is reviewed in detail in ^{22,33}). The use of ultrasound in screening could lead to increased detection of ILC at earlier stages. The difficulty of detection and diffuse growth patterns of ILC often means that ILC tumours are larger and at more advanced stages than IDC tumours at diagnosis^{7,10,18}.

1.1.1.4 Metastasis

ILC and IDC tumours both metastasise to the brain, bone, lung and liver, common metastatic sites for breast cancer. However, ILC tumours have a higher rate of metastasis to peritoneum, ovaries and gastrointestinal tract and are less likely to metastasise to the lungs and liver than IDC tumours^{12,34}. Metastases to these unusual sites appears to be due to ILC tumour cell intrinsic features as this metastatic pattern is recapitulated in intraductal xenograft models of the human ILC cell lines SUM44PE

and MDA-MB-134VI³⁵, and in genetically engineered mouse models of ILC^{36,37}. However, the molecular mechanisms driving this tropism have not been investigated.

ILC metastases display the same diffuse growth pattern as primary tumours, again making detection difficult. The unusual sites of metastases also make monitoring disease more difficult and a recent study showed a significantly reduced inclusion of metastatic ILC patients in clinical trials due to lack of 'measurable disease'³⁸. Due to the low metabolic activity of ILC tumour cells³⁹, FDG-PET/CT is less able to detect ILC primary and metastatic tumours than IDC⁴⁰. Recent trials have shown better uptake of the ¹⁸F-fluoroestradiol^{41,42} and ⁶⁸G-fibroblast activation protein inhibitor⁴³ PET/CT tracers in metastatic ILC than FDG-PET/CT. These tracers are targeted to ER α on tumour cells and FAP expressed by cancer associated fibroblasts respectively. This may lead to improved monitoring of metastatic ILC.

ILC patients tend to have a higher rate of late recurrences than IDC patients, with secondary tumours presenting between 10 and 15 years after initial diagnosis⁴⁴. This suggests an increased rate of dormancy in disseminated cells, a feature of ILC that remains to be extensively investigated.

1.1.1.5 Outcomes

Although ILC tumours display a number of features considered to be good prognostic indicators, including low grade, slow proliferation and high ER expression, generally ILC patients have been shown to have either similar^{7,19} or poorer^{9–11} outcomes relative to IDC patients. Generally, in the short term ILC patients have better overall survival (OS) than IDC patients, possibly due to the slow growing nature of ILC tumours. However, in longer term analyses, after 5-10 years ILC patients have been shown to have poorer outcomes than IDC patients^{8,9,44}, potentially reflecting the higher rate of late recurrences in ILC.

A recent retrospective analysis of a French multicentre cohort demonstrated poorer OS for metastatic ILC patients with matched clinical characteristics to metastatic IDC patients²⁹, although OS for TN-ILC patients was better, HER2+ ILC was similar and ER+/PR+/HER2- was worse than IDC. Another study found no differences in survival between metastatic ILC and IDC patients⁴⁵, though there was a shorter median OS for ILC patients approaching significance ($p=0.1$) and a smaller population of ILC patients compared to the French study.

Overall, this demonstrates that the positive prognostic factors associated with ILC do not translate to improved survival. This emphasises the need for ILC to be treated as a distinct disease from IDC, with more ILC-focused research and trials required.

1.1.2 Genomic and molecular characteristics of ILC

1.1.2.1 Mutations

ILC is defined by loss of E-cadherin, with deletions and inactivating or truncating mutations of the *CDH1* gene occurring in up to 95% of ILC tumours^{3,4}. *P53* mutations and amplification of the *MYC* locus are significantly less frequent in ILC than IDC.

After *CDH1* mutations, activating point mutations in *PIK3CA* are the most common, occurring in 48% of ILC tumours in the TCGA dataset. Over half of ILC tumours have mutually exclusive activating mutations in *PIK3CA* and *AKT1* or inactivating mutations in *PTEN*, leading to persistent activation of the Phosphatidylinositol-3-Kinase (PI3K) pathway. Comparatively, only 33% and 3% of IDC tumours have *PIK3CA* and *PTEN* mutations, respectively.

1.1.2.2 Phenotypic consequences of E-cadherin loss in ILC

The inactivation of E-cadherin leads to a wide range of ILC specific phenotypes, including the single-file discohesive invasive growth pattern of tumours. Although E-cadherin loss is a key marker of epithelial to mesenchymal transition (EMT), ILC cell lines still display epithelial markers and are very rounded rather than displaying an elongated mesenchymal-like phenotype. Classical mesenchymal markers such as vimentin are infrequently detected in the tumour epithelial compartment of ILC tumours⁴⁶.

ILC cell lines are far more resistant to anoikis in anchorage independent conditions than IDC cell lines⁴⁷. This has been shown to be due to E-cadherin loss⁴⁸ and the subsequent cytoplasmic localisation of the adherens junction protein p120. This promotes anoikis resistance via activation of Rho/ROCK signalling⁴⁹ and increased expression of *WNT11*⁵⁰ and the inhibitors of differentiation family of transcription factors *ID1*, *ID2* and *ID3* in anchorage independent conditions^{48,51}. These genes are more highly expressed in ILC than IDC and knockdown of *ID2* was shown to reduce lung metastasis of intravenously injected mouse ILC cell lines, highlighting this as a potentially useful therapeutic target in ILC.

1.1.2.3 IGF-1 and PI3K signalling

As mentioned, the PI3K pathway is frequently hyper-activated in ILC tumours due to *PIK3CA*, *PTEN* and *AKT1* mutations. E-cadherin loss also promotes increased autocrine Insulin-like growth factor 1/Insulin-like growth factor 1 receptor (IGF-1/IGF1R) signalling⁵² and subsequent downstream activation of the PI3K pathway⁵³, independent of PI3K pathway mutations. Mouse and human ILC cell lines have been shown to be more sensitive to IGF1R and AKT targeted drugs than IDC cell lines *in vitro* and, *in vivo*, AKT and mTOR inhibition suppresses ILC mouse model tumour growth^{53,54}.

Multiple studies have shown increased *IGF1* expression in ILC tumours compared to IDC tumours^{52,53,55}. Laser-capture microdissection datasets as well as analysis of ILC tumour cell lines and primary ILC cancer-associated fibroblasts (CAFs) by our group has also demonstrated increased ILC stroma expression of IGF-1¹³, promoting paracrine activation of the pathway. However, IGF-1 and IGF1R targeted therapies have been fairly unsuccessful in trials, with severe side effects on systemic insulin/glucose regulation (reviewed by Lee *et al*⁵⁶).

PI3K-targeted therapies have been shown to modestly increase progression free survival in metastatic breast cancer, particularly in *PIK3CA* mutated tumours⁵⁷, though with high rates of side effects and resistance. A number of trials are currently ongoing, investigating the combination of PI3K targeted drugs with CDK4/6 inhibitors and ETs in metastatic ER+/HER2- breast cancer, reviewed by Fuso *et al*⁵⁸. To date no studies have specifically investigated or reported on the response of ILC tumours to PI3K targeted therapies. However, due to the high frequency of *PIK3CA* mutations in ILC and hyper-activation of the pathway, PI3K-targeted therapy is an attractive potential therapeutic option for ILC.

1.1.2.4 Response to Estrogen and Endocrine-targeted therapies

The vast majority of ILC patients are ER+ and so endocrine targeted therapies (ETs) are the main therapeutic option for ILC patients. However, *in vitro* studies have shown that ILC cell lines respond differently from IDC cell lines to estrogen and ETs. An ILC-specific estrogen responsive set of genes has been identified and ILC cell lines have been shown to be more tamoxifen resistant than IDC cell lines^{59,60}, consistent with clinical findings (discussed in 1.1.3.1).

In vitro models of ET resistance in ILC cell lines have identified a number of resistance mechanisms, including upregulation of fatty acid metabolism⁶¹, increased expression of the ILC-specific estrogen response gene *WNT4* promoting proliferation and mTOR activation⁶², and the ILC-specific interaction between ER α and a co-regulator, MDC1⁶⁰. In other types of breast cancer, the TME has been shown to contribute to ET resistance (discussed in 1.2.3), however this has not been investigated in ILC.

A chromatin state unique to ILC driven by the pioneer factor FOXA1 has also been shown to promote tamoxifen resistance by allowing the retention of ER α at promoters that are lost in tamoxifen treated IDC cells⁵⁹. This ILC-specific FOXA1 driven chromatin state was found in the two ER+ ILC cell line models, SUM44PE and MDA-MB-134VI, as well as an ILC PDX model. Interestingly, this chromatin state was found to be independent of E-cadherin expression and to specifically associate with worse outcomes in Luminal A ILC but not IDC, highlighting another unique feature of ILC cell lines that requires further exploration.

1.1.2.5 Origin of ILC tumours

Non-transformed breast tissue is organised into ducts and lobules (Figure 1-2A) and the ducts are further organised into ER+/cytokeratin 8 (CK8)+ luminal epithelial cells, lining the lumen, surrounded by ER-/CK5+ basal myoepithelial cells (Figure 1-2B). Before sequencing of tumours was readily available, early pathologists suggested that the differences in the growth patterns of ILC and IDC may be due to them originating in the lobules and ducts, respectively²⁴. However, no recent evidence has supported this hypothesis, and differences in histo-morphological structures of ILC and IDC are now known to be caused by loss of E-cadherin and adherens junctions in ILC rather than a different cell of origin.

It has been proposed that breast cancer may arise from mutations in unipotent luminal or basal progenitors or from multipotent mammary stem cells (reviewed^{63–65}). Oncogenic activation of *PIK3CA* in basal progenitors lead to formation of luminal tumours in the mouse mammary gland, whereas activation in luminal progenitors gave rise to heterogenic tumours, with luminal and basal-like cells⁶⁶. To date, no research has specifically investigated the cell of origin for ILC tumours. A genetically engineered mouse models (GEMMs, discussed further in 1.1.4.2) using Cre recombinase under the control of the Whey acidic protein (*Wap*) promoter for

mammary specific knockout of *Cdh1* and *Pten* (WEPTn model, (*WapCre*;*Cdh1*^{ff};*Pten*^{ff})) leads to the development of classical ILC in the mouse mammary gland³⁷. Lineage tracing has shown the *Wap* promoter to be active in a bipotent progenitor cell, giving rise to both luminal and basal cells⁶⁷. However, the E-cadherin negative tumour cells formed in the WEPTn model exclusively display luminal features, with CK8 expression, similar to most human ILC tumours. Future research using conditional knock out of *Cdh1* and *Pten*, two of the most characteristic mutations in ILC, in different breast stem and progenitor cell types *in vivo* may help to identify a specific cell type of origin for ILC.

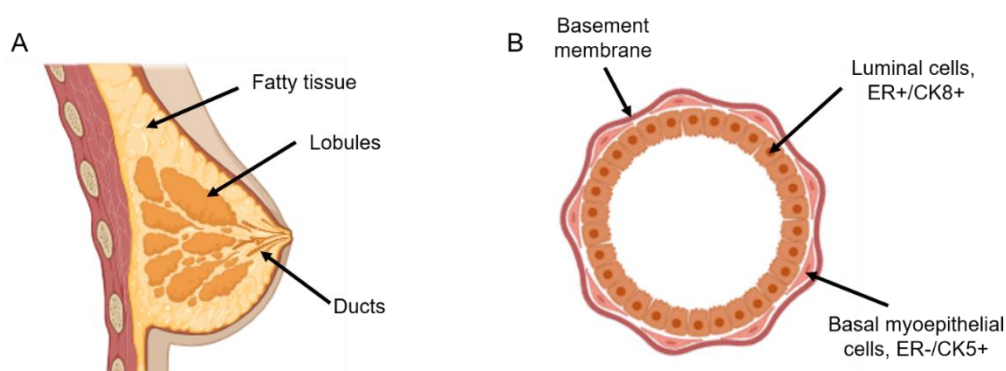


Figure 1-2. Structure of healthy mammary tissue. (A) Cross section of a healthy breast. (B) Cross section of a healthy breast duct, showing the ER+ and cytokeratin 8 (CK8+) luminal cells lining the lumen, surrounded by ER-/CK5+ basal myoepithelial cells. Made with Biorender.

1.1.3 Therapies

1.1.3.1 Endocrine targeted therapies

As over 90% of ILC patients are ER+, ILC patients are generally considered to be good candidates for endocrine therapies. However, as discussed above, ILC tumour cells *in vitro* demonstrate a divergent response to estrogen and endocrine targeted therapies and ILC patients do not show an improved response to ETs relative to IDC patients.

Analysis of the BIG 1-98 study showed that whilst there was no difference in OS or disease-free survival (DFS) for IDC patients treated with the selective estrogen modulator tamoxifen or the aromatase inhibitor letrozole, ILC patients treated with tamoxifen had a significantly worse OS and DFS than those treated with letrozole²⁵. Tamoxifen treated ILC patients also had a far worse outcome than both tamoxifen and letrozole treated IDC patients. This corroborates the *in vitro* findings that ILC cell lines are relatively tamoxifen resistant⁵⁹.

In contrast to *in vitro* studies, investigation of the response of ILC and IDC tumours to letrozole treatment showed very similar transcriptomic changes in both subtypes⁶⁸. However, this study only investigated a small number of ILC and IDC patients that all showed response to therapy. Although research investigating the differences between letrozole responsive and resistant tumours has been carried out⁶⁹, to date no studies have specifically examined the mechanisms of letrozole resistance in ILC tumours.

While ILC patients generally respond well to endocrine targeted therapies, resistance does occur and due to the poor response of ILC to chemotherapy (discussed in 1.1.3.3) there are few good follow up therapy options. In a study of matched primary and metastatic ILC tumours, 50% of metastases acquired mutations related to ET resistance, though only 11% lost expression of ER³⁴. More research is needed to understand the ILC-specific responses to estrogen and endocrine targeted therapies and how features of ILC, including signalling from the TME, influence therapy response.

1.1.3.2 CDK4/6 inhibitors

Cyclin dependent kinases 4 and 6 (CDK4/6) promote cell cycle progression from G1 to S-phase by phosphorylating and inhibiting the cell cycle checkpoint protein Rb. CDK4/6 is often deregulated in cancer, driving increased proliferation of tumour cells. The CDK4/6 inhibitors (CDK4/6i) abemaciclib, ribociclib and palbociclib are approved for treating metastatic ER+/HER2- breast cancer. Combined analysis of the MONARCH, PALOMA and MONALEESA CDK4/6i trials in combination with ETs has shown an increase in median progression free survival (PFS) for ILC patients from 16.1 months to 23.6 months⁷⁰. However, the differences in PFS between placebo and CDK4/6i treated ILC patients across all trials varied from -3.9 months to 12 months, whereas for IDC the range was from 5.9 to 15.3 months. These differences may due to smaller numbers of ILC patients included in trials limiting statistical power or could indicate a difference in efficacy of CDK4/6i for ILC patients. The ongoing PELOPS trial (NCT02764541) will specifically investigate ILC patient response to palbociclib with tamoxifen or letrozole⁷¹.

1.1.3.3 Chemotherapy

Repeatedly, studies have shown that ILC patients have decreased response rates to chemotherapy and poorer OS compared to chemotherapy-treated IDC

patients^{18,19,72,73}. Despite the large size of ILC tumours at diagnosis, the vast majority of ILC tumours are low grade with slow rates of proliferation, which could account for this discrepancy in chemotherapy response between ILC and IDC. However, ILC patients treated with chemotherapy stage matched to IDC patients still have decreased OS¹⁹ and pathologically complete response (pCR) to chemotherapy does not predict an increased OS¹⁸, suggesting other ILC phenotypes may influence chemotherapy response.

In TN-ILC, a study has shown that chemotherapy does increase OS but that the rate of pCR was still significantly lower than for TN-IDC tumours²⁸. This suggests that TN-ILC patients should receive chemotherapy, although TN-ILC patients only represent 5% of ILC patients. This emphasises the need to investigate ILC as a distinct disease and identify novel therapeutic targets for ILC therapy.

1.1.3.4 Immunotherapy

Immune checkpoint inhibitors (ICIs) are a class of drugs that relieve immunosuppression and promote cytotoxic T-cell activity against tumour cells. These are approved for use in metastatic TNBC, but TN-ILC represents less than 5% of all ILC tumours. Although analysis of bulk RNAseq data from the TCGA dataset identified an enrichment for immune cell related signatures and higher expression of immune checkpoint molecules in ILC than IDC³⁹, histological analyses have repeatedly shown reduced tumour infiltrating lymphocytes in ILC tumours^{15,34}. An ‘immune-related’ subset of ILC tumours has been identified that may benefit from ICI treatment³ and the small number of TN-ILC patients have poorer outcomes than Luminal ILC patients²³, suggesting ICIs may benefit these patients.

1.1.3.5 ILC specific clinical trials

Historically, ILC patients have either been excluded from clinical trials or results for ILC patients have not been reported separately from those of IDC patients, with both subtypes combined in an ER+ or luminal group. In trials in the metastatic setting, proportionally fewer ILC patients are included³⁸, often due to the diffuse growth pattern of ILC metastases in more unusual metastatic sites making ‘measurable disease’ more difficult to determine. Studies *in vitro* and analyses of trials where ILC-specific data is available, discussed above, have shown that ILC and IDC tumours have differing

responses to therapies. This underlines the need for increased inclusion and reporting of results for ILC patients or ILC-specific clinical trials.

To address this issue, a number of ILC-focused trials are currently ongoing or have recently reported. The ongoing ROSALINE (NCT04551495) and ROLO (NCT03620643) trials are both focused on targeting the tyrosine kinase ROS-1, the inhibition of which has shown synthetic lethality with E-cadherin loss *in vitro* and to reduce growth of primary E-cadherin negative tumours *in vivo*⁷⁴. The ROSALINE trial is using a combination of the ROS-1 inhibitor entrectinib and endocrine therapy in early stage ER+ ILC⁷⁵ whereas ROLO is using the ROS-1 inhibitor crizotinib either alone or in combination with fulvestrant in metastatic ILC⁷⁶. The PELOPS trial (NCT02764541) is also ongoing, looking at response to the CDK4/6 inhibitor palbociclib in combination with either tamoxifen or letrozole in ILC and IDC patients. The recently reported GELATo trial demonstrated partial response or stable disease in 6/23 metastatic ILC patients treated with carboplatin and the immune checkpoint inhibitor atezolizumab (anti-PD-L1)⁷⁷. However, 4/6 responsive tumours were TN-ILC, representing a relatively small population of ILC patients.

1.1.4 Models of ILC

A major issue in ILC research is the lack of *in vitro* and *in vivo* models that faithfully recapitulate the human ILC disease^{78,79}. Models of ILC have been recently reviewed in depth by Sflomos *et al*⁷⁹.

1.1.4.1 In vitro models

Although seven human cell lines have been established from ILC patient tumours, only SUM44PE and MDA-MB-134VI are Luminal A-like and ER+/HER2-, representing the majority of ILC tumours. Both of these cell lines were isolated from malignant pleural effusions from patients with luminal A ER+ tumours and have truncating point mutations in *CDH1* resulting in loss of E-cadherin expression and rounded morphology. However, they also have p53 mutations which are only present in 8% of ILC tumours³, lack mutations in the PI3K pathway, and overexpress *FGFR1* which only occurs in 9% ILC tumours, so may only represent a subset of ILC patients. Both cell lines are derived from metastases so may better reflect metastatic disease, and are very slow growing, reflecting ILC tumours but making them difficult to work with. Although all cell lines have limitations, there are many different IDC cell lines that

reflect many different genetic backgrounds. Establishing new ILC lines has had limited success, possibly due to the low grade of ILC tumours¹⁰. Work in this thesis mainly uses the SUM44PE cell line, with some experiments also carried out using the MDA-MB-134VI cell line.

The TME is important to fully understand tumour biology. Our group has isolated primary patient derived CAFs from eleven ILC and five IDC patients, the transcriptomes of which are characterised in this thesis. These represent a new *in vitro* model to investigate ILC CAFs themselves as well as the influence of CAFs on ILC tumour cells, in terms of signalling and functional changes.

1.1.4.2 Genetically engineered mouse models

The Jonkers group has developed two genetically engineered mouse models (GEMMs) of ILC, using mammary specific expression of Cre-recombinase. The KEP model (*K14-Cre;Cdh1^{ff};Tp53^{ff}*) was developed first and demonstrated that E-cadherin loss alone is not sufficient for ILC development⁸⁰. p53 mutations are infrequent in ILC^{3,81} and the KEP model most closely models the rarer and more aggressive pleomorphic subtype of ILC. KEP tumours are also ER- so this not the best model of most ILC tumours. The WEPtn model (*Wap-Cre;Cdh1^{ff};Pten^{ff}*) is weakly ER+ and has inactivation of PTEN so constitutive activation of the PI3K pathway³⁷, a common feature of ILC. These tumours more closely model classical ILC. Both the KEP and WEPtn GEMMs metastasise to ILC-enriched metastatic sites.

Tumours cells and CAFs have been isolated from both the WEPtn and KEP ILC models and are a useful additional *in vitro* ILC model. These cell lines can be re-implanted into mice either in the mammary fat pad or intraductally, although re-implanted WEPtn cells do not appear to grow as ILC tumours.

1.1.4.3 Xenografts

Engrafting ILC cell lines into the mammary fat pad has shown limited success, with very low rates of engraftment, and requires the use of estrogen supplementation which causes adverse effects on mice. Recent work by Sflomos *et al* has demonstrated that intraductal injection of human ILC cell lines SUM44PE and MDA-MB-134VI results in high rates of engraftment and closely models the development of ILC, with single-file invasive growth and metastases to ILC-enriched sites such as the gastrointestinal tract³⁵.

A number of ILC patient derived xenografts (PDX) models have also been developed^{35,82–84}. The majority of these have been established from metastatic sites, with few from primary ILC tumours. Although they maintain many of the features of the tumours they are derived from, many ILC PDXs are ER- so do not represent the majority of ILC tumours. Most ILC PDXs require estrogen supplementation when implanted into the fat pad and even then still have very low rates of engraftment. Similarly to the ER+ cell line xenografts, ER+ ILC PDXs show significantly improved engraftment without estrogen supplementation when intraductally injected and maintain their expression of ER^{35,84}. This emphasises the importance of the tumour microenvironment for tumorigenesis. However, like ILC cell line xenografts, ILC PDX xenografts are very slow growing and expansion *in vivo* can take many months to years.

Recently, zebrafish embryo xenograft assays (ZFEX) have been used to assess intra and inter-tumour heterogeneity, angiogenesis, dissemination and drug response in a wide range of cancer types^{85–91}, reviewed by Costa *et al*⁹². This provides a cheaper, faster and higher throughput alternative to mouse xenografts. This method uses far fewer cells and patient derived cell lines in the ZFEX have been shown to reflect phenotypes seen in mouse xenografts and in patients, so represents a useful preclinical model. In particular, where intraductal ILC xenografts take over a year to reach end points, the ZFEX assay is completed in 2-4 days. In this thesis, we have optimised the use of the ZFEX assay for assessing SUM44PE cell dissemination, a novel model for ILC research.

1.1.4.4 Publicly available datasets – TCGA and METABRIC

One of the most useful tools for investigating the biology of ILC is large publicly available studies in which both ILC and IDC patient tumours have been sequenced at both the genetic and transcriptomic level, with clinical characteristics available, such as molecular, histological and PAM50 subtypes and survival data. The two datasets used throughout this thesis are the Cancer Genome Atlas (TCGA)³ (<https://www.cancer.gov/tcga>) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC)⁹³, both of which have been accessed through the cbiportal website⁹⁴. Both of these datasets have mutational and bulk transcriptomic (RNAseq and microarray data) data available for over 100 ILC tumours

and over 1000 IDC tumours. The TCGA dataset also has Reverse Phase Protein Array (RPPA) data for some matched tumours, allowing correlations between a limited number of total and phospho-proteins with gene expression as well as clinical features. Alongside the information published in the initial TCGA study and available via cbioportal, further analysis of this dataset has since been published by multiple groups. For example, estimations of the stromal content of individual tumours in the TCGA determined by gene set enrichment⁹⁵ and methylation data⁹⁶ have been published, as well as the ILC-specific transcriptomic subtypes of tumours³ (see 1.1.1.2). These datasets allow validation of *in vitro* and *in vivo* experimental findings in relevant, well-annotated and easily accessible patient cohorts.

Chapter 1.2 The Tumour Microenvironment

Tumours are composed of cancer cells alongside multiple non-transformed cell types and the extracellular matrix (ECM) cells are adhered to and embedded in, forming an aberrant tissue referred to as the tumour microenvironment (TME). Whilst many tumour phenotypes can be attributed to tumour cell intrinsic features, cells of the TME can have wide ranging effects on tumorigenesis, either promoting or restraining growth, pushing tumour cells to different cell states, remodelling the ECM and driving therapy response and resistance.

The breast cancer TME consists of the ECM, innate and adaptive immune cells, endothelial cells, non-transformed epithelial cells, adipocytes, and CAFs. The TME of ILC had not been extensively investigated until recently. ILC tumours are known to have a higher stromal content than IDC tumours^{5,6}, more α -SMA⁺ CAFs⁵, fewer infiltrating lymphocytes^{15,34} and a higher ratio of pro-tumorigenic M2-like macrophages relative to M1 macrophages than IDC¹⁴ (summarised in Figure 1-3). Recent work by our group also identified a number of genes specifically enriched in the stroma of ILC, including *PAPPA*, a protease that cleaves IGF-1 binding proteins increasing local bioavailability of IGF-1 and further activating the PI3K pathway in a paracrine manner¹³.

CAFs are the dominant cell type in the TME of ILC and the current research about CAFs in ILC is discussed below (1.2.1.4). This thesis focuses on characterisation of ILC and IDC CAFs and paracrine signalling from ILC CAFs to ILC tumour cells.

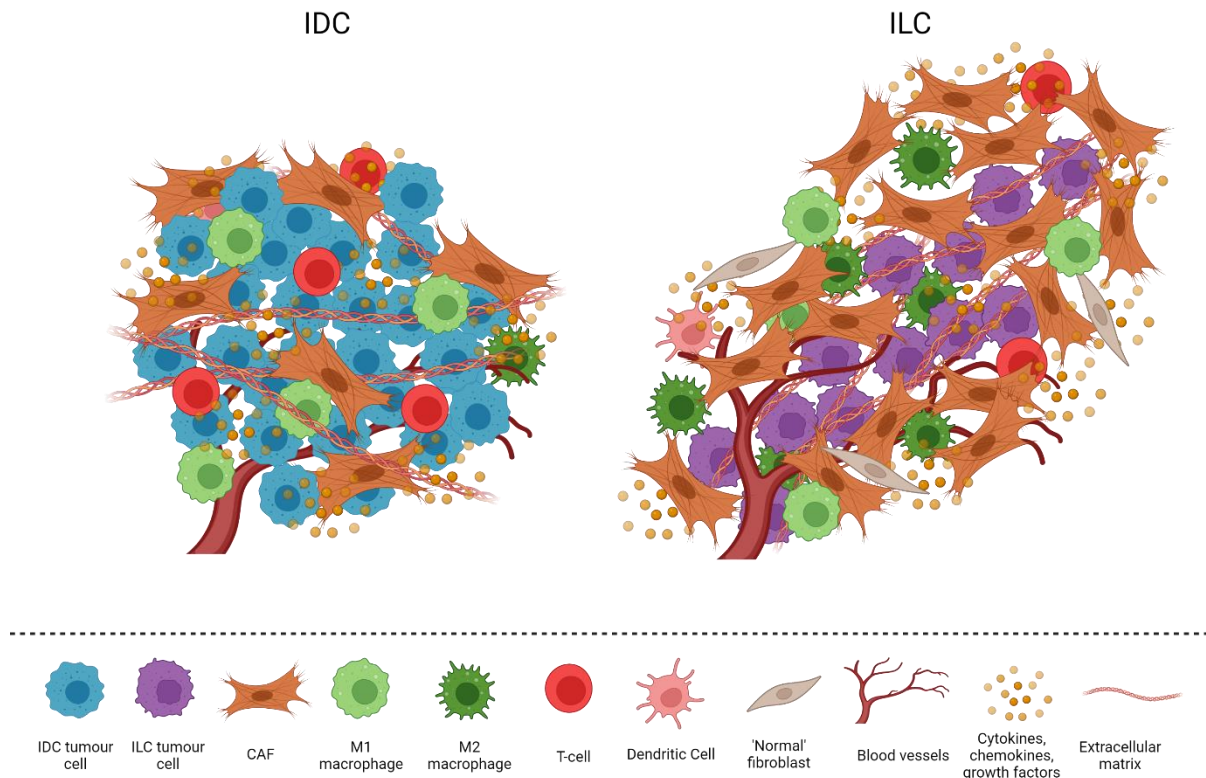


Figure 1-3. The tumour microenvironment of IDC (left, blue cells) and ILC (right, purple cells) tumours. Made with Biorender.

1.2.1 Cancer Associated Fibroblasts in Breast Cancer

Fibroblasts are resident in all healthy tissues, with diverse properties dependent on their tissue and developmental lineage (reviewed in⁹⁷). The main function of all fibroblasts is in tissue damage, in which fibroblasts are 'activated' and begin to proliferate and increase deposition and modulation of extracellular matrix. Once the damage is resolved, normal fibroblasts (NFs) generally undergo apoptosis, leaving quiescent NFs in healthy tissue. However, in cancer, persistent 'damage' signalling maintains fibroblasts in an active state, promoting their transition to cancer associated fibroblasts (CAFs).

CAFs are a heterogeneous population of stromal cells that have been shown to promote all the key hallmarks of cancer (reviewed in^{98,99}), although some tumour suppressive functions have also been reported¹⁰⁰. CAFs are defined by spindle cell shape, lack of expression of immune, endothelial or epithelial markers and expression of a combination of vimentin, α -smooth muscle actin (α -SMA), platelet-derived growth factor receptors (PDGFR α/β), CD29, fibroblast activation protein (FAP) and fibroblast

specific protein (FSP1)/S100A4. However, many of these markers are also expressed by other cell types and, to date, no unifying universal CAF-specific markers have been defined.

A summary of the origins, heterogeneity and functions of CAFs in breast cancer can be found in Figure 1-4.

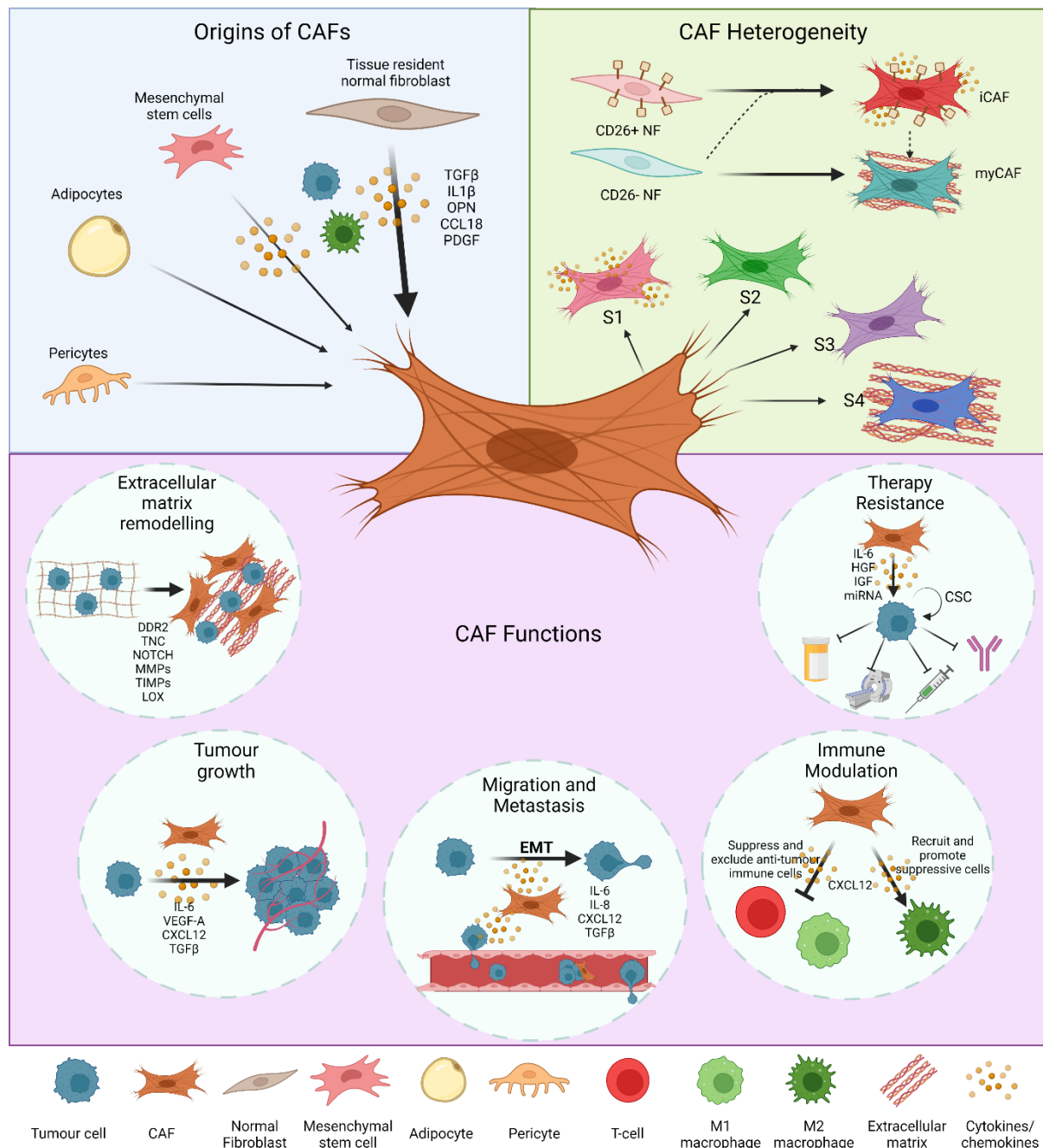


Figure 1-4. Summary of CAF origins, heterogeneity and functions in breast cancer. Made with Biorender.

1.2.1.1 Origins of breast CAFs

No definitive cell of origin has been determined for CAFs in breast cancer. The lack of CAF-specific markers and the diversity of CAF subtypes (discussed in 1.2.1.2) make lineage tracing difficult *in vivo*. In breast cancer, experimental evidence suggests that CAFs arise from tissue resident normal fibroblasts (NFs), mesenchymal stem cells (MSCs), bone-marrow derived MSCs, adipocytes and pericytes^{101–108}.

CAF activation has been shown to occur by cancer and immune cell paracrine signalling including via TGF β , PDGF, Osteopontin, IL1 β and exosomes^{109–114}, as well as increased ECM stiffness and hypoxia in the TME^{115,116}. Recent work by Houthuijzen *et al*, using a series of transplantation *in vivo* models demonstrated that the vast majority of CAFs in ILC and TNBC models originate from tissue resident fibroblasts with minimal contribution from bone marrow cells or EMT of tumour cells¹⁰¹. However, Raz *et al* demonstrated that bone-marrow derived CAFs are predominant in the TME of MMTV-PyMT tumours¹⁰², with differences in these results potentially arising from the timing of bone marrow transplantation and models used.

Adipocytes have also been proposed as the CAF cell of origin, with lineage tracing and transplantation assays showing the dedifferentiation of adipocytes towards a mesenchymal CAF-like state at tumour edges *in vivo*^{104,106}. *In vitro*, culturing of adipocytes in low pH with lactate and paracrine TGF β signalling from breast cancer cells can also induce a CAF-like state, with downregulation of adipocyte markers and upregulation of CAF markers such as α -SMA^{103,112}. Some evidence of pericyte derived or pericyte-like CAFs have been found but these represent a small proportion of CAFs mostly localised to the perivascular regions of breast tumours¹⁰⁵.

1.2.1.2 CAF Heterogeneity in breast cancer

Multiple studies have highlighted the heterogeneity of CAFs within breast tumours. Functional, spatial and transcriptomic diversity has been described in breast CAFs using single cell RNAseq (scRNAseq) and sorting by cell surface markers, both from patients and mouse models.

CAFs can broadly be divided into two categories – myofibroblastic CAFs (myCAFs), which remodel the ECM, and inflammatory CAFs (iCAFs), which secrete pro-inflammatory cytokines. Analysis of CAFs from breast tumours by cell surface markers identified 4 subtypes (CAF-S1-4), defined by differential expression of α -SMA, FAP,

FSP-1, CD29 and PDGFR β ¹¹⁷. CAF-S1 subtype, defined by high expression of all these markers, secrete high levels of CXCL12 and promote accumulation of immunosuppressive Tregs *in vitro* and is associated with increased Treg accumulation in TNBC tumours. Additional analysis of the CAF-S1 subtype by scRNAseq further delineated these CAFs into 7 clusters, with some clusters falling into the iCAF category and others showing a more myCAF phenotype, with myCAF-S1 cells promoting Treg differentiation *in vitro* and associating with ICI resistance in melanoma¹¹⁸. CAF-S1 and S4 subtypes were also identified in lymph node metastases, with CAF-S1 cells promoting invasion and EMT via paracrine TGF β and CXCL12, whereas CAF-S4 induced invasion was dependent on Notch signalling¹¹⁹, highlighting functional heterogeneity between subtypes.

Heterogeneity in resident NFs and cell of origin can also contribute to CAF heterogeneity. One study demonstrated that NFs expressing CD26 transition into immune modulatory iCAFs whereas CD26- NFs preferentially transition to ECM modulatory myCAFs during tumour progression¹⁰¹. scRNAseq of CAFs from the MMTV-PyMT model identified 3 clusters of CAFs, vCAF, mCAFs and dCAFs, and inferred that they originated from pericytes, resident fibroblasts and EMT of tumour cells respectively¹⁰⁵. A vCAF gene signature correlated with angiogenic signatures whereas the mCAF signature correlated with ECM modulating and invasion related signatures in bulk tumour RNAseq datasets.

Interestingly, many CAF subtypes have been found to be conserved across breast cancer types and in lymph node metastases (though not investigated in ILC)^{105,117–120}. Some breast CAF subtypes have also been shown to be similar to CAF subtypes identified in pancreatic ductal adenocarcinoma (PDAC) and lung adenocarcinoma^{101,105}, suggesting intra-tumoral heterogeneity and inter-tumoral similarities. However, the proportion of each CAF subtype differs across breast cancer types, with more aggressive TNBC found to have high proportions of CAF-S1 and CAF-S4 cells and luminal tumours having more CAF-S2 and S3 cells, which are also found in normal breast tissue¹¹⁷. An increased proportion of the CAF-S1 subtype in the TME of early luminal breast cancers is associated with increased risk of relapse¹²⁰, suggesting that different CAF subtypes can directly influence patient outcomes.

1.2.1.3 CAF functions in breast cancer

CAFs have been shown to promote tumour progression in breast cancer via increasing tumour growth, promoting EMT, migration and metastasis, driving immunosuppression, remodelling the ECM and enhancing therapy resistance.

Co-injection of CAFs with breast cancer cells *in vivo* enhances primary tumour growth (unpublished data by our group)^{121–123} and CAF conditioned media (CM) has been shown to increase tumour cell proliferation *in vitro*¹²⁴, as well as promoting angiogenesis via paracrine signalling by factors such as IL-6, CXCL12 and VEGF-A^{125–127}.

Either direct co-culture with CAFs or stimulation with CAF CM can promote EMT in breast cancer cells as well as enhancing migration and invasion *in vivo*. This can occur via paracrine signalling, dependent on factors such as IL-6, IL-8 and TGFβ^{122,128}, or via ECM remodelling. CD146- CAFs were shown to promote breast cancer invasion and metastasis due to deposition of a metastasis-promoting ECM via Tenascin-C activity, with this subset of CAFs associating with increased lymph node metastases¹²⁹. The CAF-S4 subset has also been shown to promote breast cancer cell invasion *in vitro*, dependent on NOTCH-driven ECM remodelling¹²⁰, and CAF expression of the collagen receptor DDR2 allows CAFs to generate an invasion-permissive ECM¹³⁰. VEGF-A expression by S100A4+ CAFs in the lung metastatic microenvironment also supports colonisation of the metastatic site by increasing angiogenesis¹²⁵.

A number of CAF secreted factors have been shown to promote breast cancer therapy resistance, such as IL-6, which is discussed in detail in Chapter 1.3. Co-culture or CAF CM has been shown to promote ER+ cell resistance to endocrine-targeted therapies by a number of mechanisms, including upregulation of the IGF-1/PI3K, mitogen activate protein kinase (MAPK) and STAT3 pathways^{131–134}. STAT3 activity in subsets of CAFs drives expression of miRNAs that are secreted in exosomes, leading to downregulation of ERα in tumour cells and cancer stem cell formation (CSC) promoting ET resistance^{135,136}. Paracrine HGF and IL-6 signalling from CAFs has also been shown to promote resistance to radiotherapy *in vitro* and *in vivo*^{124,137}, with radiotherapy increasing breast cancer cell TNFα expression leading to increased CAF proliferation and secretion of HGF in a positive feedback loop.

Crosstalk between CAFs and immune cells in the breast TME can also support tumorigenesis via the formation of an immunosuppressive TME, with a number of CAF subtypes being associated with ICI resistance¹¹⁸. CAF secretion of CXCL12 in particular has been shown to attract T-cells, causing them to differentiate to immunosuppressive Tregs^{117,120}. CAF secreted CXCL12 can also attract monocytes¹⁰¹, with CAF CM promoting differentiation towards an immunosuppressive M2-like phenotype *in vitro*¹³⁸. These CAF activated M2-like macrophages could then enhance MDA-MB-231 proliferation and migration as well as suppressing T-cell activation. Signalling from immune cells can also impact CAF activity, with a population of normal fibroblasts activated to CAFs by macrophage-secreted CCL18 promoting tumour cell chemotherapy resistance¹¹³.

However, whilst depletion of some subsets of CAFs have been shown to suppress primary tumour growth and metastasis^{125,139}, CAF depletion in an early model of PDAC actually enhanced tumour growth and aggressiveness¹⁰⁰. This underlines the need to further investigate CAFs and identify specific CAF subtypes, markers and secreted proteins that drive pro-tumorigenic phenotypes in breast cancer.

1.2.1.4 Cancer associated fibroblasts in ILC

Until recently, little work had been carried out investigating CAFs in the TME of ILC, despite the high stromal content of ILC tumours^{5,6}. Previous research has highlighted some differences in CAFs and the stroma between ILC and IDC tumours. For example, the stroma of ILC tumours has been shown to have significantly higher expression of the CAF markers FSP-1/S100A4 and FAP- α than IDC tumours¹⁴⁰ and to have a significantly increased number of proliferating α -SMA+ CAFs compared to IDC tumours⁵.

High stromal content has been associated with poorer breast cancer specific survival in TNBC⁶, whereas in ILC, high stromal content has generally been associated with better outcomes. Tumours in the 'reactive-like' subtype of ILC, which have high expression of fibroblast and stromal signatures, have a better outcome than those in the proliferative subtype³. High stromal content, as determined by pathologist assessment of H&E stained tumour cores, is also associated with better outcome for ILC patients¹⁴¹. However, many ILC tumours have also been shown to retain CD34+ fibroblasts in the TME¹⁴², which are lost in IDC tumours, and these fibroblasts are

associated with improved survival¹⁴³, though another study found no staining for CD34 in ILC tumours⁵⁵. CD34+ fibroblasts negatively correlated with α -SMA+ CAFs, suggesting that the presence of specific subtypes of fibroblasts rather than CAFs as a whole may contribute to the better outcomes for ILC patients with high stromal content. This emphasises the need for more in depth analysis of the stroma of ILC and investigation of specific CAF subtypes, markers and CAF-expressed genes and their correlation with clinical features in ILC.

In the WEPTn model of ILC, the majority of CAFs were shown to originate from breast resident fibroblasts¹⁰¹. In this same model, two CAF subtypes were identified – immunosuppressive inflammatory CAFs (iCAF) subtype and extracellular matrix (ECM) modulatory CAFs (myCAFs). These CAF subtypes identified in mouse models are enriched in the stroma of both ILC and IDC patient tumours, from laser-capture microdissection (LCM) analyses, as well as in HER2+ and TNBC tumours in single-cell RNAseq datasets. This suggests that CAF subtypes are conserved across different types of breast cancer. This conservation of CAF subtypes across breast cancers was also seen for subtypes defined by Costa *et al*¹¹⁷, although these subtypes have not been investigated in ILC.

Despite the abundance of CAFs in ILC and the phenotypic distinctions between ILC and IDC tumours, very little research has previously investigated differences between the stroma in ILC and IDC and, to our knowledge, no research has investigated paracrine CAF signalling to ILC tumour cells. Recent work by our group identified a number of ECM-related genes specifically upregulated in the stroma of ILC but not IDC or normal breast, using laser capture microdissection (LCM) of ILC tumours¹³. The work in this thesis focuses on characterising primary patient derived CAFs from ILC and IDC tumours, to complement the LCM analysis, and investigating paracrine signalling from ILC CAFs to tumour cells, in order to identify ILC-enriched CAF driven signalling.

Chapter 1.3 Interleukin-6 (IL-6) and Signal Transducer and Activator of Transcription 3 (STAT3) signalling in Breast Cancer

One of the key mediators of CAF-tumour crosstalk is the cytokine Interleukin-6 (IL-6) and its downstream effector Signal Transducer and Activator of Transcription 3 (STAT3), a transcription factor. IL-6 is a pleiotropic cytokine with a vast range of functions, both in homeostasis and in cancer, and its activation of STAT3 influences proliferation, migration and inflammatory signalling. The IL-6/STAT3 pathway has been heavily implicated in breast cancer progression but has not previously been investigated in ILC.

1.3.1 IL-6/STAT3 pathway

IL-6 binds to the transmembrane IL-6 specific receptor IL6R and forms a complex. The IL-6/IL6R complex is then recognised by the gp130 transmembrane receptor and forms a trimeric complex. gp130 can also recognise other IL-6 cytokine family members bound to their specific cognate receptors, such as IL-11, IL-31, oncostatin M (OSM) and leukaemia inhibitory factor (LIF). The trimeric IL-6/IL6R/gp130 complex then dimerises and this allows the Janus kinases (JAKs) JAK1, JAK2 and Tyk2 to be recruited to the cytoplasmic region of gp130. The JAKs phosphorylate gp130, allowing STAT3 to bind to the receptor via its SH2 domain. JAKs then phosphorylate STAT3 at Tyrosine 705 (pSTAT3/active STAT3)^{144,145}. pSTAT3 can then homodimerise and translocate to the nucleus, where it drives expression of a wide range of genes involved in proliferation, stem-ness, inflammation, EMT and therapy resistance, as well as upregulating both *IL6* and *STAT3*, driving a positive feedback loop. STAT3 activity can be inhibited via PIAS3 (protein inhibitor of activated STAT3) which binds pSTAT3 preventing STAT3 DNA-binding activity¹⁴⁶, or via STAT3-driven expression of SOCS3 (suppressor of cytokine signalling 3) which inhibits JAK2 kinase activity, providing negative feedback to the pathway¹⁴⁷.

IL-6 signalling primarily activates STAT3 but can also lead to more transient activation of STAT1. Similarly, interferon signalling mostly acts through STAT1 phosphorylation at Tyr701 leading to dimerization and transcription factor activity, but can also transiently activate STAT3¹⁴⁸. IL-6 induces homodimerisation of pSTAT3 but heterodimers of pSTAT3 and pSTAT1 can also form and drive expression of a slightly

different set of genes. Secreted forms of IL6R, generated by alternative splicing or cleavage of the transmembrane IL6R, can also bind IL-6 and subsequently gp130, leading to STAT3 activation in cells that do not express the IL6R, known as the trans-signalling pathway. A number of growth factors have also been shown to drive STAT3 activation, such as EGF, FGF-2 and IGF-1^{132,149–151}, as well as activating the PI3K and MAPK pathways. STAT3 can also be activated downstream of PI3K β by the TEK kinase BMX¹⁵² and by SRC, Integrin, NOTCH and TGF β signalling^{151,153–155}.

Tyr705 phosphorylation is the key post-translational modification that promotes STAT3 dimerisation and transcription activity, and is the focus of this thesis. However, STAT3 can also be phosphorylated at Ser727 by the MAPK and mTOR pathways, promoting expression of a subset of STAT3 target genes and influencing STAT3 activity in the mitochondria. The post-translation modifications of STAT3 and its role in the mitochondria have been reviewed by Avalle *et al*¹⁴⁷.

A summary of mechanisms of STAT3 activation can be found in Figure 1-5.

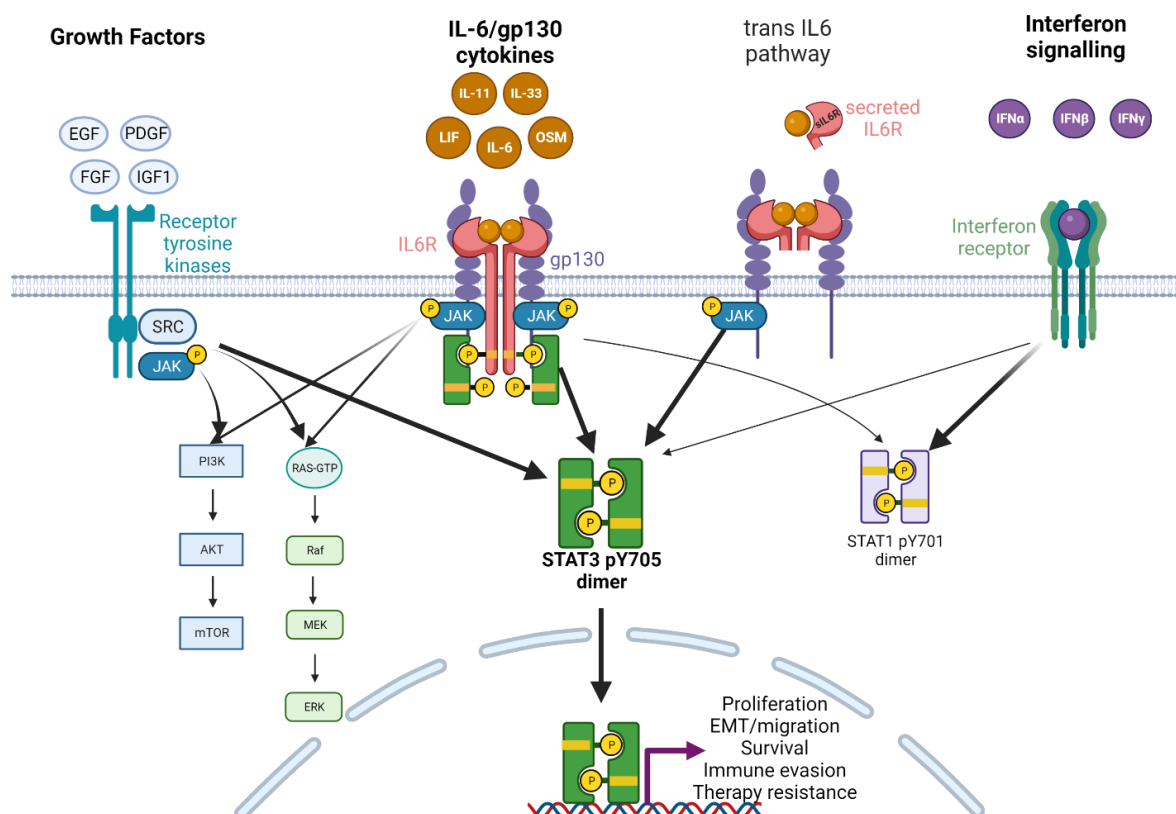


Figure 1-5. STAT3 activation. Binding of growth factors, IL-6/gp130 family cytokines or interferons to their respective transmembrane receptors results in phosphorylation of STAT3 at Tyrosine 705, allowing pSTAT3 dimers to form. Dimers of pSTAT3 then translocate to the nucleus and drive gene expression.

1.3.2 IL-6/STAT3 in breast cancer

Though both *IL6* and *STAT3* are very rarely mutated in breast cancer, with only 2% of breast tumours in the TCGA having mutations or amplifications in these genes (cbioportal.org)⁹⁴, the IL-6/STAT3 pathway is frequently deregulated, with many tumours across all breast cancer subtypes showing high expression of active nuclear pSTAT3^{156,157}. This constitutive activation of the pathway suggests an importance in breast cancer progression.

IL6 and pSTAT3 are not expressed in basal conditions by ER+ cell lines, including ILC cell lines¹⁵⁸. This may be due to the reciprocally repressive relationship between ER α and IL-6¹⁵⁹. TNBC cell lines, such as MDA-MB-231 and cells derived from the 4T1 mouse model, do express IL-6, driving autocrine activation of pSTAT3^{128,160,161}. In ER+ IDC cell lines *in vitro* and *in vivo*, paracrine IL-6 signalling from cells of the TME is mainly responsible for STAT3 activation^{124,158}, although sustained exogenous IL-6 stimulation can induce IL-6 expression in ER+ IDC cell lines¹⁶² and some ET resistant IDC cell line models express IL-6, driving autocrine activation of the pathway¹⁶³. Tumour cell expression of IL-6 is associated with more aggressive disease, lymph node involvement and poorer survival^{160,161}, and active STAT3 has been shown to correlate with poorer survival in breast cancer¹⁶⁴, though not previously specifically investigated in ILC tumours.

1.3.2.1 Breast cancer progression

In vivo studies using TNBC cell lines and the ER+ IDC cell lines MCF7 and T47D have demonstrated that IL-6 driven by autocrine IL-6 or paracrine IL-6 signalling from the TME increases primary tumour growth^{128,155,158,165–167} and STAT3 knockdown in tumour cells also suppresses tumour growth as well as altering the TME^{160,168}. Treatment of mice with IL6 or JAK inhibitors, such as the IL6R targeted monoclonal antibody tocilizumab or the JAK1/JAK2 inhibitor ruxolitinib, has been shown to suppress IL-6/STAT3 driven primary tumour growth *in vivo*. Mechanistically, IL-6/STAT3 signalling has been shown to drive growth by upregulation of pro-proliferative, anti-apoptotic and cell-cycle genes, such as MYC, BCL2 and Cyclin D1 respectively^{155,165,169}, as well as driving immunosuppression (discussed in 1.3.2.2) and angiogenesis.

IL-6 signalling promotes angiogenesis both by directly inducing VEGF expression in CAFs or endothelial cells¹²⁷ and by promoting VEGF expression downstream of STAT3 activation in breast cancer cells^{160,165}. Interestingly, paracrine IL-6 signalling from MCF10A cells with an activating mutation in *PIK3CA* promotes increased blood vessel permeability and sprouting *in vitro*, also dependent on MCF10A secretion of soluble IL6R¹⁷⁰.

Increased IL-6 and STAT3 signalling also promotes increased metastasis *in vivo* in TNBC and IDC models^{154,160,163,165} and stromal IL-6 expression at the invasive edge of breast tumours associates with increased lymph node involvement¹⁶⁰. Treatment of an intraductal ER+ PDX with ruxolitinib significantly reduced the invasive capacity of these tumours¹⁶⁵ and, in a model of HER2 amplification, STAT3 loss had no effect on primary tumour growth but increased anoikis and decreased metastasis *in vivo*¹⁷¹.

Activation of STAT3 by IL-6, NOTCH4 and fibronectin binding promotes EMT in IDC and TNBC cell lines, with increased expression of EMT transcription factors, mesenchymal-like morphology and increased expression of MMP2 and MMP9, promoting MMP-dependent invasion^{151,154,167,172–175}. However, in an intraductal MCF7 xenograft, IL-6/STAT3 signalling has been implicated in driving mesenchymal to epithelial transition of disseminated cells in metastatic sites, with STAT3 directly binding the CDH1 promoter and promoting metastatic outgrowth of dormant disseminated cells¹⁷⁶. This emphasises the context dependent activity and complexity of the IL-6/STAT3 pathway. The impact of this pathway on ILC growth and metastasis has not been investigated, with all the studies mentioned using IDC or TNBC cell lines and mouse models.

1.3.2.2 Influence on the TME

STAT3 activation in breast cancer cells can also influence the behaviour of cells in the TME, for example by promoting the formation of an immunosuppressed microenvironment. In a model of PTEN deficient breast cancer, immunosuppression was shown to be driven by PI3K β activation of STAT3 driving expression of pro-inflammatory cytokines, leading to exclusion of cytotoxic T cells¹⁵². STAT3-null tumours were shown to contain far fewer blood vessels, α -SMA+ CAFs and immunosuppressive myeloid-derived suppressor cells (MDSCs) than control tumours¹⁶⁰. Interestingly, pre-metastatic lungs in mice with STAT3-null tumours also

had fewer MDSCs, possibly due to decreased secretion of IL-6, suggesting that STAT3 activity in primary tumours can also promote the formation of a permissive pre-metastatic niche.

Paracrine IL-6 signalling directly to immune cells, either from CAFs or breast cancer cells, can also promote immunosuppression. For example, CAF-like adipocyte-secreted IL-6 can promote macrophage differentiation towards the pro-tumorigenic M2 phenotype¹⁷⁷ and tumour cell secretion of IL-6 can drive myeloid progenitor differentiation to MDSCs¹⁶¹. IL-6 driven STAT3 activation in M2-like macrophages and MDSCs increases their ability to suppress proliferation and cytotoxicity of T-cells.

IL-6/STAT3 activity in CAFs can also contribute to breast cancer progression and therapy resistance. High STAT3 expression in the stroma of TNBC tumours is associated with poorer outcomes and tumours with high stromal STAT3 had reduced infiltration of cytotoxic T-cells and increased *VEGFA* expression¹⁷⁸. Autocrine IL-6 signalling in CAFs and paracrine IL-6 signalling from TNBC cells has been shown to sustain the active phenotypes of CAFs and promote their pro-tumorigenic activity^{127,166,179}. STAT3 in CAFs drives expression of a number of pro-proliferative and pro-migratory secreted factors, including IL-6, which promotes growth and migration both *in vitro* and *in vivo* in the 4T1 mouse model of TNBC¹²¹. STAT3 can also drive expression of miRNAs in CAFs that are secreted in exosomes and promote metastasis and ET resistance in breast cancer cell lines^{135,136,180}.

1.3.2.3 Breast cancer therapy resistance

STAT3 activation in breast cancer cells has been shown to drive resistance to chemotherapy, radiotherapy, immunotherapy, targeted therapies, such as CDK4/6 inhibitors and HER2 targeted therapies, and endocrine targeted therapies^{83,131,154,167,175,181–185}. Importantly paracrine IL-6 signalling from the TME has been implicated in driving these resistance phenotypes but this pathway has not been investigated in therapy resistance in ILC.

High STAT3 activity in cancer stem cells (CSCs) has also been shown to promote resistance to ETs, HER-2 targeted therapy and chemotherapy in breast cancer. For example, IL-6 secreted by a population of macrophage-activated CAFs increases the numbers of CSCs and chemotherapy resistance in MCF7 cells *in vitro*¹¹³, and in inflammatory breast cancer models, treatment with the JAK inhibitor ruxolitinib was

found to relieve chemotherapy resistance by suppressing a small population of mesenchymal-like pSTAT3 expressing CD44⁺/CD24⁻ CSCs¹⁶⁷. STAT3 activation in CSCs is also an important mechanism of ET resistance in ER⁺ breast cancer. ALDH⁺ CSCs isolated from ET resistant ER⁺ patients have high levels of pSTAT3 and the stem-like properties of these cells as well as their ET resistance was reduced on treatment with SFX-01, a STAT3-targeted stabilised sulforaphane⁸³. CD133^{hi} CSCs isolated from ET resistant patients increased autocrine IL-6 signalling and pSTAT3 levels in response to tamoxifen and fulvestrant treatment, and blockade of IL6R with the clinically approved anti-IL6R monoclonal antibody tocilizumab significantly suppressed metastasis and restored ET sensitivity *in vivo*¹⁶³.

STAT3 activation has been shown to drive ET therapy resistance in ER⁺ IDC models *in vitro* via multiple other mechanisms alongside CSC formation. Many *in vitro* models of IDC ET resistance also show increased autocrine IL-6/STAT3 signalling^{83,154,163} and IL-6 stimulation, either by paracrine signalling from MSCs or recombinant IL-6 protein, can also promote ET resistance via ER α downregulation^{131,175}. However, although IL-6 stimulation can drive the formation of a STAT3-ER α -FOXA1 complex in T47D cells, this was shown to regulate a distinct set of genes from the ER α -FOXA1 complex¹⁶⁵. Treatment of an ET resistant metastatic ER⁺ PDX model with ruxolitinib reduced invasiveness and pSTAT3 but did not suppress ER α expression or proliferation whereas fulvestrant treatment reduced ER α expression and proliferation but tumours were still invasive and expressed pSTAT3. Combined fulvestrant and ruxolitinib treatment significantly reduced both proliferation and invasion of these tumours, suggesting that in some ER⁺ tumours the IL-6/STAT3 pathway may not directly influence ER α signalling but may represent a functionally distinct pathway that can be targeted to reduce metastasis.

ILC cell lines and patients have been shown to have poorer responses to tamoxifen than IDC patients and to have a divergent response to tamoxifen and estrogen at the gene expression and chromatin level, as well as having reduced response to chemotherapy (discussed in 1.1.3). The influence of the TME and the important paracrine IL-6/STAT3 signalling pathway on estrogen signalling and therapy resistance has not been investigated in ILC.

1.3.3 IL-6/STAT3 targeted therapies and outcomes in breast cancer

A number of drugs and monoclonal antibodies targeting the IL-6/STAT3 pathway have been developed, summarised in Figure 1-6. IL-6 can be neutralised directly using the monoclonal antibody siltuximab, which is approved for use in Castleman's disease. The monoclonal antibodies tocilizumab and sarilumab target IL6R and are approved for use in arthritis. These have shown efficacy in reducing primary tumour growth and relieving therapeutic resistance in *in vivo* models of IDC and TNBC^{124,127,186}. Bazedoxifene is a selective estrogen receptor modulator and gp130 inhibitor that has shown to be effective *in vitro* and *in vivo* in inhibiting both TNBC and ER+ breast cancer cell growth^{187,188}. In TNBC, this was primarily through suppressing pSTAT3 whereas in ER+ IDC cell lines, this occurred via suppressing ER α signalling. The JAK inhibitors ruxolitinib and baricitinib are approved for use in myeloproliferative neoplasms and arthritis respectively, and have shown promise in preclinical models of breast cancer¹⁶⁵.

A number of drugs have been developed that directly target STAT3. STATTIC is a small molecular inhibitor that binds to the SH2 domain of STAT3, inhibiting its activity¹⁸⁹. However, STATTIC also has many off target effects and can suppress histone acetylation independent of STAT3 so is a non-specific inhibitor¹⁹⁰. TTI-101 also a small molecule inhibitor of STAT3 that works by binding to the pY pocket of the SH2 domain. TTI-101 is more specific and better tolerated *in vivo* with fewer off target effects than STATTIC¹⁹¹ and trials are ongoing in metastatic solid cancers, including ER+ breast cancer (NCT03195699 and NCT05384119). The sulforaphane derivative SFX-01 directly interacts with STAT3 and suppresses its phosphorylation. This has been shown to reduce the number of ALDH+ CSCs and metastasis of ET resistant PDX models⁸³ and a trial in combination with ETs is ongoing in metastatic ER+ breast cancer (NCT02970682). However, STAT3 has important roles in regulation of the immune system and wound healing in homeostasis and so targeting STAT3 directly could have adverse side effects. This will be assessed in ongoing clinical trials

Currently, none of these drugs or antibodies are approved for use in breast cancer, though many clinical trials are ongoing or have recently reported. To date, these trials have excluded ILC patients or not specifically reported on outcomes for ILC. Ongoing

clinical trials and IL-6/STAT3 pathway drugs have recently been reviewed by Rašková *et al*¹⁹² and Manore *et al*¹⁴⁴.

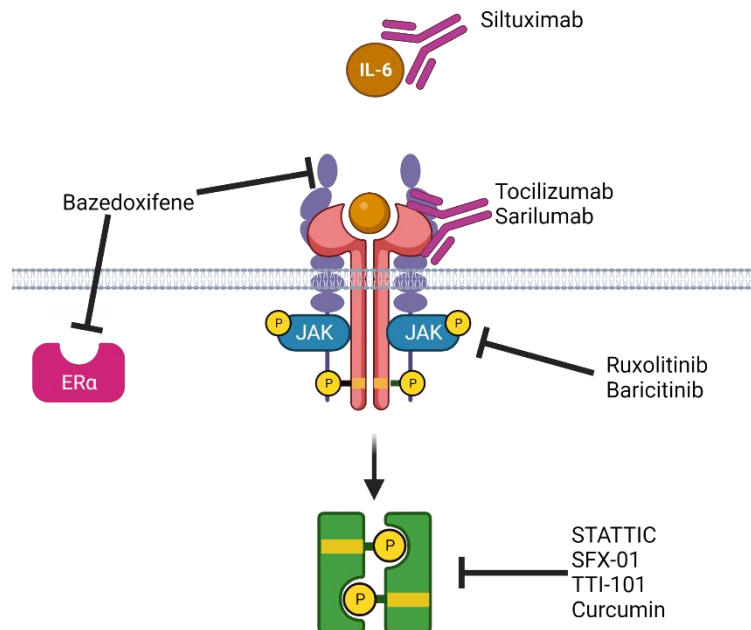


Figure 1-6. IL-6/STAT3 pathway targeted drugs and monoclonal antibodies.

Whilst experimental models both *in vitro* and *in vivo* have broadly categorised STAT3 activation as pro-tumorigenic and demonstrated the efficacy of IL-6/STAT3 targeted drugs in preclinical breast cancer models, studies of patient data gives a more mixed view. Surprisingly, nuclear STAT3 has been shown to associate with better outcomes in breast cancer^{193,194} and clinical trials using the JAK inhibitor ruxolitinib have shown limited clinical success despite strong pre-clinical evidence^{195,196}. However, some studies have shown that high levels of STAT3, STAT3 target genes or co-expression of STAT3 and IL-6 associate with poorer outcomes in breast cancer^{83,124,164,165,181} and a recent trial using SFX-01 has shown some success in ER+ ET resistant patients¹⁹⁷. This highlights the complexity of STAT3 activity and the need to identify markers or specific populations of breast cancer patients that could benefit from IL-6/STAT3 pathway targeted treatments.

Despite the central role of IL-6/STAT3 signalling in breast cancer, the role of this important signalling pathway has not previously been investigated *in vitro*, *in vivo* or in patient data in ILC.

Chapter 1.4 Hypothesis and aims

ILC is a distinct disease from IDC, yet clinically is treated the same. ILC focused research is necessary to identify ILC-enriched therapeutic targets that lead to improved outcomes for ILC patients.

The majority of research in ILC to date has focused on tumour cell intrinsic phenotypes. Despite the particularly high stromal content of ILC tumours and the strong evidence for cells of the TME promoting tumour progression in other subtypes of breast cancer, the TME has not been extensively investigated in ILC.

Our hypothesis is that CAFs in the TME of ILC tumours contribute to the distinct phenotypes and progression of this under-researched subtype of breast cancer and paracrine signalling from CAFs to tumour cells may represent potential therapeutic targets for treatment of ILC.

The specific aims of this thesis (summarised in Figure 1-7) are:

- Characterise the transcriptomes of primary patient-derived ILC and IDC CAFs
- Identify differentially expressed genes between ILC and IDC CAFs
- Identify signalling pathways activated by paracrine signalling from ILC CAFs to tumour cells
- Determine paracrine CAF driven gene expression changes in ILC tumour cells
- Determine the functional consequences of paracrine CAF signalling in ILC, including the influence on proliferation, migration and therapy resistance
- Validate *in vitro* findings in publicly available ILC and IDC datasets

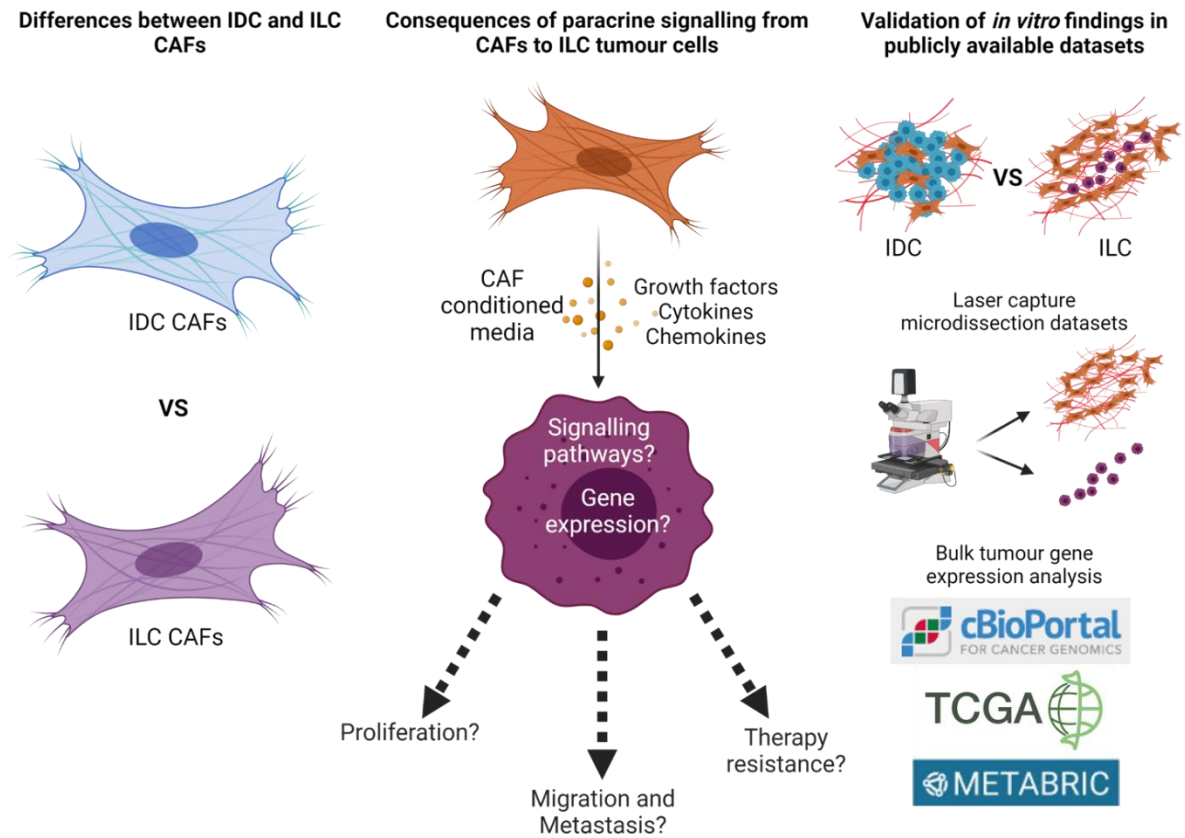


Figure 1-7. Summary of the aims of this thesis.

Chapter 2 Materials and Methods

Chapter 2.1 *In vitro* methods

2.1.1 Cell culture

SUM44PE (BioIVT), MDA-MB-134VI (ATCC), SKBR3 (a kind gift from Dr Robb Hollis, IGC, University of Edinburgh) and primary ILC and IDC cancer associated fibroblast (CAF) cells (see 2.1.1.1) were maintained in humidified air incubators with 5% CO₂ at 37°C. Details of the media each cell line was cultured in are in Table 2-1 and Table 2-2, including growth factor depleted media used for starvation before stimulation and generating conditioned media. Primary ILC and IDC CAFs were cultured on collagen coated plates (PureCol, Advanced BioMatrix, 5005, see 2.1.1.2). All cell lines were regularly tested for mycoplasma.

All cells were passaged using TrypLE (Thermo Fisher, #12605010) to detach cells and TrypLE was neutralised using appropriate complete media or in the case of SUM44PE cells, SUM44PE media + 2% FBS. SUM44PE cells and primary CAFs were frozen in Cryostor media (Stem Cell, #07933) and MDA-MB-134VI and SKBR3 cells were frozen in FBS + 10% DMSO.

Cell Line	Complete media	Growth factor depleted media
SUM44PE	Ham's F12 BSA Fraction V (1 g/l) Ethanolamine (5mM) HEPES (10 mM) Hydrocortisone (1 µg/mL) Insulin (5 µg/mL) Sodium Selenite (50 nM) Apo-transferrin (5 µg/mL) Triiodo thyronine (10 nM)	Phenol red free DMEM/F12 BSA Fraction V (1 g/l) Ethanolamine (5mM) HEPES (10 mM) Sodium Selenite (50 nM) Apo-transferrin (5 µg/mL)
MDA-MB-134VI	1:1 DMEM:L15 10% FBS 1% P/S	1:1 DMEM:L15
SKBR3	McCoy's 5a media 10% FBS 1% P/S	N/A

Primary ILC and IDC CAFs	DMEM/F12 10% FBS 1% P/S	Phenol red free DMEM/F12
-----------------------------	-------------------------------	--------------------------

Table 2-1. Cell culture media. FBS, fetal bovine serum, P/S, penicillin/streptomycin, BSA, bovine serum albumin, DMEM, Dulbecco's modified eagle medium.

Reagent	Supplier	Catalogue number
Ham's F12	Thermo Fisher	11-765-054
DMEM/F12	Sigma Aldrich	D8437
DMEM	Sigma Aldrich	D6429
L15	Thermo Fisher	11415064
McCoy's 5a	Thermo Fisher	26600023
Phenol red free DMEM/F12	Thermo Fisher	21041025
Fetal bovine serum (FBS)	Thermo Fisher	10270106
Penicillin/Streptomycin (P/S)	Thermo Fisher	15140122
Bovine serum albumin (BSA) Fraction V (7.5%)	Thermo Fisher	15260037
Ethanolamine	Sigma Aldrich	E0135
HEPES	Sigma Aldrich	H3375
Hydrocortisone	Sigma Aldrich	H4001
Insulin	Sigma Aldrich	19278
Sodium Selenite	Sigma Aldrich	S9133
Apo-transferrin	Sigma Aldrich	T2252
Triiodo Thyronine	Sigma Aldrich	T5516

Table 2-2. Cell culture media reagent

2.1.1.1 CAF isolation

CAFs were isolated by previous members of the lab (Kenneth MacLeod, Katy Teo and Laura Gómez-Cuadrado) from ILC and IDC tumours from the NHS Lothian Tissue Governance Committee under approval number 15/ES/0094. Briefly, within an hour of surgery, tumours were processed by chopping into small pieces in sterile conditions using scalpels in digestion media (DMEM/F12, collagenase-hyaluronidase enzymes (Stem Cell, #07912), P/S and fungizone) then digested for 2 hours at 37°C in a shaker. Digested tumours were strained through a 100 µm then a 40 µm strainer and then differentially centrifuged following the protocol by Speirs *et al*¹⁹⁸. Epithelial cells were isolated from the pellet after centrifugation of digested tumours at 100 xg for 2 min, then CAFs were isolated from the supernatant after centrifugation at 200 xg for 4 min.

Isolated CAFs were resuspended in DMEM/F12 + 10% FBS + 1% P/S and plated onto collagen coated plates.

2.1.1.2 Collagen coating plates

Bovine collagen I (PureCol, Advanced BioMatrix, #5005) was diluted 1:30 in distilled sterile distilled H₂O to 100 µg/mL and 4 mL of this solution was used to coat a T75 flask, coating the flask with roughly 5 µg/cm² of collagen. Plates were incubated at 37°C for 30 min then washed twice with sterile distilled H₂O then twice with PBS and stored at 4°C until use.

2.1.2 Conditioned media

CAF cells were grown to confluency in collagen coated T75 flasks and then washed twice with PBS and twice with phenol red-free DMEM/F-12 (Thermo Fisher, 21041025) and maintained in 8 mL phenol-red free DMEM/F-12 for 72 hours. Media was then collected and debris pelleted at 4°C, 1000 xg for 10 min. The supernatant was then filtered through a 0.22 µm filter and stored at -80°C until use. CAF cells were also collected and snap frozen and stored at -80°C.

2.1.3 Stimulation of cell lines

SUM44PE cells were starved in growth factor depleted media (see Table 2-1) for 24 hours before stimulation with recombinant proteins, antibodies, drugs (Table 2-3) or CAF conditioned media (CM, see 2.1.2). CAF CM was diluted 1:1 to in 2X SUM44PE growth factor depleted media (containing twice the concentration of additives as in Table 2-1) so the final concentration of all additives in the media was the same for all conditions.

For experiments up to and including 24 hours, recombinant proteins and CAF CM were diluted in growth factor depleted media and for experiments between 24 hours and 1 week, in complete media. For long term experiments (>24 hours), media was refreshed every 48 hours.

Reagent	Concentration	Supplier	Catalogue no.
Recombinant human IL-6	10 ng/mL	Biolegend	570804
Anti-IL-6	1 µg/mL	Biolegend	501125
Recombinant human FGF2	1 ng/mL	Biolegend	571506
Infgratinib (FGFRi)	2 nM	Selleck Chemicals	S2183
Recombinant IGF1	100 nM	Sigma Aldrich	I8779
OSI-906 (Lisitinib)	50 µM	Selleck Chemicals	S1091

Table 2-3. Recombinant proteins, antibodies and drugs used for in cell culture. IL-6, Interleukin-6, FGF2, Fibroblast growth factor 2, IGF1, Insulin-like growth factor 1.

2.1.4 siSTAT3

The C12-200 lipoparticle encapsulated siRNAs were a kind gift from Lidia A Valle at the University of Turin, Italy^{128,199}. SUM44PE cells were incubated with 1 µg/mL of C12-200 lipoparticles containing siRNA targeted to luciferase (siCTRL) or STAT3 (siSTAT3) in complete SUM44PE media for 48 hours (Table 2-4). Cells were then washed with PBS and stimulated as indicated.

siRNA	Target	Sense strand sequence	Anti-sense strand sequence
siCTRL	Luciferase	cuuAcGcuGAGuAcuucGAdTsdT	UCGAAGuACUcAGCGuAAGdTsdT
siSTAT3	STAT3	GccucAAGAuGAccuAGAdTsdT	UCuAGGUcAAUCUUGAGGCdTsdT

Table 2-4. C12-200 lipoparticle siRNA sequences. Lowercase indicates 2'-O-methyl modified nucleotides and phosphorothioate linkages are indicated with an s.

2.1.5 Western blotting

Cells were lysed on ice using cold RIPA buffer with protease and phosphatase inhibitors and protein concentration was determined using the BCA assay (Thermo Fisher, #23228). Lysates were normalised to 1 µg/µl by diluting in RIPA buffer and 4X NuPage LDS buffer (Invitrogen, #NP0008) with 10 mM DTT, then boiled for 10 min at 65°C. Samples were loaded into 4-15% mini-PROTEAN TGX precast gel (BioRad, 456-8085) and run in TGS running buffer at 140V for 45 min, then transferred at 25 V, 2.5 A for 10 min onto TurboBlot PVDF membrane (BioRad, #1704157). Blocking was carried out in 5% BSA in TBS-T for 45 min at room temperature, then incubated overnight at 4°C with primary antibodies (Table 2-7) diluted in block buffer. Membranes were then washed and incubated for 1 hour with HRP-conjugated secondary antibodies (Table 2-7) and visualised in BioRad imaging system with ECL detection

reagents (BioRad, 170-5061). Images were analysed in BioRad Image Lab 6.0.1 software. Western blot buffer details are in Table 2-5.

To detect both phosphorylated and total expression of a protein (eg STAT3 pY705 and total STAT3), anti-phospho protein antibodies were applied first. Once visualised, membranes were washed twice with TBS-T then twice with distilled H₂O then incubated for 10 min at room temperature in 1X ReBlot Plus Strong Antibody Stripping Solution (Millipore, #2504) diluted in distilled H₂O, washed twice in TBS-T and then blocked. Antibodies against total proteins were then applied overnight and the protocol was carried out again, as above.

Buffer	Reagent	Concentration	Supplier	Catalogue no.
RIPA buffer	Tris HCl pH 8	50 mM	Millipore	648317
	NaCl	150 mM	Sigma Aldrich	S96255
	Triton X-100	1%	Sigma Aldrich	T9284
	Sodium Deoxycholate	0.5%	Sigma Aldrich	30970
	SDS	0.1%	Sigma Aldrich	L3771
	Distilled H ₂ O	To 1L		
Lysis buffer	RIPA buffer	1 ml		
	PMSF	1.25 mM	Sigma Aldrich	P7626
	Aprotinin	0.1%	Sigma Aldrich	A6279
	NaF	0.5 mM	Sigma Aldrich	S7920
	Na ₃ VO ₄	0.1 mM	Sigma Aldrich	S6508
10X Tris-buffered saline (TBS) Adjusted to pH 7.6	Tris base	200 mM	Millipore	648310
	NaCl	1500 mM	Sigma Aldrich	S9625
	Distilled H ₂ O	To 1L		
Wash buffer (TBS-T)	10X TBS	100 ml		
	Tween-20	1 ml	Fisher	233360010
	Distilled H ₂ O	To 1 L		
TGS running buffer	10X Tris/Glycine/SDS	10 %	BioRad	161-0772
	Distilled H ₂ O	To 1 L		
Block buffer	10X TBS	50 ml		
	BSA	5%	Calbiochem	12659-500GM
	Distilled H ₂ O	450 ml		

Table 2-5. Western blot buffers. TBS, Tris buffered saline, TGS, Tris/Glycine/SDS, SDS, sodium dodecyl sulfate, BSA, bovine serum albumin.

2.1.6 Immunofluorescence

Cells were seeded onto glass coverslips and after experiments were carried out, cells were washed once with 1X TBS (**Table 2-5**) and fixed for 10 min at room temperature (RT) with PIPES/formaldehyde fixation buffer. Formaldehyde was quenched for 10 min in 0.1 M glycine then washed twice in TBS + 0.1% Triton X-100 (TBS-Tr). Samples were incubated for 30 min at RT in block buffer (TBS-Tr + 2% BSA) then incubated overnight at 4°C with primary antibodies (**Table 2-7**) diluted in block buffer. Samples were washed three times in TBS-Tr, then incubated with fluorophore-conjugated secondary antibodies (**Table 2-7**) and Phalloidin Atto-647N (Sigma Aldrich, #65906) diluted in block buffer for 45 min at RT in the dark. Samples were washed three times in TBS-Tr, once in distilled H₂O then mounted onto coverslips using Vectashield with DAPI (4',6-diamidino-2-phenylindole) (Vector Laboratories, #H-1200). Details of all buffers in **Table 2-6**. Confocal microscopy was carried out using an Olympus FV3000 Confocal Laser Scanning Microscope and Fluoview software.

All image analysis was carried out using ImageJ. To determine nuclear intensity of STAT3 and ER α , a mask of the nuclei on the DAPI channel was generated and applied to the STAT3/ER α channel. The mean gray value of STAT3/ER α in the nucleus of each cell in the image and average across the entire image was calculated in at least three fields of view (FoVs) per experiment. The macro used for this analysis can be found in **Appendix 1 Figure Ap 1-1**. The mean gray values were normalised to the mean across all control cells and differences in relative mean gray values were determined by two-way unpaired Mann-Whitney or t-tests in Graphpad Prism for STAT3 and ER α intensity respectively.

Buffer	Reagent	Concentration	Supplier	Catalogue no.
Wash buffer (TBS-Tr)	10X TBS	10%	See Table 2-5	
	Triton X-100	0.1%	Sigma Aldrich	T9284-500ML
	Distilled H ₂ O	To 100 ml		
Block buffer	BSA	2%	EMD Millipore	12659-1KG
	10X TBS	10%	See Table 2-5	
	Triton X-100	0.1%	Sigma Aldrich	T9284-500ML
	Distilled H ₂ O	To 50 ml		
PIPES/Formaldehyde fixation buffer	Formaldehyde	3.7%	Thermo Fisher	28906
	PIPES	100 mM	Sigma Aldrich	P6757
	EGTA	10 mM	Millipore	324626
	MgCl ₂	1 mM	Sigma Aldrich	M8266
	Triton X-100	0.2%	Sigma Aldrich	T9284-500ML
	Distilled H ₂ O	To 10 ml		

Table 2-6. Immunofluorescence buffers

Antibody	Application	Species	Concentration	Supplier	Catalogue No.
Primary Antibodies					
Stat3	WB IF	Rabbit	1:1000 1:500	CST	12640
Stat3 pY705	WB	Rabbit	1:1000	CST	9145
GAPDH	WB	Rabbit	1:1000	CST	5174S
ERK1/2	WB	Rabbit	1:1000	CST	9102S
ERK1/2 pT202/Y204	WB	Rabbit	1:1000	CST	9101S
ER α	IF	Rabbit	1:200	CST	13258
ALDH1A3	WB	Rabbit	1:1000	Thermo Fisher	PA5-29188
Cofilin	WB	Mouse	1:10 000	CST	5175S
Vinculin	WB	Mouse	1:1000	Sigma Aldrich	V9131
Secondary Antibodies					
Anti-rabbit- HRP	WB	Goat	1:5000	CST	7074
Anti-mouse- HRP	WB	Goat	1:5000	CST	7076
Alexa-Fluor- 488 Anti-rabbit	IF	Goat	1:500	Invitrogen	A11008
Alexa-Fluor- 594 Anti-rabbit	IF	Goat	1:500	Invitrogen	A11012
Alexa-Fluor- 488 Anti-mouse	IF	Goat	1:500	Invitrogen	A11001
Alexa-Fluor- 594 Anti-mouse	IF	Goat	1:500	Invitrogen	A11005
Ato- Phalloidin- 647	IF	-	1:400	Sigma Aldrich	65906- 10NMOL

Table 2-7. Antibodies used for Western blotting (WB) and immunofluorescence (IF).

2.1.7 RT-qPCR

RNA was extracted from cells using the Qiagen RNeasy mini kit (Qiagen, #74106) with DNase I digestion (Qiagen, #79256) following manufacturers' instructions. cDNA was generated using the Superscript First Strand-Synthesis System for RT-PCR kit (Thermo Fisher, #11904018) with 1 µg of RNA using random hexamer priming. Quantitative PCR was carried out using 6.25 µl SYBR select master mix (Thermo Fisher, # 4472908) with 100 nM forward and reverse primers (see Table 2-8) and 1 µl cDNA. Three biological replicates were analysed and duplicates of each sample were run. Conditions for qPCR – hot start for 10 min at 95°C, then 15 sec at 95°C, annealing for 30 sec at 56-62°C depending on primer T_m , 30 sec at 72°C for 40 cycles, followed by melt curve - 95°C for 15 sec, 60°C for 1 min, then +0.3°C/min for 15 min. Expression was normalised to GAPDH expression and presented as fold change relative to control sample ($2^{-\Delta\Delta C_t}$).

Target	Forward primer sequence	Reverse primer sequence	T_m	Amplicon length
STAT3	GGAACAAGCCCCAACC GA	CTAAATCAGGGGTCCCA ACTGT	61°C	171 bp
S100A8	TGCTAGAGACCGAGTGTC CTC	TGCCACGCCCATCTTTAT	59°C	134 bp
S100A9	TCCCACGAGAAGATGCAC GA	TGGCCACTGTGGTCTTAG GG	61°C	92 bp
MUCL1	GCGCCTTGCCTTCTCTTA GG	GCTGTTGTCGGATTCTGG GC	62°C	140 bp
SERPINB5	GCCTCCACATCCAGGTCT TTGT	TTGCTAGTTGCAGGGCAT CCAT	62°C	107 bp
GAPDH	GGACCTGACCTGCCGTCT AG	TGGTGCTCAGTGTAGCCC AG	62°C	111 bp

Table 2-8. Primers used for RT-qPCR.

2.1.8 Ethynyl-2'-deoxyuridine (EdU) proliferation assay

1x10⁵ cells were plated on glass coverslips in 12-well plates, with three repeats per condition. EdU (Sigma, #900584) was dissolved in PBS with 10% DMSO to make 50 mM stock solution. 10 µM EdU solution was added to the media for the stated amount of time. Cells were then washed 3x PBS, fixed in 4% formaldehyde (Thermo Fisher Scientific, #28908) for 10 min at 37°C then washed 2x PBS. The reaction was stopped with 50 mM NH₄Cl at RT for 10 minutes and washed 2x PBS. Cells were permeabilised for 10 min at RT in PBS + 0.1% Triton X-100 then washed 3x PBS. Cells were labelled using 500 µL of the TAMRA-azide EdU labelling solution per well (Table 2-9) for 30 min at RT in the dark. Cells are then washed in PBS and placed onto microscope slides with Vectashield with DAPI (Vector Laboratories, #H-1200). Slides were imaged using an Olympus FV3000 Confocal Laser Scanning Microscope and Fluoview software at 40x, with 6 random FoVs being analysed per sample and three samples analysed per condition. The number of DAPI+ and EdU+ nuclei per FoV were counted in ImageJ, using the macro in Appendix 1 Figure Ap 1-2, and the overall percentage of EdU+ cells was determined by taking the average of the 18 FoVs per condition. Data was plotted and statistics carried out in Graphpad Prism.

Reagent	Concentration	Supplier	Catalogue No.
CuSO ₄ .5H ₂ O	2 mM	Sigma Aldrich	C2284
TAMRA-azide	8 µM	Jena BioScience	CLK-AZ109
L-Ascorbic Acid	20 mg/mL	Sigma Aldrich	A92902
PBS	To 1 mL	-	-

Table 2-9. EdU labelling solution

2.1.9 ELISA

CAF conditioned media was collected as described above (see 2.1.2). Conditioned media was collected from SUM44PE cells after 24 hours maintenance of confluent cells in growth factor depleted media. Conditioned media was centrifuged, filtered and frozen as above. CAF conditioned media was diluted 1:10 in ELISA diluent, all other conditioned media used undiluted. Cells were dissociated from flasks, washed in PBS then snap frozen and stored at -80°C. Cells were then lysed and a BCA assay was performed to determine the protein concentration of the lysates – this concentration was used to normalise protein amount detected by the ELISA.

IL-6 enzyme-linked immunosorbent assay (ELISA) experiments were carried out following manufacturer's instructions, using Biolegend, ELISA MAX Deluxe Set Human IL-6, #430504. Absorbance at 450 nm and 570 nm was quantified using the Tecan Spark 20M plate reader and SparkControl software. Absorbance at 570nm was subtracted from absorbance at 450 nm and the background absorbance of media and diluent A was subtracted from final readings. The concentration of IL-6 in conditioned media was interpolated from the standard curve of IL-6 concentrations in Graphpad Prism, using least squares fit line interpolation, then multiplied by the dilution factor and normalised to the cell pellet concentration. Differences in relative IL-6 secreted by ILC and IDC CAFs determined by Mann-Whitney test in Graphpad Prism.

2.1.10 CAF derived matrices

The protocol for generating CAF derived matrices (CDMs) was adapted from the method published by Kaukonen *et al*²⁰⁰. 24-well plates or 96-well imagelock plates (Sartorius, #BA-04857) were first incubated with 0.2% gelatin (Sigma Aldrich, #G1393) diluted in PBS overnight at 4°C, then washed 3x in PBS and cross-linked using 1% glutaraldehyde (Sigma Aldrich, #354400) diluted in PBS for 30 min at room temperature. Plates were then washed and incubated with 1M glycine (Sigma Aldrich, #G8898), then washed in PBS and incubated in DMEM/F12 + 10% FBS + 1% P/S for 30 min at 37°C. For the 24-well plate, 4x10⁴ ED2334 ILC CAFs at passage 7 were seeded per well and for the 96-well image lock plates, 6500 ED26 ILC CAFs at passage 6 were seeded per well. Once confluent, the media was changed to DMEM/F12 + 10% FBS + 1% P/S with 50 µg/mL ascorbic acid (Sigma Aldrich,

#A92902) and media was changed every day for 14 days. Ascorbic acid media was made fresh every day.

To remove CAFs and retain the CDMs laid down, CDMs were denuded using extraction buffer (20 mM NH_4OH , 0.5% Triton X100 in PBS) at 37°C for 3-5 min, until no intact cells could be seen. Half the extraction buffer was gently removed and replaced with PBS twice, then remaining DNA was degraded using 10 U/mL DNase1 (Qiagen, #79256) at 37°C for 30 min. Plates were washed twice very gently with PBS then stored in PBS + 1% P/S at 4°C for up to 3 months.

Before SUM44PE cells were seeded onto the CDMs, the matrices were washed with PBS twice gently then incubated with heat denatured (heated to 75°C for 3 min then cooled) 2% BSA (Thermo Fisher #15260037) for 1 hour at 37°C. Plates were then washed in PBS and 6000 or 1000 SUM44PE cells were seeded per well for 24-well and 96-well plates respectively. Cells were allowed to adhere overnight then stimulated as indicated in experiments and brightfield images were taken at 10X using the Incucyte S3 system.

The aspect ratio of SUM44PE cells adhered to CDMs over time was measured in imageJ using the macro in Appendix 1 Figure Ap 1-3. Images were first cropped to remove borders then the background was blurred and images were thresholded to leave the outline of cells. These outlines were then converted to masks and the aspect ratio for each individual cell in each image and the mean aspect ratio of all cells larger than 50 μm^2 in an image was measured. A cell with an aspect ratio >1.7 was considered to be mesenchymal-like¹⁷⁶. An example of a brightfield image and the cell outlines generated by the macro can be found in Appendix 1 Figure Ap 1-4.

Significant differences in aspect ratio were determined in Graphpad Prism with multiple Welch's t-tests and correction by the two-stage linear step-up procedure of Benjamini, Kreiger and Yekutieli (BKY) for comparisons between time points and by a two-way ANOVA with Sidak's correction for comparisons between groups.

2.1.11 Migration assay

SUM44PE cells were seeded at 1000 cells per well in collagen coated (see 2.1.1.2) 96-well imagelock plates. Cells were left to adhere then washed once in PBS and starved in growth factor depleted media overnight. Cells were then stimulated with recombinant human IL-6 (10 ng/mL, Biolegend, #570804), CAF conditioned media or CAF conditioned media + anti-IL-6 (1 µg/mL, Biolegend, #501125) for 24 hours or 72 hours. Images were taken every 15 min on an Incucyte S3. Migration was tracked using the Trackmate²⁰¹ plugin in ImageJ using the LoG detector, with median filter and subpixel localisation, and LaP tracker. Automated thresholds were used. Significant differences in mean speed between conditions were determined by unpaired t-tests or ANOVA with Tukey's multiple comparison test in Graphpad Prism.

2.1.12 ALDEFLUOR assay

SUM44PE cells were cultured in growth factor depleted media +/- IL-6 (10 ng/mL) and SKBR3 cells were cultured in complete media for 6 days. Cells were then collected and stained using the ALDEFLUOR kit (Stem Cell Technologies, #01700), following manufacturer's instructions. Cells were collected and diluted to a concentration of 4×10^5 SKBR3 cells/mL and 5×10^5 SUM44PE cells/mL in ALDEFLUOR assay buffer. 5 µL of the ALDEFLUOR reagent was added per millilitre of cell suspension and half of each sample was then transferred and incubated with 5 µL Diethylaminobenzaldehyde (DEAB), the ALDEFLUOR inhibitor. Cells were incubated in the dark for 30 min at 37°C. Samples were then resuspended in 300 µL of ALDEFLUOR assay buffer and incubated with a final concentration of 100 ng/mL DAPI.

Samples were analysed using the BD LSR-Fortessa X-20 analyser. Data was analysed using FlowJo software. The percentage of live (DAPI negative) single ALDEFLUOR+ cells in each sample was determined relative to the matched DEAB treated sample. Differences in percentage of ALDH+ cells, laser intensity and ALDH1A3 expression were determined by two-way t-tests in Graphpad Prism.

2.1.13 FPPA

4×10^5 CAFs were seeded on collagen coated 10 cm dishes in complete media. After 24 hours, media was changed to 4 mL DMEM/F12 phenol red free for 72 hours. Conditioned media was collected as above (2.1.2) and stored at -80°C until use. Cells

were also collected and lysed using RIPA buffer, with protein concentration determined using the BCA assay. Three independent repeats were collected for each cell line.

The FPPA process was carried out by Kenneth Macleod as previously published²⁰² using validated capture-detection antibodies (Appendix 2.1). Briefly capture antibodies were spotted in quadruplet onto an Avid Nitrocellulose slide with 185 µm pins using a Quanterix-2470-microarrayer, slides were blocked then incubated with conditioned media samples overnight at 4°C. Slides were then washed and incubated with biotin-labelled secondary antibodies 1:500 in block buffer. These were detected with fluorescently labelled streptavidin and slides were analysed using an Innopsys 710 slide scanner and Mapix software. Average signal intensity for each protein was determined and median background intensity was subtracted. Intensity of spots on a media only slide was then subtracted from each sample.

Intensity was normalised to cell pellet concentration determined by BCA and proteins with a signal above 0 in at least 3 samples were considered above threshold. Log₂(mean intensity for each cell line) for each antibody above threshold was compared between ILC and IDC CAFs in Graphpad Prism with multiple Welch's t-tests and correction by BKY two-stage linear step-up procedure, with a false discovery rate (FDR)<0.05 considered significant.

2.1.14 RPPA

For the RPPA experiment, four independent biological repeats were carried out for each condition. 2.6×10^6 SUM44PE cells were grown in 10 cm dishes in SUM44PE media for 4 days, then starved for 24 hours in growth factor depleted SUM44PE media. The media was removed and ED38 ILC CAF CM mixed 1:1 with growth factor depleted media was added to the cells for 30 min and 24 hr. Control cells remained in growth factor depleted media. Cells were washed twice in PBS, which was completely removed, and then snap frozen and stored at -80°C until all repeats had been carried out. All plates were lysed together.

Cells were lysed for 10 min on ice using 100 µL MD Anderson lysis buffer (Table 2-10) then centrifuged at 13 000 rpm, 10 min, 4°C. Protein concentration was normalised to

2 mg/mL using the BCA assay (Thermo Fisher Scientific, #23227), and lysates were stored at -80°C.

The RPPA process was carried out by Kenneth Macleod as previously published²⁰³. Briefly, a serial dilution of each sample in quadruplet was printed onto an Avid Nitrocellulose slide (Grace Biolabs) with 185 µm pins, using the Aushon 1740 Arrayer platform. Antigen retrieval was carried out for 10 min with 1X Reblot Strong (Millipore, #2504), blocked for 10 min (Superblock T20 in TBS (Thermo Scientific #37536)) then incubated for 1 hour with primary antibodies (Appendix 2.2) diluted 1:250 in blocking solution. Slides were then washed and incubated with Dylight-800-labelled anti-species antibodies (Thermo Scientific, #SA5-10036 and #SA5-10176) diluted 1:2500 in block solution for 30 min. Slides were washed then scanned using an Innopsys 710 slide scanner and images were analysed using Mapix software. One slide was stained with fast green protein dye (Thermo Fisher, #A16520.22) to determine total protein concentration in each spot and scanned at 800 nm. Average signal intensity for each antibody was detected and background intensity was subtracted then normalised to total protein fast green stain intensity.

Significant changes in protein levels normalised to total protein fast green stain were assessed in Graphpad Prism 8 by two-way ANOVA with correction for multiple tests by BKY two-stage linear step-up procedure, with FDR<0.05.

Reagent	Concentration	Supplier	Catalogue No.
Triton X-100	1%	Sigma Aldrich	T9284
HEPES pH 7.4	50 mM	Sigma Aldrich	H3375
NaCl	150 mM	Sigma Aldrich	S9625-10KG
MgCl ₂	1.5 mM	Sigma Aldrich	63068
EGTA	1 mM	Sigma Aldrich	E3889
NaF	100 mM	Sigma Aldrich	S7920-100g
Na ₄ P ₂ O ₇	10 mM	Sigma Aldrich	30970-25G
Na ₃ VO ₄	1 mM	Sigma Aldrich	S6508
Glycerol	10%	Thermo Fisher	56815
Complete Mini	1 tablet	Thermo Fisher	5892791001
PhosStop	1 tablet	Sigma Aldrich	4906837001

Table 2-10. MDA Anderson lysis buffer

Chapter 2.2 RNAseq experiments and analysis

2.2.1 Primary ILC and IDC CAFs sample preparation

Primary ILC and IDC CAFs were grown to confluency on collagen coated T75 flasks in complete media. Three biological repeats of each cell line were collected, except for ED2334 (n=2). Once confluent, cells were detached, snap frozen and stored at -80°C until all samples were collected. RNA was extracted from all samples at the same time using Qiagen RNeasy kit (#74104) with DNase I (#79256) treatment. Sample collection and RNA extraction was carried out by Laura Gómez-Cuadrado.

2.2.2 CAF CM stimulation of SUM44PE cells sample preparation

To generate CAF conditioned media, 4×10^5 ED26 CAFs p5 were seeded on collagen coated 10cm dishes and after 24 hours, were washed twice with PBS, twice with DMEM/F12. CAFs were then starved for 72 hours in 4 mL DMEM/F12 and media was collected. For each biological replicate, conditioned media from two plates of ED26 p5 CAFs collected at the same time was pooled together and used to stimulate SUM44PE cells. CM was centrifuged at 400xg, 4°C, 10 min, then filtered through a 0.22 µm filter before freezing at -80°C until use.

2×10^6 SUM44PE ILC cells were seeded in 10 cm dishes. After 24 hours, cells were starved in growth factor depleted media for 24 hours. Cells were then treated for 24 hours with 6 mL of either growth factor depleted SUM44PE media or ED26 p5 CAF CM mixed 1:1 with 2X SUM44PE growth factor depleted media +/- 1 µg/mL anti-IL-6 neutralising antibody (Biolegend, #501125). Four independent replicates were carried out, with independent CAF CM samples (all from passage 5 ED26 ILC CAFs) used for each replicate. Samples were snap frozen and stored at -80°C until all replicates were collected. RNA was extracted from all samples at the same time using Qiagen RNeasy kit (#74104) with DNase I (#79256) treatment.

2.2.3 Lexogen 3`-mRNA QuantSeq

Library preparation and sequencing for both CAF and SUM44PE RNAseq experiments was carried out at the Wellcome Trust Clinical Research Facility at the University of Edinburgh. Library preparation was carried out using the Lexogen QuantSeq 3`-mRNA-Seq Library Prep kit (FWD) (Lexogen Inc, #015). Unfortunately an error during library preparation of the SUM44PE samples, in which two samples were given the

same index adapter, meant that the sequencing from these samples could not be demultiplexed and were excluded from further analysis. Single-read sequencing was performed on the NextSeq 550 platform (Illumina Inc, #SY-415-1002) using the NextSeq 500/550 Mid-Output v2.5 (150 cycle) Kit (#20024904).

The Illumina Bluebee platform was used for sequence alignment and counts, using the STAR Aligner and HTSeq-count, respectively. Differential gene expression was carried out using the EdgeR²⁰⁴ package in R version 4.2.3. For both the CAF and SUM44PE RNAseq analysis, the data was filtered to remove genes with a count-per-million (CPM) less than 1 in more than 3 samples and the data was then normalised by trimmed mean of M-values (TMM)²⁰⁵. For analysis of genes encoding secreted proteins in the CAF RNAseq dataset, raw data was filtered for genes annotated as secreted using the Gene Ontology SLIM database (GO terms: extracellular space (GO:0005615), extracellular region (GO:0005576), extracellular matrix (GO:00031012)), excluding genes annotated as part of extracellular exosomes (GO:0070062)²⁰⁶. The data was then filtered to remove genes with a CPM<5 in more than 2 samples and normalised by TMM. Differentially expressed genes for all three analyses were determined by a generalised linear model (glmFit) likelihood ratio test with FDR<0.05 considered significantly different.

2.2.4 Gene set enrichment analysis and ssGSEA

Gene set enrichment analysis (GSEA) was carried out using GSEA_4.2.2 desktop application²⁰⁷ with gene set permutation. FDR<0.05 was considered significant. Single sample GSEA (ssGSEA) was carried out using the GSVA²⁰⁸ package in R version 4.2.3.

Details of the gene sets that are not part of the Hallmarks²⁰⁹ or Gene Ontology curated gene sets can be found in Appendix 3.

2.2.5 IL6-dependent gene signature and score

Differentially expressed (FDR<0.05) genes were included in the IL-6 dependent gene signature (IL6GS) in Chapter 4 if they had a more than two-fold increase in expression on the addition of CAF CM ($\log_2(\text{CM}/\text{control}) > 1$) and a less than two-fold increase relative to control in cells treated with CAF CM + anti-IL-6 ($\log_2(\text{CM} + \text{anti-IL-6}/\text{control}) < 1$).

Z-scores for each gene in the IL6GS was calculated across all ER+ IDC (n=555) and ER+ ILC (n=191) tumours in the TCGA RNAseq dataset. The IL6GS score for each tumour sample was the sum of the z-scores for each IL6GS gene.

2.2.6 Network analysis

All networks were generated in Cytoscape 3.9.1²¹⁰. Networks of differentially expressed genes in Chapter 3 were generated using Genemania²¹¹ (accessed June 2021) and networks of significantly enriched (FDR<0.01) Gene Ontology gene sets in Chapter 4 were generated using EnrichmentMap²¹² (accessed February 2023).

2.2.7 Heatmaps

Heatmaps of the most variable genes in the CAF and SUM44PE RNAseq datasets were made using heatmap.2 in R 4.2.3. All other heatmaps show the z-scores across all samples for each gene or antibody, with unbiased hierarchical clustering by centred correlation and average linkage in Cluster 3.0²¹³ and heatmaps made using Java Treeview²¹⁴.

2.2.8 Data availability

Results of differential gene expression analysis, sample IDs and counts-per-million gene expression for each sample for the CAF and SUM44PE RNAseq datasets and a table of significantly enriched Gene Ontology Biological Process gene sets for the SUM44PE RNAseq can be found following this link: [Esme Bullock Thesis Documents](#).

Chapter 2.3 Publicly available datasets

Gene expression data from The Cancer Genome Atlas (TCGA) and METABRIC datasets were accessed using cbiportal.org⁹⁴ in September 2020, with the TCGA Firehose dataset used. All analysis was performed using ER+ ILC and ER+ IDC tumour data only. Consensus measurement of purity estimations (CPE)²¹⁵, ESTIMATE⁹⁵ (<https://bioinformatics.mdanderson.org/estimate/disease.html>) and EdEC (<http://genboree.org/theCommons/documents/569>)⁹⁶ stromal scores were accessed in October 2021. The transcriptomic subtypes of ILC tumours and gene signatures for the 'reactive-like' and 'immune related' subtypes were determined by Ciriello *et al*³. TCGA RPPA level 4 median normalised and batch corrected data was downloaded from <https://tcpaportal.org/tcpa/download.html> in April 2020.

Details of the ER+ IDC and ER+ ILC TCGA samples used, their transcriptomic subtypes, availability of stromal scores and RPPA data and the IL6 gene signature score for each sample can be found at this link: [Esme Bullock Thesis Documents](#).

Laser-capture microdissection data for ILC and IDC tumours was obtained from Gómez-Cuadrado *et al*, 2022¹³ (GSE148398) and Oh *et al*, 2015 (GSE68744)²¹⁶. Primary ILC CAF microarray data was also obtained from Gómez-Cuadrado *et al*, 2022 (GSE148156). The analysis of primary ILC and IDC CAF RNAseq data is published in Gómez-Cuadrado *et al*, 2022¹³.

Chapter 2.4 Zebrafish experiments

2.4.1 Zebrafish husbandry

Tg(Fli:eGFP) zebrafish maintenance was carried out under project license P8F7F7E52 and UK Home Office regulations, UK Animals (Scientific Procedures) Act 1986, amended in 2013, and European Directive 2010/63/EU and experiments were approved by the Home office and AWERB (University of Edinburgh Ethics Committee). All work carried out with zebrafish from 5 days post fertilisation (dpf) was carried out by Adelaide Young (PIL: I92692011). Adult Tg(Fli1:eGFP) fish were maintained at 28.5°C in 14/10 hour light/dark cycles. Prior to injection, embryos were maintained at 28.5°C and after injection, at 32°C in E3 media until harvested.

2.4.2 Zebrafish embryo xenograft assay

The zebrafish embryo xenograft (ZFEX) assay was adapted and optimised from a method developed by the Fior lab⁸⁵.

2.4.2.1 SUM44PE cell labelling

SUM44PE cells were cultured for 7 days in complete media +/- IL-6 (10 ng/mL) with media refreshed every two days. On the day of injection, cells were detached using TrypLE. 2×10^6 cells were resuspended in 1.5 mL PBS + 3 μ L Vybrant CM-Dil dye (Thermo Scientific, #V22888) and incubated at 37°C for 15 min then washed twice in sterile PBS. Dil-dyed SUM44PE cells were resuspended in PBS at a concentration of 5×10^5 cells/ μ L and stored on ice until injection.

2.4.2.2 Embryo injections

Two-day post fertilization (2dpf) Tg(Fli:eGFP) embryos were anaesthetised in 1X Tricaine and aligned on an agarose plate. Dil-dyed SUM44PE cells were backloaded into a pulled borosilicate glass capillary microinjection needle (World Precision Instruments, #TW100-4), pulled using a Laser-Based Micropipette Puller (Sutter Instrument, #P-2000). 5-10 nL of cell suspension containing roughly 200-400 cells were then injected into the perivitelline space of each embryo using a microinjector (World Precision Instruments, Pneumatic Pico pump #PV820). Injections were carried out by Adelaide Young. Embryos were then transferred to a clean petri dish with 1X Tricaine for 5 minutes to allow wounds to close and then Tricaine was replaced with

fresh E3 media and embryos were incubated at 32°C. Tricaine and E3 media details are in Table 2-11.

2.4.2.3 4 hours post injection screen

4 hours post injection, embryos were anaesthetised in 1X Tricaine and screened using the GALA mesoscope. Any embryos with no SUM44PE cells injected or very small tumours, embryos with yolk sac injected cells, dead embryos, embryos with cardiac oedema or yolk sac defects were rejected. Embryos with SUM44PE cells already outside of the primary site were separated and considered to be 'CIRC' embryos. Embryos with SUM44PE cells within the perivitelline space and no defects were used for the ZFEX assay.

2.4.2.4 1 dpi – tumour size screen

24 hours after injection (1dpi/3dpf), anaesthetised embryos were screened using the GALA mesoscope and sorted into groups depending on tumour size. Embryos with tumours in the PVS smaller than the eye were assigned to the (+) group, tumours the same size as the eye in the (++) group and tumours larger than the eye in the (+++) group. Any dead embryos or embryos with yolk sac defects or cardiac oedema were excluded from further analysis.

2.4.2.5 2 dpi – screen for embryos with disseminated cells

48 hours post injection (2dpi/4dpf), embryos were anaesthetised and each individual embryo was assessed for the presence of disseminated tumour cells on the GALA microscope. Embryos were considered to have disseminated tumour cells (MET group) if any Dil-dyed SUM44PE cells could be detected outside of the yolk sac using the GALA mesoscope at 3.2x. The percentage of embryos with disseminated cells was calculated for embryos injected with untreated SUM44PE cells and those injected with IL-6 treated SUM44PE cells. A two-way paired t-test was carried out in Graphpad Prism to determine any significant differences in number of embryos with disseminated cells between the groups.

Embryos were euthanized in 25X Tricaine. Embryos were then fixed at 2dpi/4df in PIPES/paraformaldehyde (PIPES/PFA) fixation buffer (Table 2-11) for 4 hours then stored at -20°C in methanol.

All screening was carried out on a GALA mesoscope at 3.2x with RFP laser exposure time fixed at 150 ms and 77% laser intensity. GFP laser exposure was fixed at 50 ms and 77% laser intensity. Representative images were taken at 3.2x and 8x.

2.4.3 Whole mount embryo processing and microscopy.

Embryo processing was carried out by Aleksandra Rozyczko using a protocol adapted from the Fior lab⁸⁵. Embryos that were fixed in PIPES/PFA fixation buffer and stored in methanol were rehydrated in 75%, 50% and 25% methanol diluted in PBS + 0.1% Triton (PBS-Tr), then washed 4x in PBS + 0.1% Triton for 5 min. Embryos were then washed 1x in distilled H₂O and incubated for 7 min in ice cold acetone at -20°C, washed 2x in PBS-Tr and incubated in blocking solution for 1 hour at RT then overnight at 4°C. Embryos were then washed twice in PBS-Tr for 10 min then 4x in PBS + 0.05% Tween (PBS-T) for 30 min each. Embryos were incubated with 5 µg/mL Hoechst (Thermo Fisher, #62249) in block buffer for 1 hour at RT then overnight at 4°C. Embryos were then washed 4x in PBS-T for 15 min each, fixed in PIPES/PFA fixation buffer for 20 min, washed 1x in PBS-T for 5 min. 10 embryos per slide were mounted on their sides in Mowial mounting solution (Sigma Aldrich, #81381) between two coverslips sealed with Vaseline and clear nail varnish and attached to a microscopy slide with nail varnish.

Embryos were imaged at 20x on the Andor Dragonfly confocal microscope. Z-stacks with 5 µm steps were taken of successive regions spanning the entire embryo and these images were stitched together to form one large image of the entire embryo, with no overlapping regions. A fixed exposure time and laser intensity was used for each image: 405 Hoechst/DAPI - 222.5 ms and 75%, 488 GFP - 85 ms, 50%, 561 RFP – 112.5 ms, 40%.

2.4.4 EdU staining of embryos

EdU (10 mM with 2% DMSO) was added to the E3 media of uninjected controls and SUM44PE injected Tg(Fli:eGFP) embryos at 4 dpf/2 dpi for 24 hours. Embryos were then euthanized in 25X Tricaine and fixed in PIPES/PFA fixative for 4 hours and stored in methanol at -20°C. Embryos were rehydrated and stained using the Alexa Fluor-647 azide dye following manufacturers' instructions (Invitrogen, #C10340). EdU stained embryos were imaged at 20x on the Andor Dragonfly confocal microscope as above with fixed exposure times and laser intensity - 405 Hoechst/DAPI - 222.5 ms

and 75%, 561 RFP – 112.5 ms, 40%, 647 Far red – 20 ms and 4%. An unpaired two-way t-test was carried out in Graphpad Prism to determine any significant differences in EdU positivity between the groups.

2.4.5 Whole mount embryo image analysis

2.4.5.1 Disseminated cell quantification

3D projections of the entire embryos were analysed in Imaris 9.7.2. A surface was rendered manually of the entire yolk sac region and surfaces were generated automatically for the GFP-labelled vasculature. For the vasculature surface, smooth surface detail was set at 1.5 μm with background subtraction and local contrast at 3 μm . A threshold of 4.5 μm^2 was used and regions were only kept with voxels greater than 1000. SUM44PE cells were detected automatically as 'spots' by the Imaris software. XY-dimension of spots was set as 3 μm and Z-dimension as 6 μm with background subtraction. Spots were filtered for a mean intensity greater than 240. The total number of spots throughout the entire embryo was recorded and spots outside of both the yolk and vasculature surfaces were designated as disseminated cells. The absolute number of disseminated SUM44PE cells and the percentage of disseminated SUM44PE cells relative to total SUM44PE cells was determined for each embryo. Significant difference was determined by a paired t-test in Graphpad Prism.

2.4.5.2 EdU quantification

EdU+ spots were detected automatically in Imaris with XY dimensions at 4 μm and Z-dimension as 6 μm with background subtraction and filtered for a mean intensity greater than 300. EdU spots less than 3 μm from the centre of SUM44PE Dil+ spots were considered overlapping and designated as EdU+ SUM44PE cells.

2.4.5.3 Angiogenesis quantification

For analysis of angiogenesis, a mask was manually drawn around the primary tumour in the perivitelline space and a mask of the vasculature in that region was automatically made using the same settings as above. The volume of vessels within the primary tumour mask was measured. An unpaired two-way t-test was carried out in Graphpad Prism to determine any significant differences in volume of blood vessels between the groups.

Solution	Reagent	Concentration	Supplier	Catalogue No.
E3 media (50X stock)	NaCl	250 mM	Sigma Aldrich	S9625
	KCl	8.4 mM	Sigma Aldrich	P5405
	CaCl ₂ .2H ₂ O	16.5 mM	Sigma Aldrich	C7902
	MgSO ₄ .7H ₂ O	165 mM	Sigma Aldrich	PHR2604
	Sterile H ₂ O	To 10 L		
1X E3 media	50X E3 media	2 %		
	0.01% Methylene Blue	0.3 %	Sigma Aldrich	457250
	Sterile H ₂ O	To 1 L		
Tricaine 25X stock and euthanasia Adjusted to pH 7	Tricaine powder	2 g	Stratech	C3619-APE
	Tris Base	20 mM	Melford	B2005
	MilliQ H ₂ O	500 mL		
PIPES/PFA fixative	PIPES	100 mM	Sigma Aldrich	P6757
	MgSO ₄	1 mM	Sigma Aldrich	M8266
	EGTA	2 mM	Millipore	324626
	Paraformaldehyde	1.5%	Thermo Fisher	043368-9L
	Distilled H ₂ O	To 1 mL		
Block buffer	BSA	1 %	Calbiochem	12659-500GM
	DMSO	1 %	Sigma Aldrich	D2650-100ML
	Triton X100	0.05 %	Sigma Aldrich	T9284-500ML
	Goat Serum	15 µL/mL	Abcam	ab7481
	PBS	To 50 mL		

Table 2-11. Zebrafish experiment solutions

Chapter 2.5 General

2.5.1 Statistics

Statistics were carried out using Graphpad Prism 9.5.1 or using R studio with R 4.2.3, where indicated in figure legends.

2.5.2 Schematics and Diagrams

All schematics and diagrams were generated using Biorender.com, using a license held by Esme Bullock.

Chapter 3 Characterising Invasive Lobular and Invasive Ductal Carcinoma Patient-derived Cancer Associated Fibroblasts

The primary patient derived cancer-associated fibroblasts (CAFs) used in this work were isolated from patient samples by previous members of the Brunton group – Katy Teo, Kenneth Macleod and Laura Gómez-Cuadrado. RNA for the CAF RNAseq analysis was isolated by Laura Gómez-Cuadrado. The forward-phase protein array was carried out by Kenneth Macleod at the HTPU facility in the Institute of Genetics and Cancer and the library preparation and Lexogen 3`-Quantseq was carried out at the Wellcome Trust Clinical Research Facility at the University of Edinburgh.

Parts of this work are published in the following papers^{13,101}:

Gómez-Cuadrado L, Bullock E, Mabruk Z, et al. Characterisation of the Stromal Microenvironment in Lobular Breast Cancer. *Cancers (Basel)*. 2022;14(4):904. doi:10.3390/cancers14040904

Houthuijzen, J. M. *et al.* CD26-negative and CD26-positive tissue-resident fibroblasts contribute to functionally distinct CAF subpopulations in breast cancer. *Nat. Commun.* **14**, 183 (2023). doi: 10.1038/s41467-023-35793-w

Chapter 3.1 Introduction

Despite the multitude of differences between Invasive Lobular Carcinoma (ILC) and Invasive Ductal Carcinoma (IDC) tumours, at the histological, clinical and molecular levels, therapeutically, ILC and IDC patients are treated the same. Although ILC is usually slow growing and low grade at diagnosis, patients often develop late stage recurrences and, despite having favourable clinicopathological characteristics compared to IDC, ILC patients have been shown to have either similar or worse survival than IDC patients^{7,8,19,217}.

ILC tumours have a particularly high stromal content relative to IDC tumours⁶ and cancer-associated fibroblasts (CAFs) form the majority of this non-cancerous tumour microenvironment⁵⁵. CAFs have been shown to promote the proliferation¹³², invasion¹¹⁹ and metastasis^{101,121} of cancer cells in breast cancer, as well as remodelling the extracellular matrix (ECM)^{130,218}, driving therapy resistance^{133,136,219} and promoting the formation of an immunosuppressive microenvironment^{101,117,219}. Most research into ILC has focused on tumour cell autonomous features or used bulk RNAseq of entire tumours, losing any information on reciprocal signalling between tumour cells and the tumour microenvironment. Interrogating the unique features of ILC, including the stroma, may help to identify novel therapeutic targets that can specifically benefit ILC patients.

One issue with ILC is the lack of models available to study this unique disease. Members of our lab have isolated primary patient-derived CAFs from 11 ILC and 5 IDC patients. These provide a relevant model to understand ILC biology and the interactions between CAFs and tumour cells *in vitro*. RNAseq analysis of these 16 CAF lines was carried out in order to characterise the transcriptome of these primary ILC and IDC CAFs.

The analysis in this chapter forms part of the first in depth study into the tumour microenvironment of ILC¹³, focused on cancer-associated fibroblasts, although since this work was published, studies using single cell RNAseq in mouse models and human tumours have further investigated the stromal¹⁰¹ and immune¹⁴ microenvironments of ILC tumours. RNAseq analysis shows that, while ILC and IDC CAFs on the whole do not differ dramatically, a number of genes are differentially expressed between primary patient-derived ILC and IDC CAFs. Combining the

RNAseq analysis with laser-capture microdissection (LCM) datasets and bulk tumour RNAseq from TCGA and METABRIC datasets identified genes specifically and consistently enriched in the stroma of ILC tumours, highlighting further differences between these two subtypes of breast cancer.

Chapter 3.2 Results

3.2.1 RNA-seq analysis of 11 primary ILC and 5 primary IDC CAF lines

Previous work by our group generated a laser-capture microdissection (LCM) dataset from 17 ILC tumours and identified a number of genes enriched specifically in the stroma of ILC tumours and not in IDC or normal breast¹³. To complement this work, transcriptomic analysis of CAFs isolated by differential centrifugation from 5 ER+ IDC and 11 ER+ ILC chemo-naïve tumours and grown *in vitro* for a maximum of 10 passages was carried out. Three biological replicates of each cell line were analysed (except for ED2334, where n=2), focusing on differentially expressed genes between ILC and IDC CAFs (clinical details of tumours in Table 3-1).

Principle component analysis (PCA) plots showed that IDC and ILC CAFs do not cluster separately (Figure 3-1A), though samples from the same cell lines do cluster together (Figure 3-1B). IDC and ILC CAFs also did not separate in unbiased hierarchical clustering of the 500 most variable genes across the whole dataset (Figure 3-1C), although again, samples from the same cell lines clustered together indicating a consistency between biological repeats. This demonstrates that there is not a clear distinction between ILC and IDC CAFs overall at the transcriptomic level.

CAF ID	Histological subtype	Allred Score (ER)	HER2 status	Size (mm)	Grade
ED2334	ILC	8	-	11	2
ED2425	ILC	6	-	38	2
ED26	ILC	8	-	8	2
ED28	ILC	8	-	20	2
ED29	ILC	8	-	NA	2
ED30	ILC	8	-	22	2
ED31	ILC	8	-	20	2
ED32	ILC	8	NA	70	2
ED34	ILC	8	+	13	2
ED37	ILC	7	-	NA	2
ED38	ILC	2	-	NA	2
ED2227	IDC	7	-	14	2
ED2240	IDC	8	-	13	2
ED2263	IDC	8	-	13	3
ED2327	IDC	8	-	18	2
ED9B	IDC	8	-	11	1

Table 3-1. Clinical details of ILC and IDC CAFs. NA = data not available.

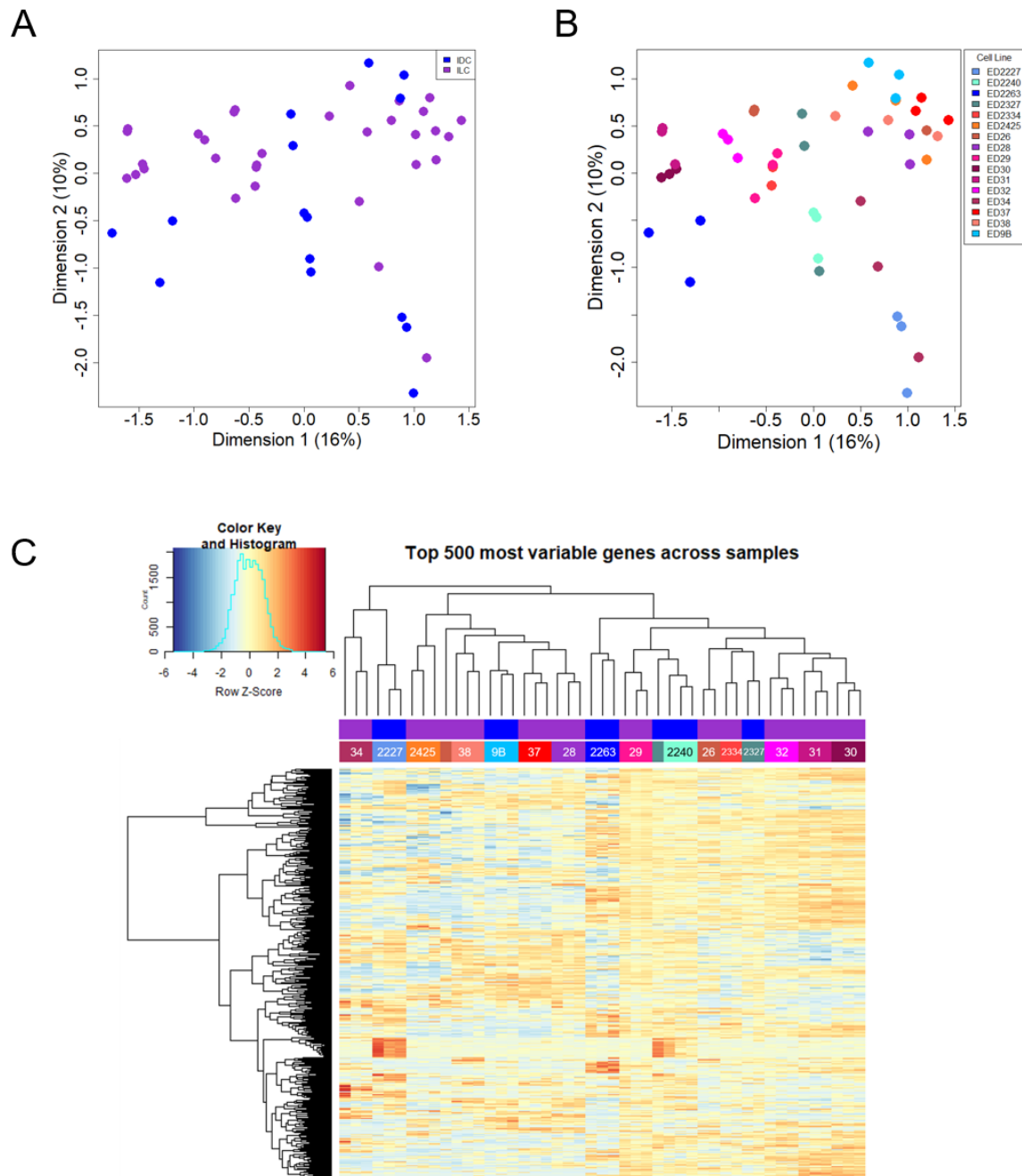


Figure 3-1. ILC and IDC CAFs are fairly similar at the transcriptional level. PCA plots of all samples in the RNAseq dataset labelled by (A) subtype (IDC=blue, ILC=purple) and (B) cell line. (C) Heatmap of 500 most variable genes across the whole dataset made using heatmap.2 in R. Unbiased hierarchical clustering of genes (rows) and samples (columns), with subtype labelled in top bar (IDC in blue, ILC in purple) and individual cell lines labelled in bottom bar.

3.2.2 Differential gene expression in ILC and IDC CAFs

Differential gene expression analysis identified 153 genes (likelihood ratio test $FDR < 0.05$, glmFit in EdgeR package²²⁰), with significantly higher expression of 117 genes in IDC CAFs and 36 genes in ILC CAFs (Figure 3-2A). 10 of the 36 genes upregulated in ILC CAFs were significantly more highly expressed in ER+ ILC tumours than ER+ IDC tumours in the TCGA dataset (Table 3-2). Interrogation of these differentially expressed genes in the LCM datasets from ILC and IDC tumours^{13,216} demonstrated that *CECR1* and *MYH3* were also significantly more highly expressed in the tumour stroma (TS) relative to the tumour epithelium (TE) in ILC tumours than in IDC tumours (Table 3-2).

Unbiased hierarchical clustering of differentially expressed genes (DEGs) separated the ILC CAFs from the IDC CAFs and the DEGs formed 4 clusters (Figure 3-2B). Whilst the genes upregulated in ILC CAFs seem to be fairly consistently upregulated across all ILC cell lines (ILC cluster) and a cluster of 20 genes is consistently upregulated across all IDC CAFs (IDC cluster), the two other clusters of IDC CAF upregulated DEGs are dominated by high expression in ED2263 and ED2227 respectively. Genes highly expressed by ED2227 seem to also be upregulated across all IDC CAFs with lower expression in ILC CAFs, whereas genes in the ED2263 cluster are almost exclusively highly expressed in this cell line. ED2263 also clustered separately from other cell lines in the PCA plots (Figure 3-1) indicating that this cell line is quite distinct from the other CAFs.

Functional annotation of the DEGs in DAVID²²¹ showed that they are enriched for plasma membrane, secreted (extracellular region and space) and extracellular matrix (ECM) related genes ($FDR < 0.05$, Figure 3-2C), though no GO biological process or molecular function terms were significantly enriched in these genes. Gene set enrichment analysis (GSEA) of Hallmark gene sets²⁰⁹ demonstrated that ILC CAFs are significantly enriched for MYC targets and mTORC1 signalling whilst TNF α signalling via NF κ B and Hypoxia gene sets are significantly enriched in IDC CAFs (Figure 3-2D), highlighting differences at both the individual gene level and functional gene set level. The enrichment for mTORC1 signalling in ILC CAFs is particularly interesting as the upstream IGF-1 and PI3K pathways have been shown to be enriched and hyperactivated in ILC tumours and cell lines^{3,52,222}, via cell autonomous

and paracrine signalling¹³, and suggests that targeting the downstream mTORC1 pathway could potentially impact both the tumour epithelium and CAFs within ILC tumours.

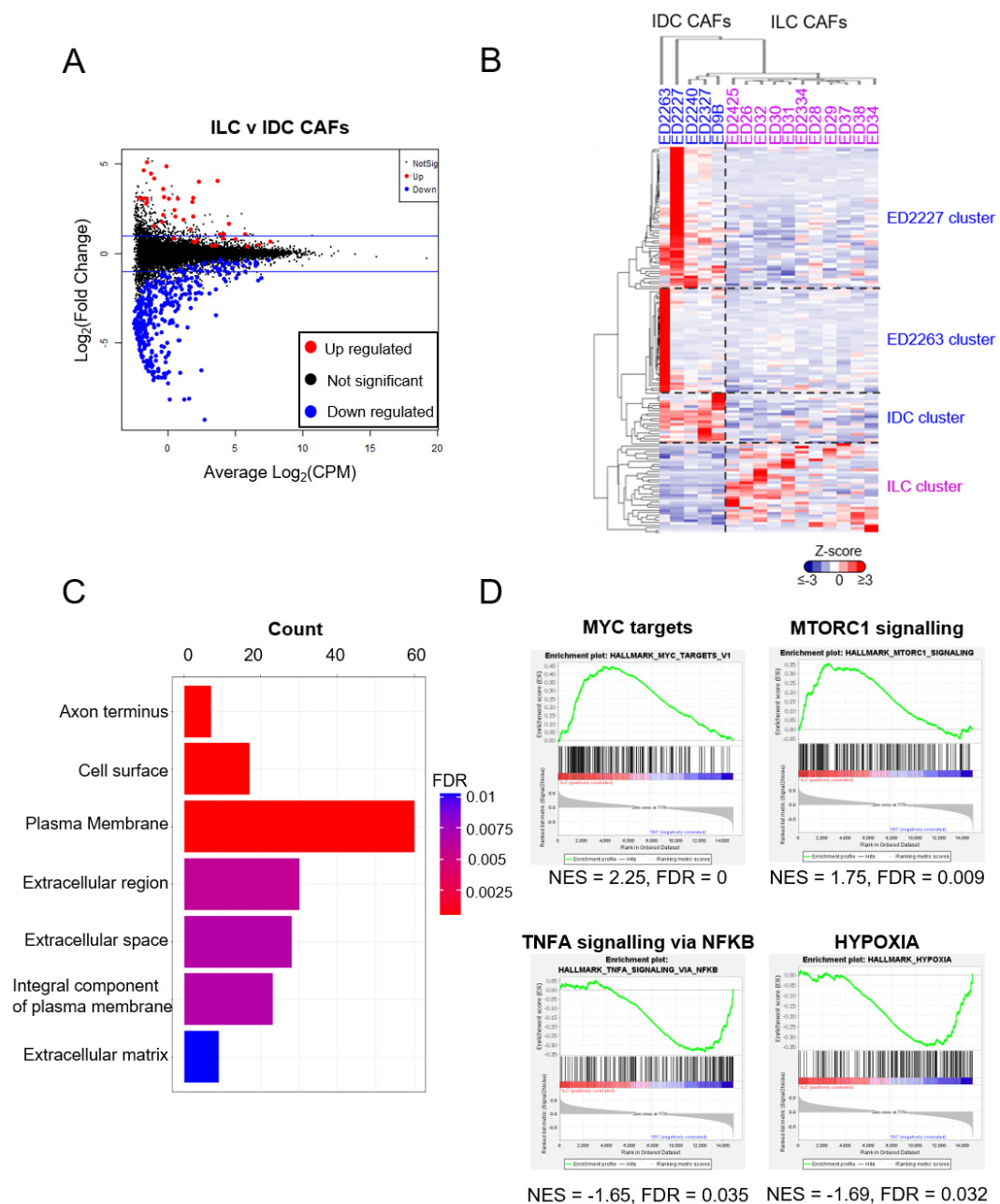


Figure 3-2. Differential gene expression in ILC and IDC CAFs. (A) Plot of differentially expressed genes between ILC and IDC CAFs, significantly upregulated genes (increased expression in ILC CAFs) labelled red and downregulated (increased expression in IDC CAFs) labelled blue, FDR<0.05 generalised linear model (glmFit) likelihood ratio test in EdgeR package. (B) Heatmap of differentially expressed genes generated using Cluster 3.0²¹³ and Java TreeView²¹⁴. Z-score expression values shown, with each column representing the average expression for each cell line. IDC CAFs labelled in blue, ILC in purple, clusters of genes upregulated in IDC CAFs labelled in blue, ILC CAFs in purple. (C) Plot of significantly enriched (FDR<0.05) GO cellular component annotation of differentially expressed genes in DAVID⁴. (D) GSEA plots of significantly enriched (FDR<0.05) Hallmark gene sets in ILC CAFs (NES>0) and IDC CAFs (NES<0), NES = normalised enrichment score. (B) and (D) adapted from¹³.

	CAF RNA-seq		LCM datasets log ₂ (TS/TE)		TCGA dataset ER+ ILC v ER+ IDC tumours
Gene	Log ₂ (ILC/IDC)	FDR	ILC	IDC	Subtype with highest expression (significance level)
<i>SORCS3</i>	5.09	0.018	-0.01	0.13	IDC
<i>CHGA</i>	4.66	0.018	0.00	0.10	IDC (***)
<i>GATD3A</i>	4.46	0.017	NA	NA	ns
<i>PURPL</i>	4.2	0.02	NA	NA	NA
<i>MTRNR2L8</i>	4.06	0.002	NA	NA	NA
<i>AREG</i>	4.03	0.045	-0.77	-1.25	IDC (*)
<i>MYH3</i>	3.11	0.013	0.82	0.03	ILC (****)
<i>ATP2A3</i>	3.08	0.007	0.00	0.05	ns
<i>HOXD11</i>	3.06	0.013	-0.09	-0.02	ILC (**)
<i>TAF4A5</i>	3.06	0.042	NA	NA	ns
<i>ADAMTS16</i>	3.04	0.04	0.11	0.05	ILC (***)
<i>TFAP2A</i>	2.9	0.008	-0.58	0.01	IDC (*)
<i>KBTBD11</i>	2.89	0.032	0.62	0.83	ILC (****)
<i>NPW</i>	2.79	0.004	-0.03	NA	ns
<i>LINC00473</i>	2.43	0.024	0.34	NA	NA
<i>PDE10A</i>	2.43	0.024	-0.16	-0.17	ILC (*)
<i>IL1A</i>	2.17	0.008	0.01	-0.40	ns
<i>FHAD1</i>	2.07	0.022	0.24	0.02	ns
<i>FAM225B</i>	1.76	0.036	0.00	NA	ILC (*)
<i>MMP16</i>	1.64	0.019	0.05	0.19	ILC (**)
<i>IDH1-AS1</i>	1.52	0.048	NA	NA	NA
<i>CECR1</i>	1.34	0.033	1.59	0.14	ILC (*)
<i>PLAAT3</i>	1.09	0.011	NA	NA	ns
<i>UNC5B</i>	1.07	0.036	1.45	2.05	ns
<i>PTGFRN</i>	1.05	0.039	0.60	-0.12	ns
<i>RAC2</i>	0.83	0.009	0.07	0.18	ILC (**)
<i>SUV39H2</i>	0.82	0.048	0.45	-0.03	IDC (****)
<i>TENM3</i>	0.82	0.024	0.07	NA	ILC (**)
<i>SNHG17</i>	0.74	0.014	-0.24	NA	ns
<i>DANCR</i>	0.67	0.007	-0.39	NA	ns
<i>RNU4ATAC</i>	0.66	0.021	-0.32	NA	ns
<i>ZNF175</i>	0.6	0.021	0.09	0.31	ns
<i>PAPSS2</i>	0.48	0.046	1.06	0.12	ns
<i>ZNF518B</i>	0.44	0.039	-0.33	NA	ns
<i>FKBP4</i>	0.42	0.023	-0.33	0.17	IDC (****)
<i>SNHG25</i>	0.4	0.045	NA	NA	NA

Table 3-2. 36 genes upregulated in ILC CAFs and their expression in LCM datasets and TCGA. Log₂(ILC CAF/IDC CAF expression) and FDR significance determined by glmFit in EdgeR. Log₂(tumour stroma/tumour epithelium) (TS/TE) for ILC tumours (n=17)¹³ and IDC tumours (n=32)²¹⁶. Expression of these genes in 113 ER+ ILC and 326 ER+ IDC tumours from the TCGA dataset. Significance determined by Wilcoxon U test with Benjamini-Hochberg (BH) correction, with subtype with the highest expression of the gene indicated and significance level in brackets (adjusted p<0.05 *, p<0.01 **, p<0.001 ***, p<0.0001 ****). ns = not significant, NA = data not available.

3.2.3 Estrogen receptor expression in ILC and IDC CAFs

Although the estrogen receptor is primarily expressed by epithelial cells in breast cancer, expression of Estrogen Receptor 1 (*ESR1*) and G-protein coupled Estrogen Receptor 1 (*GPER1*) has also been demonstrated in CAFs isolated from breast^{223–226}, prostate^{227,228} and cervical cancers²²⁹. Interestingly, both *ESR1* and *GPER1* were significantly more highly expressed in IDC CAFs than ILC CAFs (Figure 3-3A and B, *ESR1* logFC = -1.59, FDR = 0.032 and *GPER1* logFC = -0.92, FDR = 0.019).

ESR1 expression is relatively low across all CAF lines and forms part of the ‘IDC cluster’ in Figure 3-2B, indicating it is relatively consistently upregulated across all IDC CAF lines. Also, in the laser capture microdissection datasets, IDC tumours had a higher ratio of *ESR1* expression in the stroma relative to the tumour epithelium compared to ILC tumours (Figure 3-3B). This implies higher relative stromal expression of *ESR1* in the IDC microenvironment than in ILC, supporting the findings from our *in vitro* CAF RNAseq analysis. *GPER1* is more highly expressed than *ESR1* in all CAF lines and is particularly highly expressed in the ED2263 IDC CAF line, forming part of the ‘ED2263 cluster’ in Figure 3-2B. However, even excluding ED2263, mean *GPER1* expression is higher in IDC CAFs than ILC CAFs. *GPER1* was not included in the IDC LCM dataset and so could not be assessed here. As both genes have relatively low expression, validation of *GPER1* and ERα protein expression in cultured ILC and IDC CAFs by immunofluorescence is necessary.

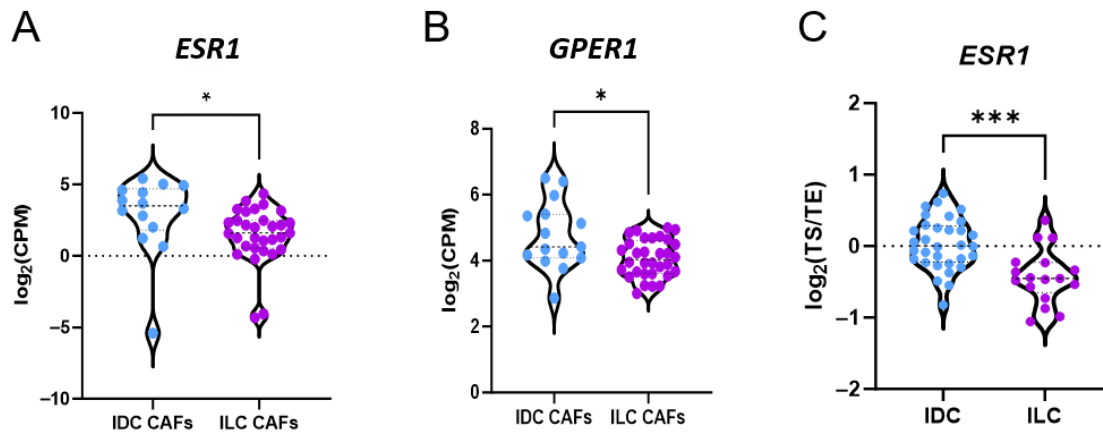


Figure 3-3. ESR1 and GPER1 expression in ILC and IDC CAFs. (A) *ESR1* and (B) *GPER1* expression in IDC (blue, $n=5$ cell lines with 3 replicates per cell line) and ILC (purple, $n=11$ cell lines with 3 replicates per cell line) CAFs in the RNAseq dataset, each point representing one sample. Significance determined by EdgeR glmFit, $FDR < 0.05$. (C) $\log_2(\text{tumour stroma/tumour epithelium})$ (TS/TE) for IDC tumours ($n=32$) and ILC tumours ($n=17$), each point representing one tumour, two-tailed Mann-Whitney test, $p=0.0003$.

3.2.4 ILC and IDC CAFs do not differ in expression of CAF subtype signatures

In breast cancer, multiple subtypes of CAFs have been identified with differing phenotypes and associations with clinical features, by expression of CAF cell-surface markers and through single-cell RNAseq analysis of mouse models and patient tumours^{101,117,118}.

Expression of CAF subtype signatures from Houthuijzen *et al*¹⁰¹ and Costa *et al*¹¹⁷ was analysed in the CAF RNAseq dataset by single-sample GSEA (ssGSEA)²⁰⁸. The two subtypes from Houthuijzen *et al*, inflammatory CAFs (iCAF) and myofibroblastic CAFs (myCAF), were defined by single cell RNAseq (scRNAseq) of fibroblasts from an ILC mouse model (*WapCre;Cdh1^{fl/fl};Pten^{fl/fl}*, WEPTn), with iCAFs representing immune modulatory CAFs and myCAFs representing more ECM modulatory CAFs. Costa *et al* defined four subtypes of CAFs (S1-S4) by analysis of CAF surface markers in Luminal and Triple Negative (TNBC) tumour samples by FACs, with S1 and S4 representing tumour-infiltrated myofibroblasts. S1 CAFs were shown to be immunosuppressive whereas S4 CAFs were enriched for cytoskeletal and oxidative phosphorylation gene sets.

Looking at expression of iCAF and myCAF signatures (Figure 3-4A) and also S1 and S4 signatures (Figure 3-4B), showed that ILC and IDC CAFs do not separate out, and many cell lines show either high expression of both signatures or low expression of both signatures rather than mapping to a distinct population. This suggests that the cell lines represent heterogeneous populations, with samples potentially containing multiple CAF subtypes. Interestingly, ED2263, the IDC CAF cell line that clustered separately from other cell lines in the PCA plots, has a very high expression of the CAF-S4 gene signature but low expression of the CAF-S1 signature whereas all other IDC CAF lines had high expression of both the CAF-S1 and S4 signatures (Figure 3-4B). This suggests that ED2263 is dominated by CAFs in the S4 subtype rather than a heterogeneous population of both and may explain why this cell line is distinct from other IDC CAFs. No significant differences in expression of signatures were found between ILC and IDC CAF for any of the signatures (Figure 3-4C), although ILC CAFs generally had a lower score for all four subtype signatures.

Unbiased hierarchical clustering of the CAF cell lines based on ssGSEA subtype scores separated the cell lines into three clusters: CAF-S4 high, myCAF/S1 high and iCAF/S1 high. There is some overlap in genes in the CAF-S1 signature with genes in both the iCAF and myCAF signatures (see Appendix 4), which could explain why it was hard to separate these out. Labelling of the PCA plot with these clusters showed that expression of these signatures could explain some of the differences in dimension 1, with S4-high CAF cell lines generally clustering separately from the myCAF/S1 and iCAF/S1 high CAFs, except for ED2327 samples, which had particularly high expression of all four signatures, and ED9B which also had relatively high CAF-S1 signature expression.

Overall, there was a high variability of scores for each CAF subtype signature, demonstrating the heterogeneity between these cell lines. The fact that many of the cell lines show high or low expression of multiple of these gene signatures suggests that multiple CAF subtypes and phenotypes may be present within each cell line. To investigate this further, more in depth analyses should be carried out, such as spatial transcriptomics or scRNAseq of freshly isolated CAFs from ILC and IDC tumours.

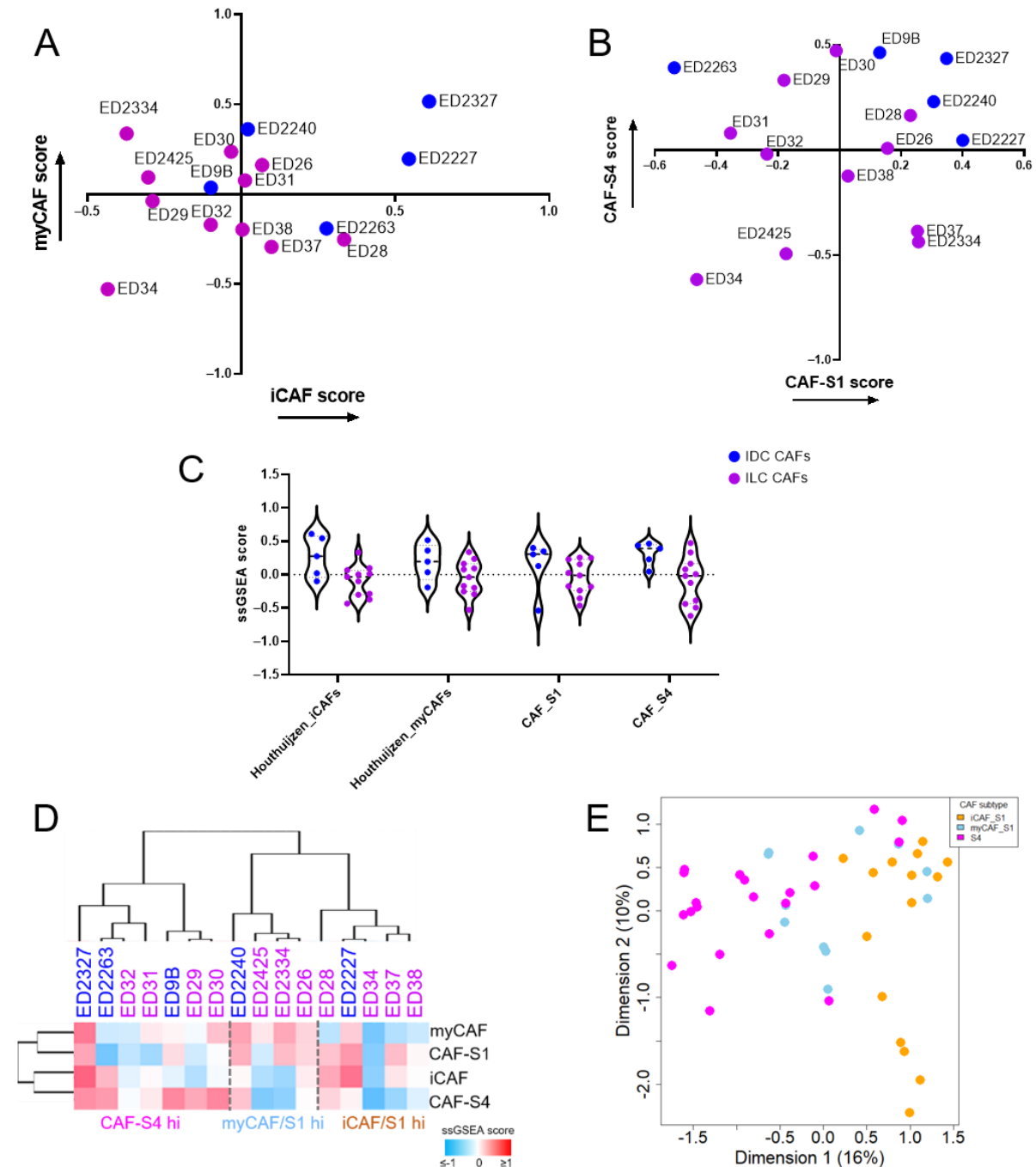


Figure 3-4. CAF subtype signatures in ILC and IDC CAFs. Scores for expression of CAF subtype signatures were generated using ssGSEA in the GSVA package in R²⁰⁸. Co-ordinate plots of subtype scores for (A) iCAF and myCAF ILC mouse model signatures from Houthuijzen *et al*¹⁰¹ and (B) human breast cancer CAF S1 and S4 signatures from Costa *et al*¹¹⁷. (C) Expression of the signatures in IDC CAFs (blue) and ILC CAFs (purple). There was no significant difference in expression of signatures between ILC and IDC CAFs by multiple Mann-Whitney tests with correction by the Benjamini, Krieger and Yekutieli two-stage step-up method (BKY) in Graphpad Prism. (D) Unbiased hierarchical clustering of CAF subtype ssGSEA scores in Cluster 3.0 and Java Treeview. (E) PCA plot of CAF RNAseq data (as seen in Figure 3-1) labelled by clusters determined in (D).

3.2.5 Integrating ILC and IDC CAF RNAseq data with laser-capture microdissection datasets

Previous work in the lab by Laura Gómez-Cuadrado identified genes specifically enriched in the stroma of ILC tumours compared to IDC tumours²¹⁶ and normal breast²³⁰ using a laser-capture microdissection (LCM) analysis of 17 ER+ ILC tumours¹³. This identified a total of 261 genes specifically enriched in the stroma of ILC tumours. Of these, 186 were expressed by primary ILC and IDC CAFs in the RNAseq dataset and 27 were differentially expressed between them before correction, with 10 genes more highly expressed in ILC CAFs and 17 in IDC CAFs (Figure 3-5A and Table 3-3). Network analysis using Genemania in Cytoscape identified interactions between 18 of these genes by known genetic and physical interactions (Figure 3-5B).

Investigating expression of these 27 DEGs in ER+ ILC and ER+ IDC tumours in the TCGA and METABRIC datasets showed that 20 genes were significantly more highly expressed in ILC tumours, including 6 that were also significantly more highly expressed in ILC CAFs (Table 3-3). This analysis again highlighted *CECR1* and *MYH3* but also identified *AASS*, *AKT3*, *EDA2R* and *MBP* as genes significantly enriched in isolated ILC CAFs, bulk ILC tumours and the ILC stroma.

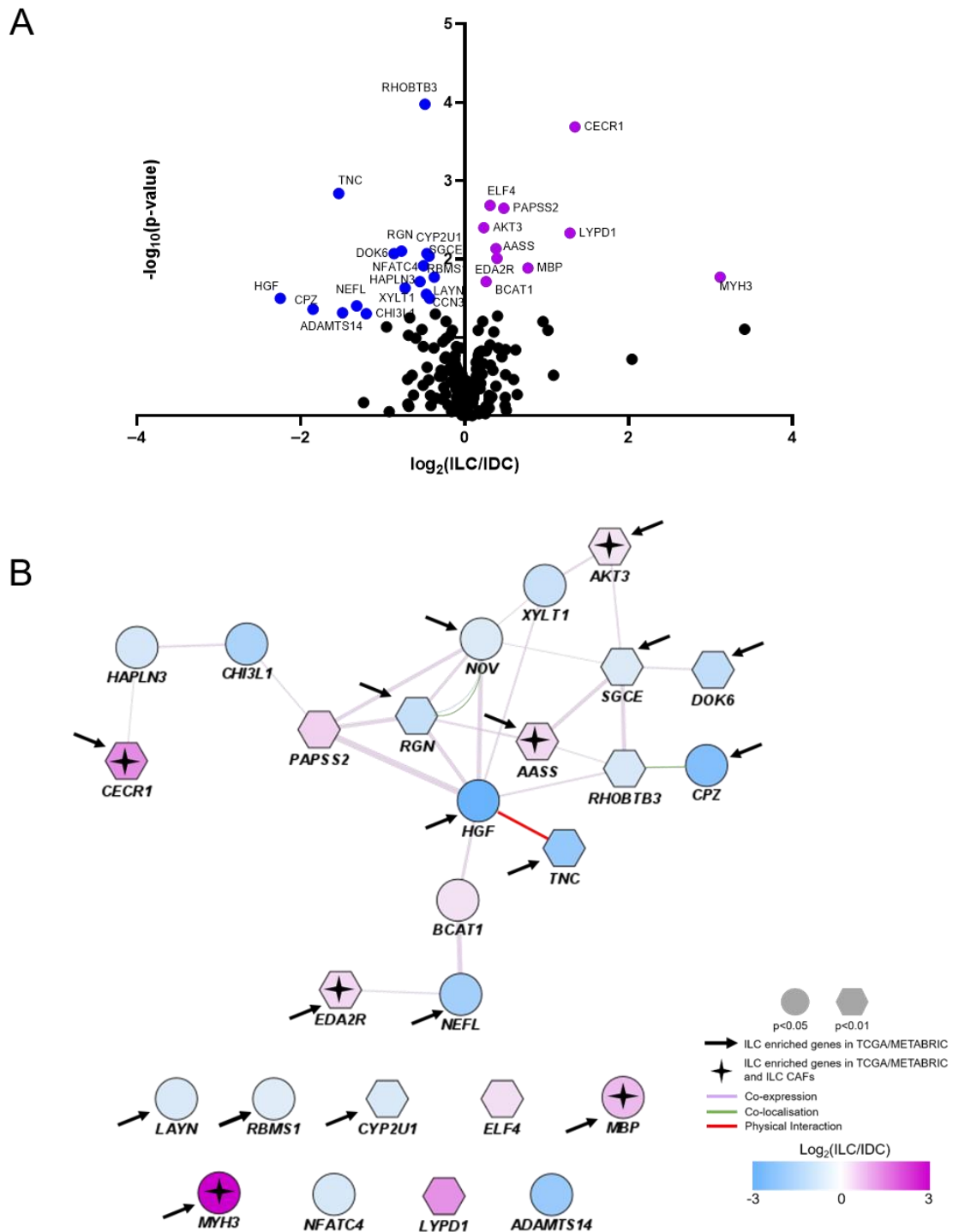


Figure 3-5. Expression of ILC stroma enriched genes identified by LCM analysis in the ILC and IDC CAF RNAseq dataset. (A) Volcano plot showing differential expression of the 176 ILC stroma-enriched genes expressed above threshold in the CAF RNAseq dataset. Genes significantly upregulated in ILC CAFs are labelled in purple ($\log_2(\text{ILC/IDC}) > 0$), in IDC CAFs in blue (Mann Whitney U test, $p\text{-value} < 0.05$). (B) Network of the 27 differentially expressed genes, genes upregulated in ILC CAFs in purple ($\log_2\text{FC} > 0$) and upregulated in IDC CAFs in blue ($\log_2\text{FC} < 0$), circles $p < 0.05$, hexagons $p < 0.01$. Arrows show genes significantly upregulated in ER⁺ ILC tumours compared to ER⁺ IDC tumours in the TCGA and/or METABRIC datasets and crosses indicate genes up in ILC CAFs and tumours. Network made using Genemania in Cytoscape, showing interactions due to co-expression (purple lines), co-localisation (green) and physical interaction (red), adapted from¹³

	CAF RNAseq		LCM datasets Log ₂ (TS/TE)		METABRIC and TCGA ER+ ILC v ER+ IDC		
Genes	Log ₂ (ILC/IDC)	p-value	ILC	IDC	METABRIC	TCGA microarray	TCGA RNA-Seq
<i>RHOBTB3</i>	-0.49	0.000	0.79	-0.15	ns	ns	IDC (***)
<i>CECR1</i>	1.35	0.000	1.59	0.14	ns	ILC (*)	ILC (*)
<i>TNC</i>	-1.54	0.001	1.07	-0.01	ns	ILC (*)	ILC (*)
<i>ELF4</i>	0.31	0.002	0.95	-1.53	ns	ns	IDC (*)
<i>PAPSS2</i>	0.48	0.002	1.06	0.12	NA	ns	ns
<i>AKT3</i>	0.23	0.004	0.64	0.13	ns	ILC (**)	ILC (***)
<i>LYPD1</i>	1.28	0.005	0.91	0.03	IDC (**)	ns	ns
<i>AASS</i>	0.38	0.007	0.77	0.12	ILC (**)	ILC (***)	ILC (***)
<i>RGN</i>	-0.77	0.008	0.94	0.34	ILC (*)	ILC (**)	ILC (***)
<i>CYP2U1</i>	-0.46	0.009	1.20	0.07	ILC (**)	ILC (*)	ILC (**)
<i>DOK6</i>	-0.86	0.009	0.70	-0.01	ILC (***)	ILC (**)	ILC (***)
<i>SGCE</i>	-0.44	0.009	0.72	0.11	ILC (**)	ILC (**)	ILC (***)
<i>EDA2R</i>	0.39	0.01	0.73	0.13	ns	ns	ILC (**)
<i>NFATC4</i>	-0.5	0.012	1.16	-0.21	NA	ns	ILC (**)
<i>MBP</i>	0.77	0.013	0.83	0.01	ILC (*)	ILC (*)	ILC (**)
<i>RBMS1</i>	-0.37	0.017	0.82	0.03	ns	ILC (**)	ILC (***)
<i>MYH3</i>	3.12	0.017	0.86	-0.01	ns	ILC (*)	ILC (***)
<i>HAPLN3</i>	-0.55	0.019	0.81	-0.04	ns	ns	ILC (***)
<i>BCAT1</i>	0.26	0.019	0.89	0.15	IDC (***)	ns	IDC (**)
<i>XYLT1</i>	-0.73	0.023	0.97	0.07	IDC (*)	ns	ns
<i>LAYN</i>	-0.47	0.028	1.56	0.05	ILC (*)	ILC (*)	ILC (***)
<i>HGF</i>	-2.25	0.032	1.85	-0.13	ns	ILC (**)	ILC (***)
<i>NOV</i>	-0.43	0.032	1.08	0.02	ns	ILC (***)	ILC (***)
<i>NEFL</i>	-1.32	0.039	0.76	0.14	ns	ILC (***)	ILC (***)
<i>CPZ</i>	-1.85	0.044	0.51	-0.01	ILC (**)	ILC (*)	ILC (***)
<i>ADAMTS14</i>	-1.49	0.048	1.83	0.13	IDC (**)	ns	ns
<i>CHI3L1</i>	-1.2	0.05	0.50	0.45	NA	ns	ILC (***)

Table 3-3. 27 ILC stroma-enriched genes differentially expressed between ILC and IDC CAFs.

Log₂ fold change for expression in ILC CAFs and IDC CAFs, Log₂(ILC/IDC)>0, up in ILC CAFs, <0, up in IDC CAFs. P-value represents the results of Mann Whitney U test before correction. Log₂ fold change between tumour stroma (TS) and epithelium (TE) in ILC (n=17)¹³ and IDC²¹⁶ (n=32). For expression of these genes in ER+ ILC and ER+ IDC tumours in METABRIC (ILC n=125, IDC n=1116) and TCGA datasets (microarray ILC n=40, IDC n=344, RNAseq ILC n=191, IDC n=555), ILC or IDC indicates significantly higher expression in that subtype with level of significance in brackets. METABRIC and TCGA statistics by Mann Whitney U test with BH correction, adjusted p<0.05 *, p<0.01 **, p<0.001 ***, ns = not significant, NA=gene not in dataset. Genes highlighted in purple are upregulated in ILC CAFs, ILC stroma and ILC tumours in METABRIC and/or TCGA datasets. Adapted from¹³

3.2.6 ILC and IDC CAF secretome

Paracrine signalling via secreted proteins is an essential method of communication between the stroma and tumour cells and secreted proteins represent accessible potential therapeutic targets as well as potential biomarkers. To investigate the secretome of ILC and IDC CAFs, the RNAseq dataset was limited to 1571 genes encoding for proteins annotated to be secreted, extracted from the Gene Ontology Slim database²⁰⁶. This had little impact on the distribution of samples in the PCA plot (Figure 3-6A), with samples from the same cell lines still clustering closely together and in the similar regions as in the full dataset, with ED2263 samples still clustering relatively separately from the other samples.

Differential expression analysis in this limited dataset identified 37 DEGs between ILC and IDC CAFs, with 7 of these genes upregulated in ILC CAFs (Figure 3-6B and Table 3-4). Genetic and co-expression interactions were detected between *MMP16* and *IL1A*, and *CECR1* was shown to have a genetic interaction with *ADAMTS16* as well as a physical interaction with *IL1A* by network analysis (Figure 3-5C). Of the 7 genes upregulated in ILC CAFs, again, only *CECR1* was significantly enriched in the stroma of ILC tumours¹³ and *CECR1* and *ADAMTS16* were significantly more highly expressed in ILC tumours than IDC tumours in the TCGA dataset.

	CAF RNA-Seq data		METABRIC and TCGA ER+ ILC v ER+ IDC		
Genes	Log ₂ (ILC/IDC)	FDR	METABRIC	TCGA microarray	TCGA RNAseq
<i>NPW</i>	2.8114	0.0099	NA	ILC*	ns
<i>IL1A</i>	2.1933	0.0193	ns	ns	ns
<i>MMP16</i>	1.6798	0.0284	ns	ns	ns
<i>CHGA</i>	4.6764	0.0319	ns	ns	IDC****
<i>ADAMTS16</i>	3.0412	0.0466	ns	ILC*	ILC****
<i>ADA2</i>	1.3518	0.0495	ns	ILC*	ILC**
<i>AREG</i>	4.0835	0.0495	ns	ns	IDC

Table 3-4. Expression of genes significantly upregulated in ILC CAFs encoding for secreted proteins. Significantly enriched secreted protein genes in the CAF RNAseq dataset with log₂(expression in ILC CAFs/IDC CAFs) (glmFit, FDR<0.05). Expression of these genes in the METABRIC (ILC n=125, IDC n=1125) and TCGA datasets (microarray ILC n=40, IDC n=344, RNAseq ILC n=191, IDC n=563). ILC or IDC indicates significantly higher expression in that subtype with level of significance in brackets, Mann Whitney U test with BH correction, adjusted p<0.05 *, p<0.01 **, p<0.001 ***, ns = not significant, NA=gene not in dataset. Adapted from¹³



Figure 3-6. Analysis of genes encoding secreted proteins in the CAF RNAseq dataset. (A) PCA plot of CAF RNAseq samples after filtering dataset to only include 1571 genes encoding for secreted proteins, labelled by cell line. (B) Unbiased hierarchical clustering of differentially expressed secretome genes between ILC and IDC CAFs (glmFit, FDR<0.05), made using Cluster 3.0 and Treeview. (C) Genemania network of differentially expressed secretome genes, made using Genemania in Cytoscape accessed November 2023. Genes upregulated in ILC CAFs in purple, in IDC CAFs in blue, arrows indicate significantly more highly expressed in ILC tumours in TCGA and/or METABRIC datasets and crosses show genes also upregulated in ILC CAFs. Purple lines indicate interactions by co-expression, green lines by genetic interactions and red by physical interactions. (B and C) adapted from¹³

To further investigate the secretome of ILC and IDC CAFs, conditioned media (CM) was collected from four ILC CAF lines (ED26, ED28, ED34 and ED2334) and two IDC CAF lines (ED9B and ED2227) and submitted for forward phase protein analysis (FPPA). 42 out of the 64 antibodies in the assay detected protein above background level (Figure 3-7) and of these, no proteins were significantly differentially expressed between ILC and IDC CAFs after correction (multiple Welch's t-tests with BKY correction in Graphpad Prism). However, 9 proteins had a more than two-fold higher expression in ILC CAF CM than IDC CAF CM, including IL-1 α and Amphiregulin which are also significantly more highly expressed in ILC CAFs at the transcriptional level (Table 3-2 and Figure 3-6). Before correction, Osteopontin expression was also significantly higher in ILC CAF CM (Table 3-5).

The top 20 most highly expressed proteins in the FPPA dataset (Table 3-5) include four IGF binding proteins and IGF-1. The IGF-1 pathway has been shown to be very important for driving proliferation and tumour progression in breast cancer, with a number of groups identifying this pathway as particularly important in ILC^{52,222}. Work by Laura Gómez-Cuadrado has also identified a paracrine IGF1/PAPP-A/IGFBP axis in ILC¹³ and this FPPA data further highlights the importance of CAFs in regulating IGF1 signalling. This is explored in more detail in Chapter 6.

14/20 of the genes (*IGFBP5*, *TIMP1*, *CCL1*, *DPP4*, *CXCL16*, *CCL5*, *IL6*, *CCL11*, *IGF1*, *CCN3*, *MMP3*, *CSF2*, *VEGFA*, *IGFBP5*) encoding these proteins are involved in 'regulation of cell migration' (GO biological process, GO:0030334) and more specifically, *CXCL1*, *IL6*, *SPP1*, *CCL1*, *DPP4*, *PF4*, *CXCL16*, *CCL5*, *CCL11*, *CCN3* and *VEGFA* are involved in 'leukocyte chemotaxis'. This suggests that these CAFs may regulate both tumour cell migration and the immune system, although this gene ontology analysis is limited and biased by the selection of antibodies used in the FPPA. The majority of these genes were significantly more highly expressed in ILC tumours than IDC tumours in the TCGA dataset (Table 3-5), despite not being differentially expressed between ILC and IDC CAFs at the protein or gene expression level. This again may be due to the higher stromal content in ILC tumours than in IDC tumours⁶.

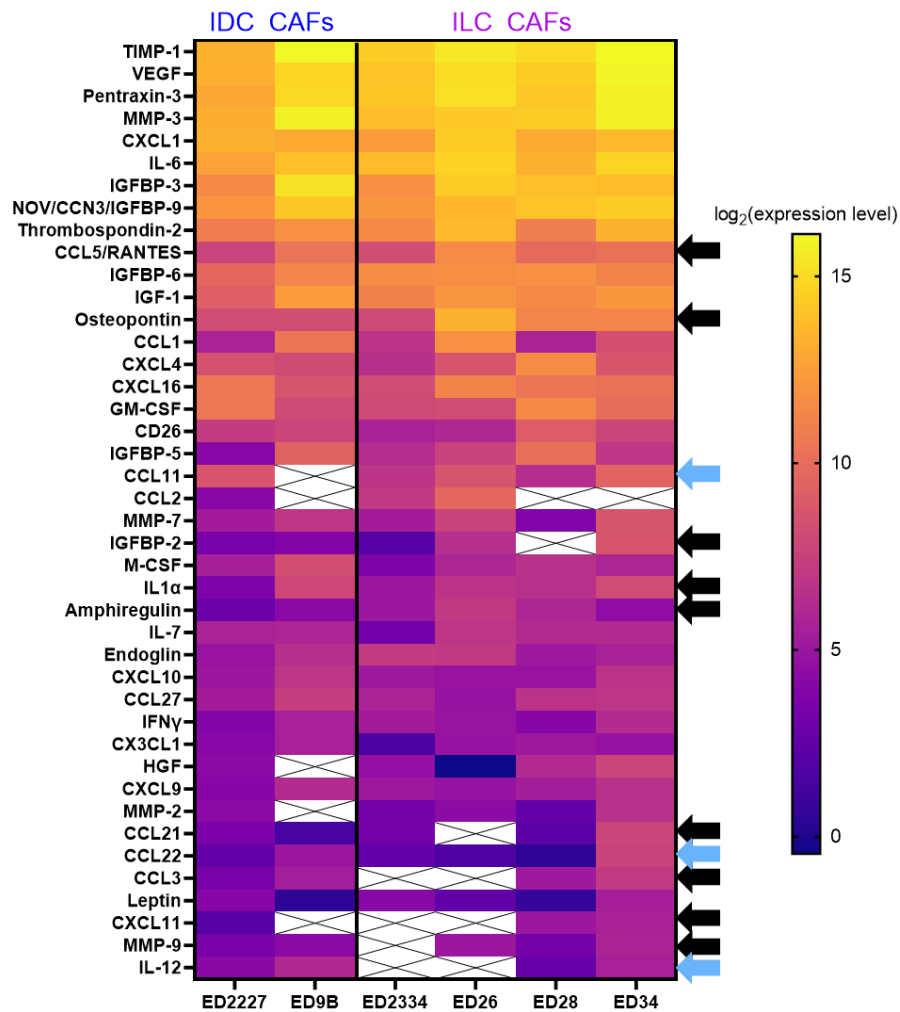


Figure 3-7. ILC and IDC CAF secretome analysed by FPPA. Heatmap showing $\log_2(\text{expression level})$ of the 42/64 secreted proteins expressed above background in CAF conditioned media in 2 IDC CAF lines (ED2227 and ED9B) and 4 ILC CAF lines (ED2334, ED26, ED28 and ED34). Ordered by mean expression level across ILC CAFs. Empty boxes with crosses indicate no expression above background in that cell line, arrows indicate genes with more than two fold higher expression in ILC CAF CM (black) or in IDC CAF CM (blue). N=3 biological replicates for each cell line, except ED34 and ED9B for which n=2 due to low concentrations of proteins in final replicates.

		FPPA data		CAF RNA-seq dataset		TCGA RNAseq
Protein	Gene	Log ₂ (FC)	p-value	Log ₂ (FC)	FDR	ER+ ILC v ER+ IDC
TIMP-1	<i>TIMP1</i>	0.32	0.73	0.12	0.85	ILC (**)
VEGF	<i>VEGFA</i>	0.69	0.37	-0.38	0.54	IDC (**)
Pentraxin-3	<i>PTX3</i>	0.72	0.29	0.3	0.85	ILC (****)
MMP-3	<i>MMP3</i>	-0.03	0.98	-1.15	0.49	ILC (****)
CXCL1	<i>CXCL1</i>	0.26	0.79	0.88	0.69	ILC (****)
IL-6	<i>IL6</i>	0.88	0.19	-0.49	0.56	ILC (****)
IGFBP-3	<i>IGFBP3</i>	0.05	0.96	-0.7	0.44	ILC (**)
NOV/CCN3/IGFBP-9	<i>CCN3</i>	0.23	0.76	-0.43	0.63	ILC (****)
Thrombospondin-2	<i>THBS2</i>	0.90	0.25	0.66	0.43	ns
CCL5/RANTES	<i>CCL5</i>	1.40	0.26	-0.69	0.84	ILC (****)
IGFBP-6	<i>IGFBP6</i>	0.99	0.21	-0.2	0.59	ILC (****)
IGF-1	<i>IGF1</i>	0.42	0.68	-1.61	0.11	ILC (****)
Osteopontin	<i>SPP1</i>	3.36	0.01	1.61	0.18	IDC (****)
CCL1	<i>CCL1</i>	0.80	0.60	NA	NA	ns
CXCL4	<i>PF4</i>	0.22	0.83	1.32	0.49	ILC (**)
CXCL16	<i>CXCL16</i>	0.31	0.75	-0.09	0.96	ILC (****)
GM-CSF	<i>CSF2</i>	-0.53	0.53	NA	NA	ns
CD26	<i>DPP4</i>	-0.33	0.76	-0.26	0.78	ILC (****)
IGFBP-5	<i>IGFBP5</i>	0.42	0.79	-0.45	0.77	ILC (*)
CCL11	<i>CCL11</i>	-1.63	0.22	-1.11	0.66	ns

Table 3-5. Expression of the 20 most highly expressed proteins in ILC CAF conditioned media.

All log₂(FC) = log₂ fold change of ILC/IDC samples. FPPA data - Log₂(mean expression) of ILC (n=4 cell lines, 3 repeats (except ED34, 2 repeats)) and IDC CAFs (n=2 cell lines, 3 repeats of ED2227, 2 repeats ED9B) is shown and the p-value from multiple unpaired Welch's t-tests before correction performed in Graphpad Prism. The ILC and IDC CAF RNA-seq data shows Log₂(ILC/IDC) expression and FDR calculated by glmFit likelihood ratio test in EdgeR. Expression of the genes encoding these proteins was assessed in ER+ ILC (n=191) and ER+ IDC (n=563) tumours in the TCGA RNAseq dataset, with 'ILC' indicating genes significantly enriched in ILC tumours, 'IDC' – genes enriched in IDC tumours. Significance level shown in brackets with FDR calculated by multiple unpaired Welch's t-tests in Graphpad Prism with BKY correction, FDR<0.05 *, FDR<0.01 **, FDR<0.0001 ****.

Chapter 3.3 Discussion

The purpose of this work was to investigate differences between ILC and IDC CAFs and to complement the findings of the laser-capture microdissection analysis of ILC tumours¹³. PCA plots and unbiased hierarchical clustering of the most variable genes did not clearly separate the ILC and IDC CAFs from each other. In many ways, the lack of separation between the ILC and IDC CAFs is to be expected, as they are the same cell type isolated from ER+ breast cancer and CAF signatures have been shown to be conserved across breast cancer subtypes^{101,117}. However, 153 differentially expressed genes were identified between ILC and IDC CAFs with a number validated as enriched in the stroma of ILC tumours relative to IDC tumours in the LCM and bulk tumour datasets.

Although we fulfilled the aims of this project, this analysis is limited in that we are using bulk RNAseq analysis of cultured CAFs and so we lose any information on intra-cell line heterogeneity, reciprocal signalling from the tumour epithelium and spatial heterogeneity. Whilst culturing these CAFs provides a relevant and useful model for studying ILC biology *in vitro*, some studies have shown extended culturing of CAFs can alter their behaviour, transforming them towards an iCAF phenotype²¹⁹. We have tried to limit the effect of this by only using CAFs at low passage. Validation of this data using our LCM dataset as well as publicly available bulk RNAseq datasets allows us to corroborate our findings but further studies analysing primary CAFs immediately after isolation from ILC tumours by scRNAseq or use of spatial transcriptomics may provide further in depth analysis and information on interactions with other cell types in the TME of ILC.

3.3.1 The prognostic role of the stroma in ILC

ILC tumours have been shown to retain CD34+ fibroblasts which are not found in IDC tumours¹⁴², though one study detected no CD34+ fibroblasts in the stroma of 145 luminal ILC tumours by IHC⁵⁵. The presence of CD34+ fibroblasts associates with improved overall and disease-free survival in ILC¹⁴³. CD34+ fibroblasts also negatively correlate with α -SMA+ CAFs in the TME of ILC tumours, potentially suggesting poorer outcomes for ILC patients with increased numbers of α -SMA+ CAFs, which has been found in IDC tumours but not investigated in ILC²³¹. CD34 was not detected above threshold in any of our CAF lines in the RNAseq dataset and was

not enriched in the stroma of ILC tumours in the LCM dataset either (data not shown). However, our CAF lines could potentially contain a sub-population of CD34+ fibroblasts that the bulk RNAseq analysis did not have the resolution to detect.

A recent study by Forsare *et al* demonstrated that high stromal content was associated with a good prognosis in univariate analyses of ILC patients at 10 years follow-up¹⁴¹, though no significance was found in multivariate analyses or at 25 years follow-up. In addition, ILC tumours in the ‘reactive-like’ subtype characterised by high stromal content were shown to have a better outcome than those in the proliferative subtype³. High expression of two ILC-stroma enriched genes identified by the LCM analysis, *PAPPA* and *TIMP2*, is also associated with improved outcomes in ILC tumours¹³. However, in the Forsare *et al* study, stromal content was estimated by eye by pathologists using H&E stained tumour cores rather than specific staining for cell types, and the high stromal content of ‘reactive-like’ ILC tumours is based on bulk RNAseq analysis. Similarly the survival analysis for ILC-stroma enriched genes was based on bulk RNAseq data rather than investigating stroma-specific expression of these genes. The association of specific CAF subtypes or CAF expressed genes in the stromal compartment of ILC with clinical outcomes should be further investigated.

3.3.2 ILC stroma and CAF enriched genes

Six genes (*AASS*, *AKT3*, *CECR1*, *EDA2R*, *MBP* and *MYH3*) were found to be significantly enriched in the stroma of ILC tumours in the LCM dataset as well as more highly expressed in ILC CAFs and ILC tumours (Figure 3-5). *AASS*, *MBP* and *MYH3* encode for a lysine degrading enzyme, myelin basic protein and myosin heavy chain 3, respectively. To date, none of these genes have been implicated in the progression of breast cancer. In contrast, *EDA2R* encodes for a TNF α super family receptor and has been proposed as a tumour suppressor in breast cancer as its expression is reduced in breast cancer cell lines and tumours and overexpression induces apoptosis²³². However, its role in fibroblasts has not been explored.

CECR1, also known as *ADA2*, encodes for a secreted adenosine deaminase isoenzyme and was significantly differentially expressed between ILC and IDC CAFs in the complete RNAseq dataset. High ADA2 activity has previously been reported in macrophages and endothelial cells^{233–235}, whilst breast cancer cells had lower ADA2 activity. This work demonstrates that CAFs in the TME of ILC can also provide a

source of ADA2, although validation of ADA2 protein expression and activity is needed. ADA2 activity has been shown to promote the polarisation of macrophages towards the pro-tumorigenic M2-phenotype²³³ and *CECR1* expression significantly correlates with M2-macrophage signatures in breast cancer and glioma^{233,236}. The high stromal expression of *CECR1* in ILC tumours could contribute to the higher level of infiltrated M2-macrophages found in ILC tumours¹⁴.

Conversely, high expression of *CECR1* has generally been shown to be associated with better outcome in many cancer types including breast cancer and to negatively correlate with immunosuppressive myeloid-derived suppressor cells and positively correlate with cytotoxic T-cells²³⁷. However, this analysis did not distinguish between subtypes of breast cancer and *CECR1* was found to be significantly associated with poorer survival in ILC patients in the SCANB dataset, though no impact on survival was seen in TCGA and METABRIC datasets¹³. ADA2 activity in plasma also increases with grade of triple-negative breast cancer, although this trend was not seen in ER+ breast cancers²³⁶. The consistency of elevated *CECR1* expression in ILC tumours and CAFs highlights this gene as a potentially interesting ILC-enriched gene. However, the conflicting data on ADA2 as either pro-tumorigenic and immunosuppressive or anti-tumorigenic and promoting immune infiltration into tumours highlights the need for work interrogating the role of TME derived ADA2 activity in ILC.

Significantly higher expression of *AKT3* was also seen in ILC CAFs and the stroma of ILC tumours relative to IDC tumours. *AKT3* is a downstream Ser/Thr kinase of the PI3K pathway and its expression in CAFs in head and neck squamous cell carcinoma promotes expression of pro-inflammatory cytokines and is associated with worse OS²³⁸, though its role in breast CAFs has not been investigated. Furthermore, ILC CAFs were also significantly enriched for mTORC1 pathway gene sets. mTORC1 activity in CAFs has also previously been shown to promote CAF secretion of pro-inflammatory and pro-migratory cytokines^{239,240}, such as CCL5 which had a more than two-fold higher expression in CM from ILC CAFs than IDC CAFs in the FPPA dataset. The FPPA data also demonstrated high expression of IGF-1 and IGF-1 binding proteins in CAF CM, all of which are significantly more highly expressed in ILC tumours (Table 3-5).

The PI3K/AKT/mTOR pathway is hyper-activated in ILC tumours, due to increased autocrine^{52,53} and paracrine¹³ signalling of activating growth factors, such as IGF-1, and 48% of ILC tumours in the TCGA dataset have activating PIK3CA mutations³. The LCM analysis also identified *PAPPA* as specifically upregulated in the stroma of ILC¹³. PAPP-A is a proteinase that cleaves IGF binding proteins increasing IGF-1 local bioavailability and increasing IGF-1 driven activation of the PI3K pathway in ILC cells *in vitro*. This suggests that targeting both the tumour epithelium and stroma with IGF-1/PI3K/mTORC targeted therapeutics may offer clinical benefit to ILC patients. However, IGF-1 targeted therapies have had poor results in trials (reviewed in⁵⁶) and high expression of *PAPPA* was shown to associate with better outcomes for ILC patients¹³. This highlights the need for more research to fully identify useful biomarkers and targets of this pathway in ILC.

3.3.3 *ESR1* and *GPER1* expression in CAFs

IDC CAFs were shown to express higher levels of both *ESR1* and *GPER1* than ILC CAFs in the RNAseq dataset, with a higher stroma to epithelial ratio of *ESR1* seen in IDC tumours than ILC tumours in the laser capture microdissection datasets (Figure 3-3). Estrogen stimulation of IDC CAFs has been shown to promote angiogenesis, dependent on GPER1 promoting HIF1 α activity and subsequent VEGF expression^{223–225}. Interestingly, IDC CAFs were significantly enriched for the Hallmark Hypoxia gene set (Figure 3-2D), which could potentially be due to increased GPER1 expression in these CAFs, though experimental validation is needed.

Generally, ER α expression (encoded by *ESR1*) is thought to be restricted to the epithelial compartment of breast tumours. However, low *ESR1* expression was detected in the RNAseq analysis and ER α has been detected in the stroma of IDC tumours by IHC²²⁶. ER α expression in prostate CAFs has been shown to promote tumour cell proliferation *in vitro*²²⁷, but to reduce infiltration of M2-like macrophages and subsequent tumour cell invasion and metastasis *in vivo*²²⁸. However, in gastric cancer, estrogen stimulation of ER α + CAFs promoted increased migration *in vitro* via upregulation of CD147, MMP2 and MMP9²⁴¹, suggesting diverging roles of ER α + CAFs in different cancer types. The role of ER α + CAFs in breast cancer has not been investigated to date.

Expression of *GPER1* and *ESR1*, in particular, is quite low in our RNAseq dataset. Validation of both ER α and GPER1 expression at the protein level in CAFs, either by immunofluorescence or western blot, is needed to validate these findings and further work is needed to determine any functional consequences. Analysis and comparison of ER α and GPER1 protein expression in the stroma of ILC and IDC tumours and correlation with clinical outcomes would also be useful. However, an analysis of GPER1 expression across 209 ILC tumours showed no association with breast cancer specific survival and no stromal staining²⁴².

3.3.4 CAF secreted proteins

Very few significant differences were found in CAF CM proteins secreted by ILC and IDC CAFs in the FPPA analysis as a limited number of cell lines and antibodies were used for this experiment, reducing the statistical power of these experiments. As only two IDC CAF lines were included in the FPPA, it may reduce the validity of statistical analysis, as there may be high levels of variation between patient samples, independent of breast cancer subtype. However, this data gives a descriptive overview of the secretome of ILC and IDC CAFs and, in combination with analysis of the CAF RNAseq and TCGA datasets, allows us to draw some tentative conclusions and provides a useful dataset for *in vitro* work.

Of the most highly expressed proteins by ILC CAFs, the vast majority were involved in cell migration and leukocyte chemotaxis, though this analysis was biased by our selection of antibodies. However, this may suggest that CAF secreted proteins could regulate both tumour cell migration as well as immune cell infiltration into ILC tumours. Immune related gene signatures are more highly expressed in ILC than IDC tumours³⁹, although another study has shown that ILC has fewer tumour infiltrating lymphocytes¹⁵. Further investigation of the role of these proteins and CAFs in the immune microenvironment of ILC is necessary. Analysis of CAF CM from a subset of our ILC and IDC CAFs by mass spectrometry is ongoing in collaboration with the Poli group at the University of Turin, allowing an unbiased view of CAF secreted proteins within the conditioned media.

Osteopontin (*SPP1*) was the only protein significantly more highly expressed in ILC CAF CM than in IDC CAF CM before correction (Table 3-5). It was also more highly expressed at the transcript level in ILC CAFs, though not significantly. Osteopontin is

a secreted phosphoprotein recognised by CD44 and integrin receptors and high serum levels have been associated with poorer outcomes in metastatic breast cancer²⁴³. Interestingly, *SPP1* is part of a stromal signature that predicts poor outcome in IDC patients²³⁰. Osteopontin expression by breast cancer cells can lead to activation of fibroblasts and mesenchymal stem cells, converting them to pro-migratory inflammatory CAFs in a paracrine manner and promoting increased tumour growth and metastasis *in vivo*^{107,110,111}. CAF secreted osteopontin has also been shown to increase 3D-growth of the IDC MCF7 cell line²³¹. It would be useful to assess the role of CAF-secreted osteopontin on the phenotypes of both CAFs and ILC tumour cells in future.

IL1A and *AREG* are significantly more highly expressed by ILC CAFs (Table 3-2) and had a more than two fold higher expression in the ILC CAF CM than IDC CAF CM (Figure 3-7). However, neither are enriched in the ILC stroma in LCM datasets, suggesting high tumour cell expression of these genes. Amphiregulin (AREG) is an ER α regulated gene product and, once cleaved from its transmembrane domain, behaves as an epidermal growth factor receptor ligand²⁴⁴. Autocrine AREG activity in breast cancer cells has been shown to promote aromatase inhibitor resistance^{245,246} and so paracrine signalling from CAFs may also contribute to acquired aromatase inhibitor resistance in ILC. However, *AREG* was very highly expressed by just one ILC CAF line, ED34, and was not upregulated in ILC tumours or the stroma so may not be representative of all ILC CAFs.

IL-1 α stimulation of normal gastric fibroblasts was shown to promote CAF-like behaviour in these cells²⁴⁷ and paracrine IL-1 α signalling from pancreatic ductal adenocarcinoma (PDAC) tumour cells to pancreatic stellate cells converts these to an iCAF phenotype, with increased expression of pro-inflammatory cytokines such as IL-6²⁴⁸. This suggests that ILC CAF expression of IL-1 α could perpetuate the formation of iCAFs during ILC tumour progression and promote expression of key mediators of CAF-tumour crosstalk via NF κ B pathway activation, such as IL-6. However, despite the higher expression of *IL1A* in ILC CAFs, IDC CAFs were significantly enriched for the 'TNF α signalling via NF κ B' gene set and TNF α signalling has also been shown to promote iCAF formation *in vivo* in PDAC models²⁴⁸. This suggests the NF κ B pathway may be important in CAFs from both ILC and IDC tumours.

TIMP1, a tissue inhibitor of matrix-metalloproteases, was the most highly secreted protein across all ILC and IDC CAFs. The most well characterised role of TIMP1 is in inhibition of matrix metalloproteinases, generally considered to be a tumour suppressive function²⁴⁹. However, recent work has demonstrated that CAF secreted TIMP1 binding to CD63 on the surface of CAFs or tumour cells can also drive STAT3 activation^{134,135}. This was shown to promote tumour cell growth and migration *in vitro* and TIMP-1 activation of STAT3 in CAFs drives a feed-forward loop, upregulating CD63 expression. TIMP-1 activated STAT3 in CAFs also upregulated expression of miR-22, which targets and downregulates ER α and drives endocrine resistance in tumour cells, as well as upregulating expression of IL-6.

IL-6 was one of the most highly expressed proteins in the FPPA analysis of ILC and IDC CAF CM and is a pleiotropic cytokine that has been shown to be instrumental in CAF-tumour cell crosstalk in many cancer types^{113,127,128,250}. Although no differences were seen in *IL6* expression at the transcriptional level, there was a trend towards higher IL-6 protein in CAF CM from ILC CAFs than IDC CAFs (Table 3-5). The next chapters focus on paracrine IL-6 signalling from CAFs to ILC tumour cells and the consequences of this at the transcriptional (Chapters 4 and 5) and functional levels (Chapter 6).

3.3.5 CAF subtypes

The heterogeneous nature of CAFs means that targeting specific pro-tumorigenic CAF subtypes or CAF expressed genes or secreted proteins offers a better potential clinical approach than targeting the stroma as a whole. As these CAFs were isolated using differential centrifugation rather than a specific marker, we have not biased this analysis for a particular CAF subtype. Characterisation of each of these CAF cell lines by previous members of the lab has demonstrated that all cells are spindle-shaped and express vimentin and lack expression of epithelial markers¹³. A sub-population of each cell line also expresses α -SMA, demonstrating the heterogeneity of these patient-derived CAFs.

Whilst this work was implemented to look at differences between ILC and IDC CAFs as a whole, ssGSEA revealed a range of expression levels of CAF subtype signatures from Costa *et al*¹¹⁷ and Houthuijzen *et al*¹⁰¹ between CAF cell lines, with no significant differences in expression of these signatures between ILC and IDC CAFs (Figure 3-4).

This suggests that the CAF lines may contain mixed populations of CAF subtypes. However, as these CAFs have been cultured rather than analysed directly from tumour samples, their expression of subtype signatures may have been altered and so may not represent the CAF subtypes seen in WEPTn or human tumours.

The lack of differences in subtype signature expression between ILC and IDC CAFs (Figure 3-4C) is unsurprising as these CAF subtypes have been detected in many subtypes of breast cancer^{117,120} and in both primary and secondary sites¹¹⁹. The Houthuijzen iCAF and myCAF subtypes were also shown to overlap with PDAC iCAF and myCAF signatures¹⁰¹. Houthuijzen *et al* determined that the iCAF and myCAF signatures were enriched in the stroma of ILC and IDC tumours in our LCM datasets. Interestingly, there was a more significant enrichment for the myCAF signature than iCAF signature in ILC tumour stroma and our ILC CAFs in general tended to have stronger expression of the myCAF signature despite all showing high expression of CD26 and IL-6, a key iCAF marker.

The CAF-S1 and S4 subtypes are more predominant in TNBC and HER2+ breast cancer than luminal cases, which had highest levels of CAF-S2 and S3 subtypes¹¹⁷. Therefore CAF-S1 and S4 may not have been the best signatures to investigate in our CAFs, though gene signatures of the S2 and S3 subtypes have not been published. However, the S1 and S4 subtypes were detected in some luminal tumours, with CAF-S2 and S4 the dominant subtypes in early luminal tumours¹²⁰, while high numbers of the CAF-S1 subtype of CAFs associated with increased risk of recurrence. It would be useful to see if the three clusters of CAF subtypes defined in this analysis have any bearing on clinical outcomes.

This work primarily focused on CAFs within the TME of ILC and IDC, with the LCM analysis excluding immune components. Both the CAF-S1 and iCAF subtypes have been shown to be immunosuppressive, with CAF-S1 secreted CXCL12 promoting T-cell recruitment and CD26 driving immunosuppressive Treg formation¹¹⁷, and iCAF secreted CXCL12 recruiting monocytes¹⁰¹. IDC CAFs expressed significantly higher levels of *CXCL12* at the transcript level, however, in the FPPA data, *CXCL12* was not detected above threshold in CAF CM from both ILC and IDC CAFs. ILC tumours have lower tumour infiltrating lymphocytes than IDC tumours¹⁵ and recent scRNAseq analysis of immune cells from patient tumours has demonstrated a higher ratio of pro-

tumorigenic M2-macrophages to M1 macrophages in ILC than IDC tumours¹⁴. It would be interesting to determine how CAFs in general as well as specific subtypes are influencing the immune microenvironment of ILC.

Chapter 3.4 Conclusions

This analysis of isolated primary ILC and IDC CAFs identified a number of differentially expressed genes enriched in the stroma of ILC tumours, such as *CECR1*. We have also demonstrated an enrichment for mTORC1 signalling genes in ILC CAFs relative to IDC CAFs and high expression of the upstream growth factor, IGF-1, and a number of IGF-1 binding proteins in CAF conditioned media. The data generated by this work provides a valuable asset for the ILC research community. Further work is needed to validate the functional consequences of genes and proteins identified in this analysis and whether they contribute to the progression or suppression of ILC tumours, or could potentially be used as biomarkers or therapeutic targets.

Chapter 4 Paracrine IL-6/STAT3 signalling in Invasive Lobular Carcinoma

The Cancer-associated fibroblasts (CAFs) used in this chapter have been characterised in the previous chapter (Chapter x) and in Gómez-Cuadrado *et al*, 2022¹³. The reverse-phase protein array (RPPA) was carried out by Kenneth Macleod at the HTPU in the IGC and Lexogen 3'-mRNA-seq library preparation and sequencing was carried out at the Wellcome Trust Clinical Research Facility. I generated the samples and carried out the analysis for both of these experiments, with assistance from Henry Beetham in the RNAseq analysis.

Chapter 4.1 Introduction

Despite accounting for up to 15% of breast cancer cases, Invasive Lobular Carcinoma (ILC) is understudied and there are currently no ILC-specific treatments targeting the many unique features of this disease. The majority of previous studies on ILC have focused on the cancer cell-autonomous phenotypes of ILC, for example interrogating the consequences of E-cadherin loss and phosphatidylinositol-3-kinase (PI3K) pathway hyper-activation^{52,222} or the inherent differences between ILC and IDC cell lines, at the genomic and transcriptomic level^{251–253}. However, a key feature of ILC is the high stromal content and density of cancer-associated fibroblasts (CAFs) present within the tumour⁵. Studies of ILC tumour cells and components of the tumour microenvironment (TME) in isolation, such as our work on CAFs in Chapter 3 and Gómez-Cuadrado *et al*, 2022¹³, provide valuable insight into the biology of ILC but lack the important context of crosstalk between the cancer cells and surrounding stroma.

Paracrine signalling from the TME to tumour cells is well known to drive a wide range of pro-tumorigenic phenotypes. In breast cancer, CAF secreted factors, such as growth factors, cytokines and exosomes, have been shown to promote therapy resistance^{133,136,219}, tumour cell metastasis^{101,121}, cancer stem cell formation¹³⁶, immunosuppression^{254,255} and proliferation¹³². Work in the previous chapter and in Gómez-Cuadrado *et al*, 2022¹³ identified a number of differentially expressed genes encoding for secreted proteins between ILC and IDC CAFs, such as *CECR1*, and the forward-phase protein array highlighted a number of highly expressed proteins secreted into the CAF conditioned media, including the pleiotropic cytokine IL-6 (see Chapter 3). Secreted proteins are useful therapeutic targets due to their extracellular location as they can be targeted with highly specific monoclonal antibodies.

To interrogate paracrine signalling in ILC, conditioned media (CM) was collected from our primary patient-derived ILC CAFs, characterised in Chapter 3. This was then used to stimulate the SUM44PE ILC cell line and changes in signalling pathways was evaluated using a reverse-phase protein array (RPPA). This showed a significant increase in Signal Transducer and Activator of Transcription 3 (STAT3) phosphorylation at Tyr705, an activating phosphorylation site. I then identified Interleukin-6 (IL-6) as the CAF-secreted factor driving this signalling and found a significant enrichment for the IL-6/STAT3 pathway and genes activated by the

pathway identified *in vitro* in ER+ ILC tumours compared to IDC tumours in The Cancer Genome Atlas (TCGA)²⁵⁶ dataset. Although the IL-6/STAT3 pathway is well characterised in other subtypes of breast cancer, no research has investigated the pathway in ILC. The enrichment of this pathway in ILC suggests targeting paracrine IL-6 signalling could potentially be a new therapeutic option for ILC patient treatment.

Chapter 4.2 Results

4.2.1 CAF conditioned media induces activation of a number of pro-tumorigenic pathways in ILC tumour cells

In ILC, CAFs are the dominant cell type within the TME⁵⁵ and secrete a wide range of pro-tumorigenic factors (see Chapter 3). To assess the signalling driven by CAF-secreted factors, CAF CM was collected from the primary ILC CAF line ED38 and used to stimulate the ILC tumour cell line SUM44PE for 30 min and 24 hours. A reverse phase protein array (RPPA) was carried out, assessing CAF CM induced changes in 120 signalling proteins and phospho-proteins (for full list of antibodies used, see Appendix x). Twenty-two proteins were significantly changed relative to control at either 30 min or 24 hours (Figure 4-1).

CAF CM stimulation induced activation of a number of phospho-proteins and pathways linked to proliferation at 30 minutes of stimulation. For example, there was a significant increase in phospho-Rb, an inhibitory phosphorylation that allows for cell cycle progression. There was also a significant increase in phosphorylation of proteins in the IGF-1/PI3K/mTOR pathway, such as pIGF1R and p-mTOR. High expression of IGF-1 and a number of IGF-1 binding proteins was detected in CAF CM from ILC and IDC CAFs in the forward phase protein array (FPPA) experiment in Chapter 3 (Figure 3-6). *IGF1* expression has also been shown to be enriched in the stroma of ILC tumours¹³ and, in ER+ ILC tumours in the TCGA dataset, *IGF1* expression significantly and positively correlates with ESTIMATE stromal score (Pearson's $r = 0.69$, $p < 0.0001$, $n = 112$). The ESTIMATE stromal score⁹⁵, based on single sample gene set enrichment analysis, estimates the proportion of stromal content in bulk tumour samples. The IGF-1/PI3K/mTOR pathway is well known to be hyper-activated in ILC, due to mutations in the pathway³ and inactivation of E-cadherin^{52,222}. These data demonstrate further CAF-driven paracrine activation of the pathway, both *in vitro* and in patient tumours.

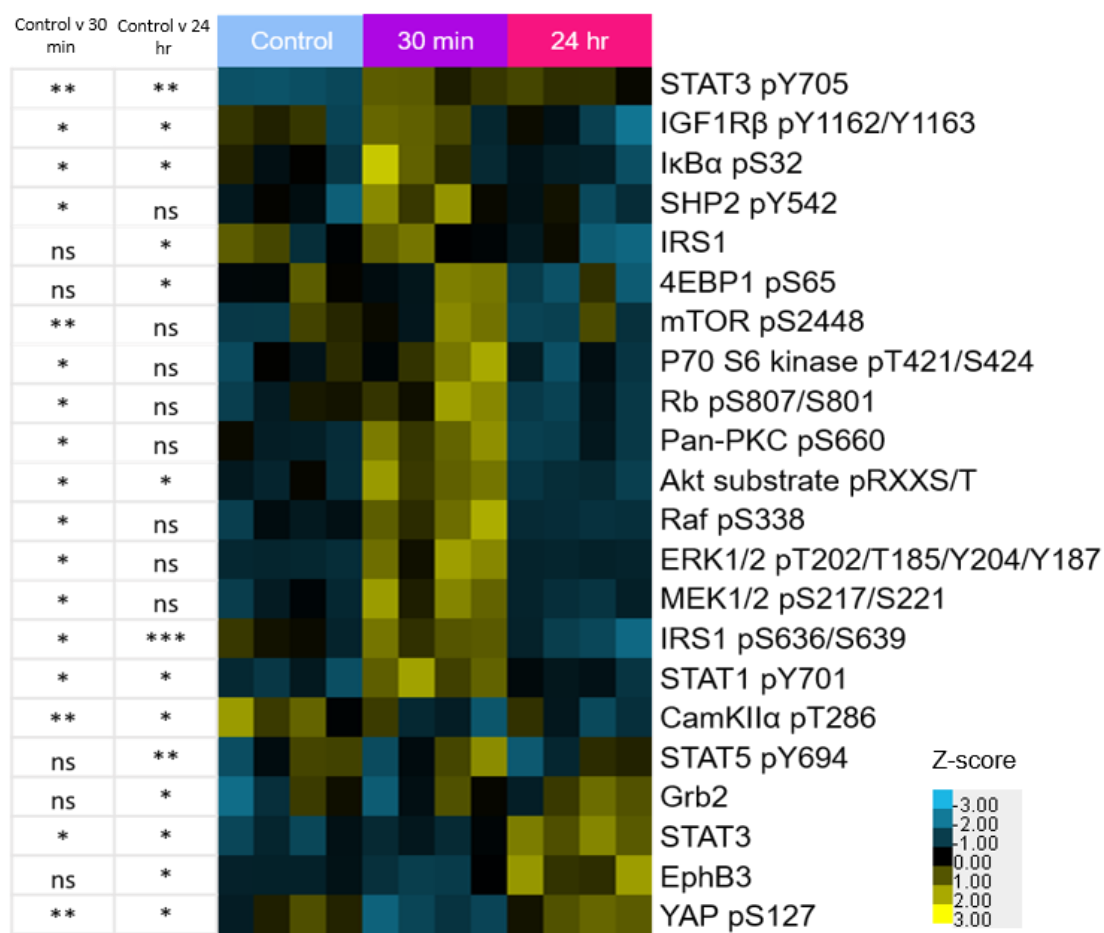


Figure 4-1 . Heatmap of proteins and phospho-proteins significantly altered by ED38 ILC CAF CM stimulation of SUM44PE cells. Showing z-scores of proteins significantly different from control samples, with unbiased hierarchical clustering of the proteins. Two-way paired ANOVA with correction by two-stage step-up BKY method in Graphpad Prism, n=4, FDR<0.05 *, FDR<0.01 **, FDR<0.001 ***.

Phospho-proteins in the mitogen-activated protein kinase (MAPK) pathway had a more than two-fold significant induction by CAF CM stimulation of SUM44PE cells in the RPPA data (Figure 4-2A). Phosphorylation of ERK1/2 was significantly upregulated at 30 minutes, with a 5-fold increase relative to unstimulated cells (Figure 4-1 and Figure 4-2A). ERK1/2 are downstream kinases of the MAPK pathway and can activate a wide range of downstream targets, resulting in increased cell survival, growth and proliferation. RAF pS338 and MEK1/2 pS217/221, upstream kinases in the MAPK pathway, also had a more than 2-fold increase in expression (Figure 4-2A), demonstrating that the CAF CM activates the MAPK pathway in ILC cells.

STAT1 pY701 and STAT3 pY705 also had a more than two-fold significant induction by CAF CM in the SUM44PE cells (Figure 4-2A). STAT1 and STAT3 are transcription factors and phosphorylation on Tyrosine 701 and Tyrosine 705 respectively leads to

hetero- or homo-dimerization of the proteins and translocation to the nucleus where they can act as transcription factors. Although the two proteins are highly similar in terms of structure and activation, STAT1 and STAT3 have divergent rolls in cancer, with STAT1 generally considered a tumour suppressor and STAT3 considered to be pro-tumorigenic²⁵⁷. STAT3 pY705 can upregulate expression of total STAT3 in a feed-forward loop. This may account for the significant increase in total STAT3 protein at 24 hours of stimulation with CAF CM (Figure 4-1). Whereas the MAPK pathway proteins and STAT1 were no longer highly phosphorylated at 24 hours, STAT3 pY705 was maintained after 24 hours. The increase in total and phospho-STAT3 was validated by Western blot using the same cell lysates as used for the RPPA experiment (Figure 4-2B and C). Of note is that no significant changes were seen in STAT3 pSer727 phosphorylation in the RPPA experiment,

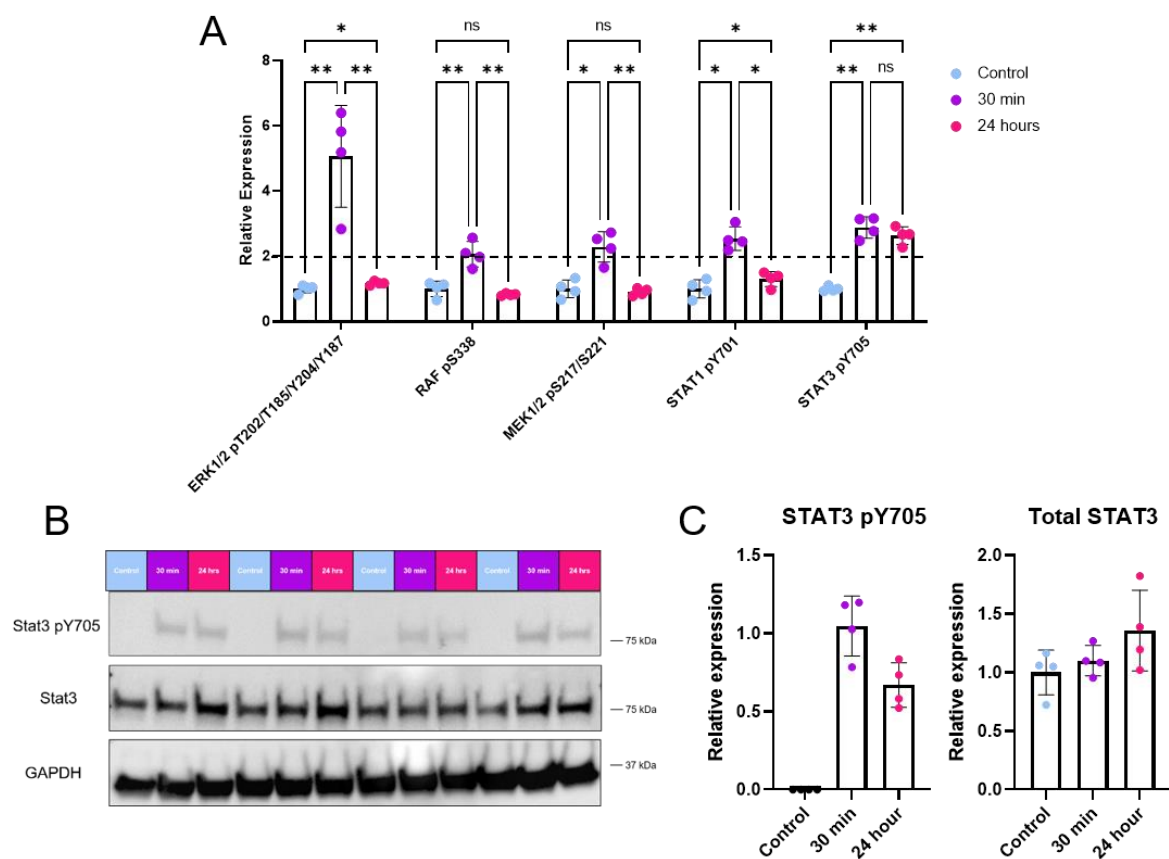


Figure 4-2. CAF conditioned media drives MAPK and STAT3 pathway activation. (A) Expression of proteins with a more than two fold significant difference compared to control samples in the RPPA, each point represents on biological replicate. Statistics in GraphPad Prism, two-way paired ANOVA with correction by two-stage step-up BKY method, FDR<0.05 *, FDR<0.01 **, FDR<0.001 ***. Validation of total and phospho-STAT3 induction by CAF CM by (B) western blot and (C) quantification of band

intensity with points showing band intensity of each biological repeat, bar representing mean $\pm$ standard deviation (sd).

To determine whether activation of these signalling proteins is relevant to ILC tumour biology, the RPPA data from The Cancer Genome Atlas (TCGA)²⁵⁶ was assessed. 122 out of 223 proteins were differentially expressed (FDR<0.01) between ER+ ILC and ER+ IDC tumours, with STAT3 pY705 being one of the most significantly upregulated proteins in ILC tumours (Figure 4-3A). pSTAT3, pERK1/2, pMEK and pRAF were all significantly more highly expressed in ILC tumours, with E-Cadherin and β -catenin significantly upregulated in IDC tumours as expected. This suggests that activation of these pathways occurs in ILC tumours and this signalling is enriched in ER+ ILC tumours compared to ER+ IDC tumours (Figure 4-3B). STAT1 pY701 was not included in this study. STAT3 pY705, pERK1/2 and pMEK all significantly and positively correlate with each other in ER+ ILC tumours (Figure 4-3C), potentially implying a common mechanism of activation. All together, these data demonstrate that secreted factors from primary ILC CAFs drive STAT3 and MAPK pathway activation *in vitro* and that these pathways are also significantly upregulated in ILC patient tumours compared to IDC tumours, suggesting a relevance to human ILC biology.

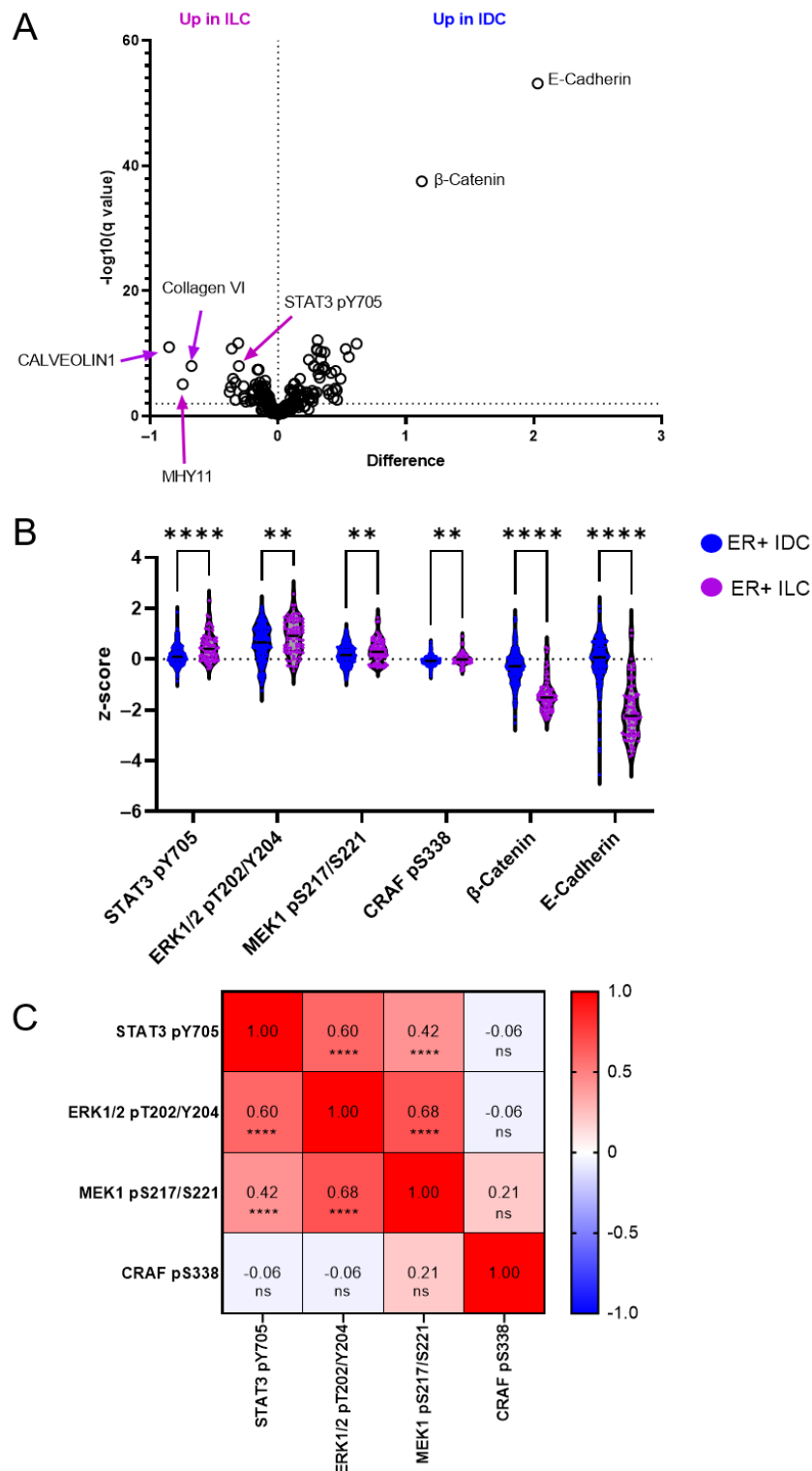


Figure 4-3. pSTAT3 and MAPK pathway activation is enriched in ER+ ILC tumours. (A) Volcano plot showing 222 proteins and their differential expression between ER+ ILC (n=86) and ER+ IDC (n=273) tumours in the TCGA RPPA L4 dataset, dotted line in FDR=0.01. (B) Expression of pSTAT3, pERK1/2, pMEK, pRAF, β -catenin and E-cadherin in ER+ ILC (purple) and IDC (blue) tumours in TCGA RPPA dataset. (A) and (B) multiple unpaired t-tests with correction by two-stage step-up BKY method, FDR<0.05 *, FDR<0.01 **, FDR<0.001 ***, FDR<0.0001 ****. (C) Correlation matrix of pSTAT3, pERK1/2, pMEK and pRAF expression in ER+ ILC tumours (n=86), Pearson's correlation, p<0.0001 ****, ns=not significant.

4.2.2 Identifying CAF-expressed STAT3 activating factors

STAT3 and the MAPK pathway activation can be driven by a wide range of CAF secreted factors including growth factors like fibroblast growth factor 2 (FGF2) and cytokines like interleukin-6 (IL-6). As STAT3 and the MAPK pathway are known to drive oncogenic phenotypes and attempts to directly target STAT3 and the MAPK proteins have been minimally successful clinically (reviewed here^{144,258}), determining the secreted factor driving this activation could allow us to potentially find a more accessible extracellular therapeutic target and inhibit the activation of multiple pro-tumorigenic downstream pathways.

Expression of reported STAT3 activating secreted factors (based on a literature review) was assessed in primary ILC CAFs in the microarray dataset in Gómez-Cuadrado *et al*, 2022 (GSE148156)¹³. These included IL6/gp130 family members, cytokines, growth factors, TGF β family members and neurotrophins^{257,259–265}. None of the reported STAT3-activating factors investigated here were differentially expressed between ILC and IDC CAFs (data not shown) though 16/22 were differentially expressed between ER+ ILC and ER+ IDC tumours in the TCGA or METABRIC datasets (Table 4-1).

The genes encoding *Interleukin-6 (IL6)* and *Fibroblast growth factor 2 (FGF2)* were two of the most highly expressed STAT3 activating factors in the CAFs (Figure 4-4 and Table 4-1). *FGF2* and *IL6* were also significantly more highly expressed in ER+ ILC tumours than ER+ IDC tumours in both TCGA RNAseq and METABRIC⁹³ microarray datasets. IL-6 is a pleiotropic cytokine that is well known to activate STAT3 in breast cancer¹⁴⁴ as well as, more transiently, STAT1 and the MAPK pathway. IL-6 was one of the most highly expressed proteins in ILC and IDC CAF CM in the FPPA dataset (Figure 3-6). In some breast cancer cell lines, FGF2 has been reported to activate STAT3 due to FGFR1 recruiting and activating JAK2 and Src²⁶³. FGFR1 is amplified in 10% of ER+ ILC tumours in the TCGA dataset and in both of the ER+ ILC tumour cell lines, SUM44PE and MDA-MB-134VI²⁶⁶. Antibodies for FGF2 were not available for the FPPA experiment.

As pSTAT3 expression was induced by CAF CM stimulation and significantly increased in ILC tumours, the high expression of *IL6* and *FGF2* in CAFs and ILC

tumours suggests that these could be the CAF secreted factors driving STAT3 activation *in vitro*.

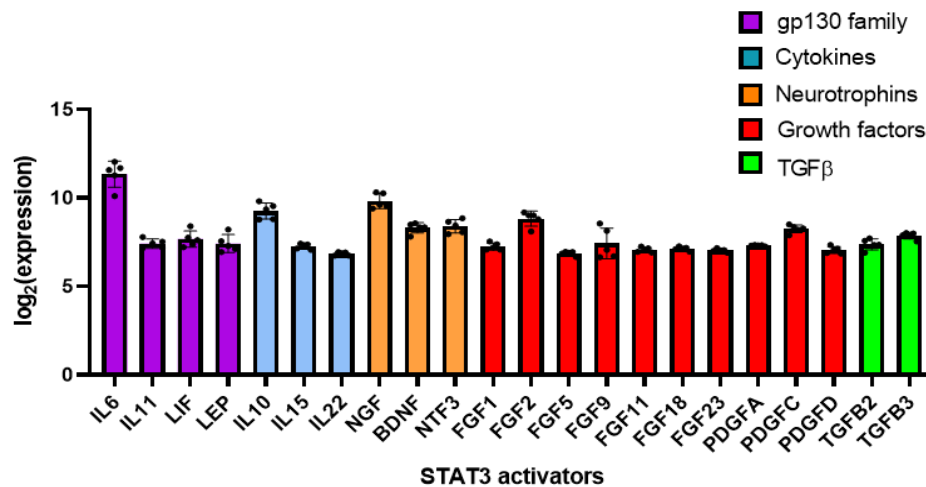


Figure 4-4. mRNA expression of STAT3 activating factors in ILC CAFs. (5 independent cell lines, mean and standard deviation plotted, GSE148156).

	CAF microarray	TCGA		METABRIC	
Gene	Log ₂ (mean) expression in primary ILC CAFs	Higher expression in ILC or IDC	Adjusted p-value	Higher expression in ILC or IDC	Adjusted p-value
IL6	11.341	ILC	1.42E-13	ILC	2.45E-03
NGF	9.827	ILC	6.59E-09	ILC	ns
FGF2	9.427	ILC	1.59E-17	ILC	2.40E-06
NTF3	8.389	IDC	7.82E-04	IDC	ns
TGFB3	7.871	ILC	9.42E-17	ILC	1.71E-06
LIF	7.690	ILC	2.90E-20	ILC	8.86E-03
LEP	7.498	ILC	1.92E-16	ILC	2.24E-08
FGF9	7.463	ILC	5.00E-15	ILC	1.33E-07
TGFB2	7.374	ILC	6.59E-09	ILC	ns
PDGFA	7.207	ILC	8.83E-05	IDC	ns
FGF11	7.081	ILC	2.75E-03	IDC	ns
PDGFD	7.078	ILC	3.09E-11	ILC	2.24E-08
IL15	7.060	ILC	2.11E-07	IDC	ns
FGF1	6.961	ILC	7.12E-04	IDC	ns
FGF18	6.920	ILC	1.11E-17	ILC	6.28E-03
FGF5	6.876	IDC	5.39E-07	ILC	ns

Table 4-1. Expression of STAT3 activating factors in ILC CAFs and in bulk tumours. Showing genes differentially expressed between ILC and IDC tumours in the TCGA or METABRIC datasets, ranked by expression in ILC CAFs from GSE148156. Mean expression of each gene across 5 independent primary ILC CAF cell lines and differential expression (Wilcoxon U test with BH correction) between ER+ ILC and ER+ IDC bulk tumours in the TCGA RNAseq dataset (ILC n=191, IDC=327) and METABRIC microarray dataset (ILC n=122, IDC n=1098). Genes highlighted in purple are significantly increased in ILC tumours in both TCGA and METABRIC. ns – not significant.

4.2.3 IL-6 in CAF conditioned media drives STAT3 activation

To determine whether IL-6 or FGF2 is driving STAT3 activation, SUM44PE cells were stimulated with varying concentrations of recombinant human FGF2 (Figure 4-5A) or IL-6 (Figure 4-5B). FGF2 induced ERK1/2 phosphorylation in a dose responsive manner, as previously reported in this cell line²⁶⁶. However, FGF2 stimulation did not result in STAT3 phosphorylation. In contrast, IL-6 stimulation drove a dose responsive increase in both pERK1/2 and STAT3 pY705 at 30 min (Figure 4-5B, left) and, at 24 hours, an increase in total STAT3 could also be seen (Figure 4-5B, right) as was detected in the RPPA experiment (Figure 4-1).

SUM44PE cells were then stimulated with ILC CAF CM +/- Infigratinib, an FGFR1 inhibitor (FGFRi) (Figure 4-5C), or +/- an IL-6 neutralising antibody (anti-IL-6, Figure 4-5D). This shows that, while FGFRi does inhibit FGF2-induced ERK1/2 phosphorylation and slightly reduces pERK1/2 driven by CAF CM, it has no effect on CAF CM induced STAT3 phosphorylation. However, the IL-6 neutralising antibody completely inhibits STAT3 Tyr705 phosphorylation induced by both recombinant IL-6 and CAF CM and reduces CAF CM induced ERK1/2 activation. This was validated in another human ILC tumour cell line, MDA-MB-134VI (MDA134, Figure 4-5E). Stimulation with recombinant IL-6 for 30 min also significantly increased expression of STAT3 in the nucleus in SUM44PE cells (Figure 4-5F). This suggesting that IL-6 induced STAT3 phosphorylation leads to STAT3 activation and translocation to the nucleus where it can act as a transcription factor. These data demonstrate that IL-6 secreted by CAFs is partially responsible for MAPK pathway activation and is solely responsible for STAT3 activation.

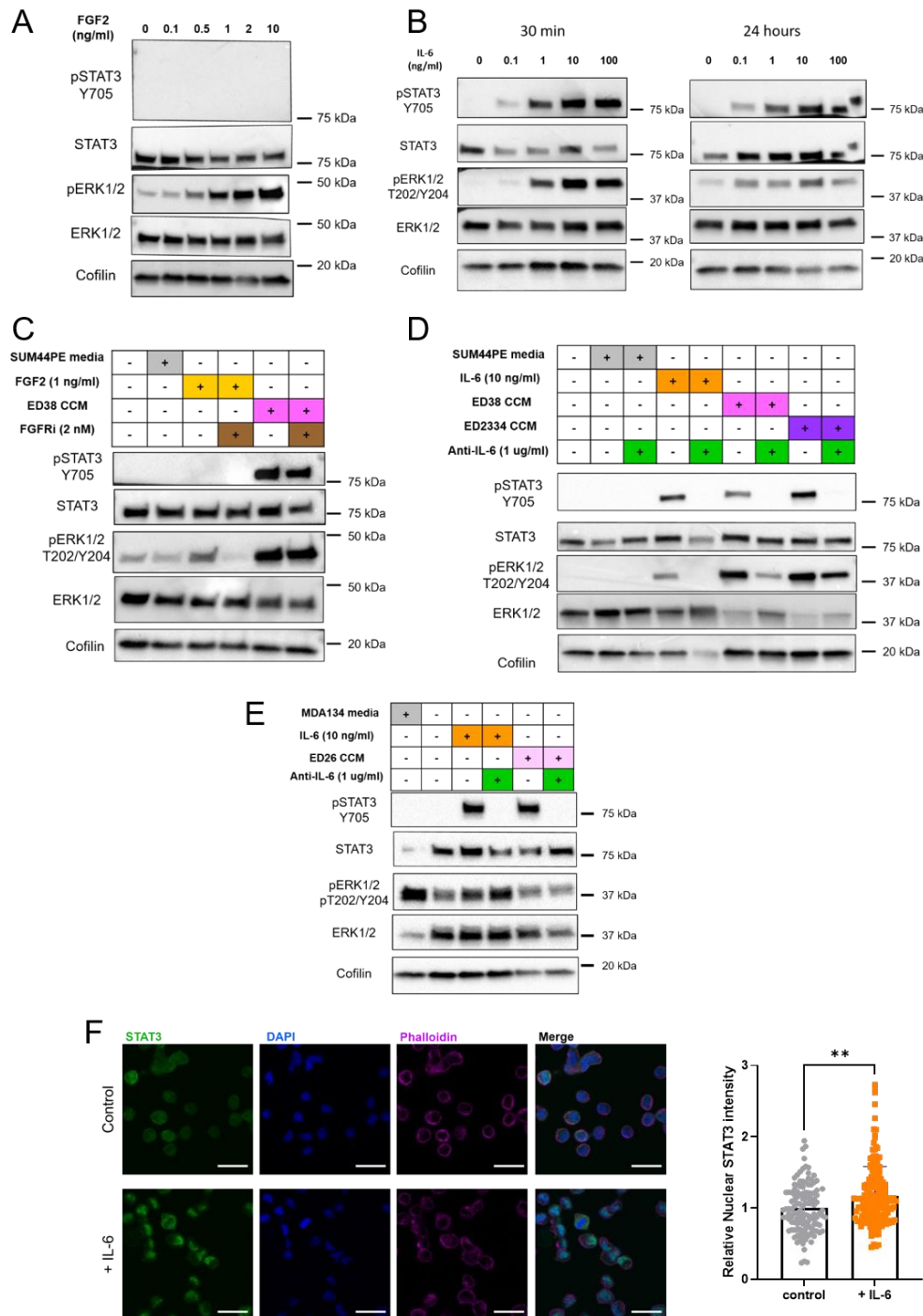


Figure 4-5. IL-6 drives STAT3 activation. SUM44PE cells were stimulated with increasing concentrations of recombinant FGF2 (A) for 30 min and IL-6 (B) for 30 min (left) or 24 hours (right). SUM44PE cells were stimulated with (C) 1 ng/ml recombinant FGF2 or ED38 ILC CAF conditioned media for 30 min +/- Infigratinib (FGFRi, 2nM) and (D) 10 ng/ml recombinant IL-6 or CAF conditioned media from two independent primary ILC CAF cell lines, ED38 and ED2334, +/- anti-IL-6 (1 µg/ml). (E) MDA-MB-134VI cells were stimulated with 10 ng/ml IL-6 or ILC CAF line ED26 CM +/- anti-IL-6 (1 µg/ml). Cell lysates from (A-E) were probed for STAT3 pY705, ERK1/2 pT202/Y204 and total proteins with Cofilin as loading control. (F) Left - Representative images of SUM44PE cells stimulated with recombinant IL-6 for 30 min and stained for STAT3 (green) with DAPI (blue) and Phalloidin (magenta), 40X magnification, scale bar 25 µm. Right - Relative nuclear STAT3 intensity (mean gray value normalised to mean of control mean gray values) was quantified in ImageJ, n=3 for each condition with three random FoVs quantified for each repeat. Each spot represents one nuclei, bar shows mean, two-tailed Mann Whitney U test p=0.0013.

To quantify the amount of IL-6 secreted by CAFs and tumour cells, an IL-6 ELISA was carried out. Conditioned media was collected from three primary ILC CAF lines (ED2334, ED38 and ED26), two primary IDC CAF lines (ED2240 and ED9B) and the ILC cell lines, SUM44PE and MDA-MB-134VI. ILC CAFs secrete roughly 10 ng/ml IL-6 into their conditioned media with IDC CAF CM containing roughly 3 ng/ml IL-6 on average (Figure 4-6A). ILC tumour cells did not secrete any detectable IL-6 into their conditioned media, suggesting an entirely paracrine IL-6 signalling pathway. The concentration of secreted IL-6 was normalised to the total protein concentration of the lysed CAF cells that the CM was collected from. There was a trend towards ILC CAFs secreting more IL-6 than IDC CAFs (Figure 4-6B), although this was non-significant and *IL6* gene expression was not significantly different between ILC and IDC CAFs (Figure 4-6C, data from CAF RNAseq experiment in Chapter 3 and Gómez-Cuadrado *et al*, 2022¹³). IL-6 was one of the most highly secreted proteins by ILC CAFs as detected in the CAF FPPA (Figure 3-6), with a non-significant trend towards higher IL-6 concentrations in CAF CM from ILC CAFs than IDC CAFs.

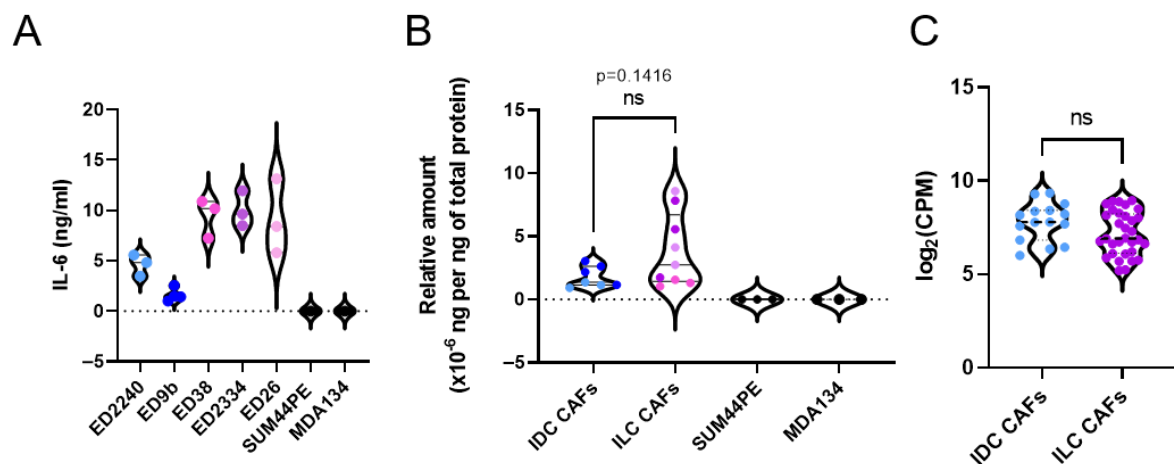


Figure 4-6. IL-6 expression in primary ILC and IDC CAFs. (A) The absolute concentration and (B) the concentration relative to cell pellet protein concentration of IL-6 in conditioned media was determined by ELISA for patient derived primary CAFs from two IDC patients (ED2240, light blue, and ED9b, dark blue), three ILC patients (ED38, pink, ED2334, purple, ED26, light pink) and the ILC cell lines SUM44PE and MDA-MB-134VI (MDA134) (black), n=3 biological replicates for each cell line. (C) *IL6* expression in 11 ILC and 5 IDC CAF lines determined by RNAseq (see Chapter 3¹³). (B) and (C) unpaired two-tailed Mann Whitney tests in Graphpad Prism).

4.2.4 IL-6/STAT3 signalling is enriched in ILC tumours and correlates with stromal content

To determine whether IL-6/STAT3 signalling is enriched in ILC tumours, ER+ ILC and IDC tumours were compared in the TCGA dataset. *IL6* is significantly more highly expressed in ILC than IDC tumours (Table 4-1 and Figure 4-7A) and significantly positively correlates with STAT3 pY705 in matched tumours from the TCGA RNAseq and RPPA datasets (Figure 4-7B). This suggests a relationship in tumours as well as *in vitro*.

As the *in vitro* data in Figure 4-6 suggests that this is an entirely paracrine signalling pathway, I wanted to assess if IL-6 is expressed in the stroma of ILC tumours. Using laser capture microdissection data from ER+ IDC²¹⁶ and ER+ ILC¹³ tumours, there is a significantly higher stromal to tumour epithelium ratio of *IL6* expression in ILC compared to IDC tumours (Figure 4-7C), suggesting that *IL6* is more highly expressed by the stromal compartment of ILC than in IDC.

ILC tumours have a higher stromal content than IDC tumours^{5,6}, due to the single-file invasive pattern of ILC growth. In TCGA data, ILC tumours have a significantly lower “consensus measurement of purity estimate” (CPE)²¹⁵ (Figure 4-7D), showing a lower ‘tumour purity’ or tumour cell content in ILC tumour samples. This also supports the increased presence of non-transformed cells such as CAFs represented in the bulk tumour analysis in ILC tumours in the TCGA RNAseq dataset. *IL6* expression in ER+ ILC and IDC tumours in the TCGA dataset was normalised to CPE scores (Figure 4-7E). Even after correction, ILC tumours have significantly higher *IL6* expression, again supporting an enrichment for IL-6 signalling in ILC.

ILC tumours have significantly higher ESTIMATE stromal scores⁹⁵ than IDC tumours, also suggesting a higher stromal content (Figure 4-7F). ESTIMATE stromal scores significantly and positively correlate with *IL6* and *STAT3* expression (Figure 4-7G and H) as well as STAT3 pY705 (Figure 4-7I) in ER+ ILC tumours. These data demonstrate that IL-6/STAT3 signalling is enriched in ILC compared to IDC and supports paracrine activation of the pathway by the stroma of ILC tumours, as we have shown *in vitro*.

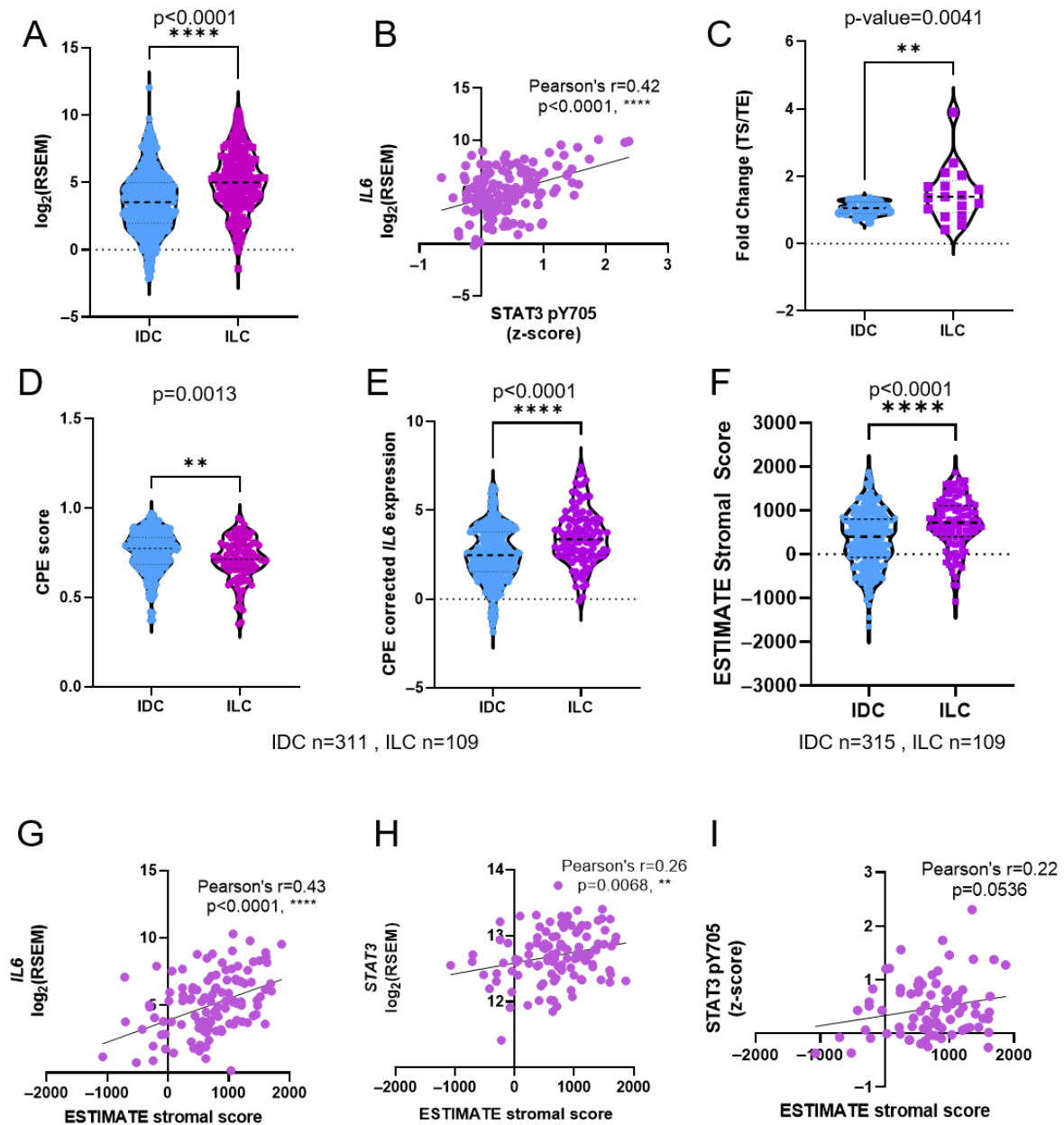


Figure 4-7. IL-6/STAT3 signalling is enriched in ILC tumours and correlates with stromal content. (A) *IL6* expression in ER+ IDC (blue, n=555) and ER+ ILC tumours (purple, n=191), TCGA dataset. (B) Correlation between *IL6* mRNA and STAT3 pY705 in ER+ ILC tumours in the TCGA RNAseq and RPPA datasets respectively, n=143. (C) Ratio of *IL6* expression in tumour stroma (TS) to tumour epithelium (TE) in IDC tumours (blue, n=32, GSE68744²¹⁶) and ILC tumours (purple, n=17, GSE148398¹³). (D) CPE scores²⁶⁷ for ER+ IDC (n=311) and ER+ ILC (n=109) tumours in the TCGA dataset and (E) *IL6* expression corrected for CPE scores. (F) ESTIMATE stromal scores in 315 ER+ IDC and 109 ER+ ILC tumours in TCGA dataset and correlation of ESTIMATE stromal score in ER+ ILC tumours with (G) *IL6* and (H) STAT3 (TCGA RNAseq, n=109) and with (I) STAT3 pY705 (TCGA RPPA, n=78). Pearson's correlation in B and G-H and unpaired t-tests in A and C-F were carried out in GraphPad Prism, ** $p < 0.01$, **** $p < 0.0001$.

4.2.5 IL-6 drives gene expression changes in ILC tumour cells

Phosphorylated STAT3 dimerises and translocates to the nucleus, where it drives gene expression. A 3'-mRNA-seq experiment was performed to determine the gene expression driven by CAF CM and IL-6 within the CAF CM in SUM44PE cells. SUM44PE cells were stimulated with CAF CM from ED26 ILC CAFs +/- anti-IL-6 for 24 hours and RNA and protein was extracted. Western blotting validated the IL-6 dependent phosphorylation of STAT3 by CAF CM in SUM44PE cells (Figure 4-8A).

Differential gene expression analysis was carried out using the EdgeR package in R²⁰⁴. Unbiased hierarchical clustering of the top 500 most variable genes across the whole dataset revealed that the control samples and the cells treated with CAF CM + anti-IL-6 clustered together and cells treated only with CAF CM clustered separately (Figure 4-8B). An ANOVA-like generalised linear model identified 110 differentially expressed genes (DEGs) between all three groups, with unbiased hierarchical clustering of these genes again showing CAF CM treated samples clustering separately from the other groups (Figure 4-8C). This suggests that the vast majority of DEGs are being driven by IL-6 within the CAF conditioned media. There are only 5 DEGs between control and CM + anti-IL-6 samples compared to 130 and 101 DEGs between CM treated samples and control or CM + anti-IL-6 treated samples respectively (Figure 4-8D) further supporting the IL-6 driven gene expression changes. Expression of some of the most highly induced genes, *S100A8*, *S100A9*, *MUCL1* and *SERPINB5*, was validated by RT-qPCR (Figure 4-8E).

A table of the results of the differential gene expression analysis can be found in "[Esme Bullock Thesis Documents](#)".

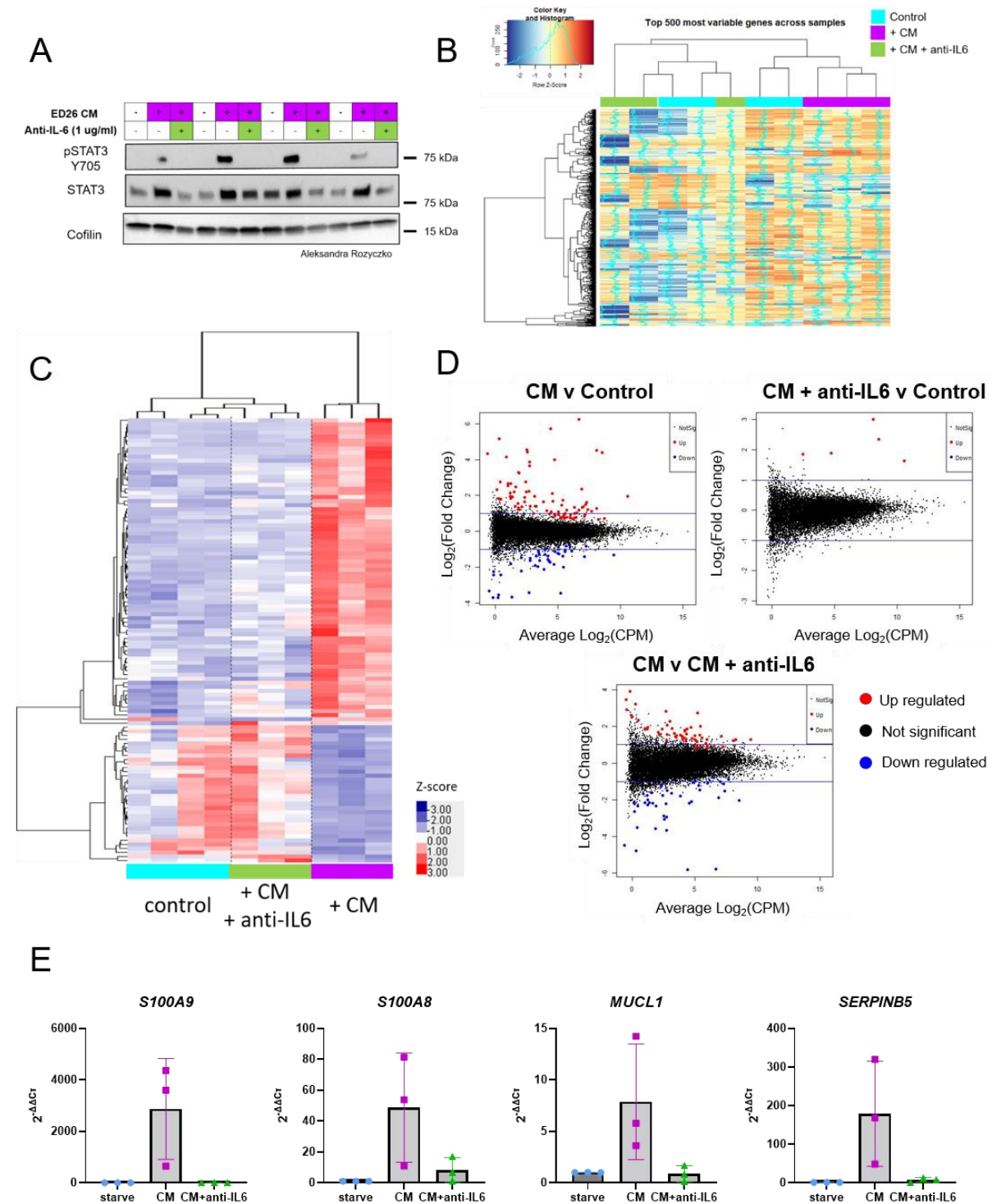


Figure 4-8. CAF secreted IL-6 drives gene expression changes in ILC tumour cells. (A) Western blot of SUM44PE cells in growth factor depleted media stimulated with ED26 ILC CAF CM +/- anti-IL-6 (1 µg/ml), lysates probed for STAT3, STAT3 pY705, ERK1/2, ERK1/2 pT202/Y204 and cofilin as loading control, run by Aleksandra Rozyczko. (B) Heatmap of the top 500 most variable genes across all samples. (C) Heatmap showing unbiased hierarchical clustering of 111 significantly differentially expressed genes between all three groups, EdgeR ANOVA-like generalised linear model FDR<0.05 and (D) plots of differentially expressed genes between control and +CM (top left), control and +CM + anti-IL-6 (top right) and + CM and +CM + anti-IL-6 (bottom). (E) Expression of *S100A9*, *S100A8*, *MUCL1* and *SERPINB5* in starved SUM44PE cells or cells stimulated with CAF CM +/- IL-6, determined by RT-qPCR.

To determine some of the possible downstream functional consequences of the CAF CM driven genes, gene set enrichment analysis (GSEA)²⁶⁸ using the Hallmarks gene sets²⁰⁹ was performed on the RNAseq dataset. This showed that the CAF CM treated cells were enriched for IL6-STAT3 signalling, interferon signalling and epithelial-to-mesenchymal transition (EMT) gene sets, whereas the control cells were enriched for proliferative and estrogen response early gene sets (Figure 4-9A). Similarly, when GSEA was carried out comparing cells stimulated with CM to cells stimulated with CM + anti-IL-6, CM cells were also enriched for EMT and inflammatory gene sets and CM + anti-IL-6 cells were enriched for proliferative and estrogen signalling gene sets (Figure 4-9B). This implies that IL-6 within the CAF CM is driving these inflammatory and EMT gene signatures rather than other CAF-secreted factors.

CAF CM treated cells were also significantly enriched for expression of genes found in the 'immune-related' and 'reactive-like' subtypes of ILC described by Ciriello *et al*³, compared to both control and CAF CM + anti-IL6 cells (Figure 4-9C and D). Five genes (*AKR1C1*, *VNN2*, *DGAT2*, *AKR1C2* and *THRSP*) significantly upregulated by CAF CM treatment are part of the 'immune-related signature' and seven genes (*SERPINB5*, *SLC28A3*, *BBOX1*, *MUCL1*, *SLC25A37*, *CLCA2* and *AKAP3*) are part of the 'reactive-like' signature. The 'reactive-like' subtype of ILC is characterised by low tumour cell content and high stromal and CAF related gene expression and the gene expression in the 'immune-related' subtype correlates with macrophage related signatures³. A third subtype, 'proliferative', is defined by low expression of the genes in reactive and immune subtypes. Interestingly, ER+ ILC tumours in the proliferative subtype have significantly lower expression of *IL6* compared to tumours classified as both reactive-like and immune-related in the TCGA dataset (Figure 4-9E). However, there is no significant difference in pSTAT3 expression between subtypes using STAT3 pY705 expression in the TCGA RPPA dataset (data not shown).

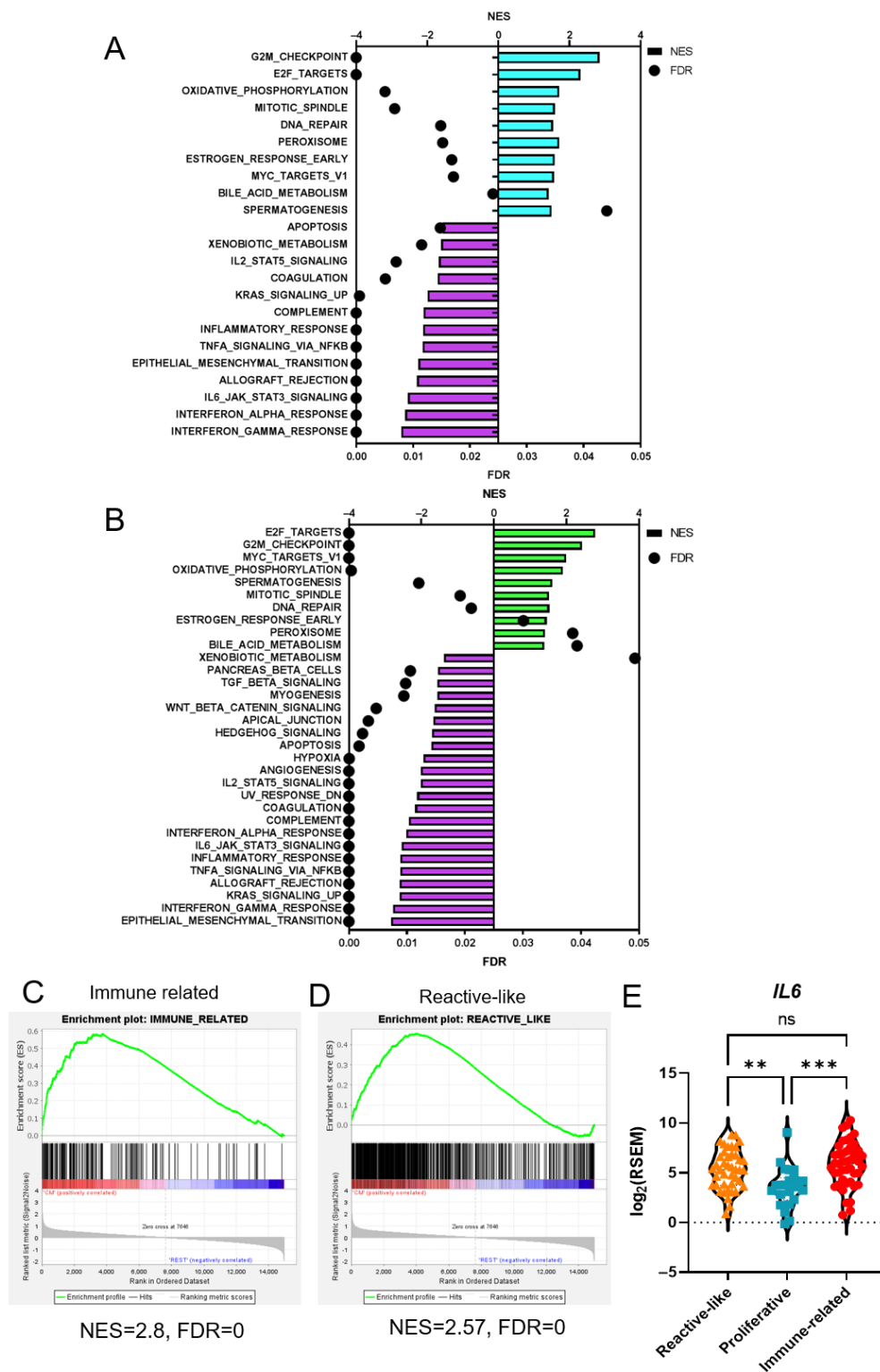


Figure 4-9. IL-6 in CAF CM drives inflammatory and migratory gene set expression. Plots showing the Hallmark gene sets enriched in (A) control (blue bars) compared to CAF CM stimulated cells (purple bars) and (B) CAF CM stimulated cells (purple bars) were compared to CAF CM + anti-IL-6 stimulated cells (green bars). Showing significantly enriched (FDR<0.05) gene sets: dots show FDR and bars show normalised enrichment score (NES). GSEA plots of Immune-related (C) and Reactive-like (D) ILC subtype gene signatures in CM treated v control and CM + anti-IL-6 defined by Ciriello *et al*^β. (E) *IL6* expression in ER+ ILC tumours in the TCGA dataset, by Ciriello subtype, reactive-like n=42, immune-related n=44, proliferative n=23, one-way ANOVA with multiple comparisons, p<0.01 **, p<0.001 ***.

GSEA was carried out comparing CAF CM treated cells to both control and CM + anti-IL-6 cells using Gene ontology biological process (GO BP) gene sets. This identified 291 gene sets significantly enriched in CM treated cells and 53 gene sets significantly enriched in control and CM + anti-IL-6 cells (FDR<0.01), representing gene sets regulated by IL-6 in the CAF CM. Details of enriched gene sets can be found in "[Esme Bullock Thesis Documents](#)".

A network of these enriched gene sets was built using EnrichmentMap²¹² in Cytoscape and gene sets with similar biological themes (multiple overlapping genes within gene sets) were clustered together using Autoannotate²⁶⁹ (Figure 4-10). This identified the clusters 'chromosome mitotic segregation', 'respiratory electron transport' and 'DNA unwinding replication' as downregulated by IL-6 in CAF CM. This further supports an IL-6 dependent suppression of the cell cycle and oxidative phosphorylation in SUM44PE cells.

The three largest clusters of gene sets enriched in conditioned media treated cells were annotated as 'immune response activation', 'migration neutrophil chemotaxis' and 'embryonic development morphogenesis'. The 'immune response activation' cluster contains 183 gene sets with a wide range of functions, including immune cell activation and migration, response to infection, cell adhesion, cytokine signalling and inflammation. 27 of the 76 genes significantly upregulated by CAF CM are included in these gene sets, with 91/183 of these gene sets containing *IL4R*. This encodes for the receptor for the cytokine interleukin-4 (IL-4) and is very highly expressed on a wide range of immune cells²⁷⁰. This could account for the high number of immune related pathways, as the majority of these gene sets are based on immune cell intrinsic signalling rather than the effect of signalling from tumour cells. Similarly, IL-6/STAT3 signalling is key to a wide range of immune cell functions²⁷¹ and IL-6 is key mediator of chronic inflammation²⁷², so these IL-6 induced genes may have different roles in cancer cells than in immune cells. A number of bacterial and protozoan response gene sets are enriched in CAF CM treated cells in comparison to both control and CM + anti-IL-6 cells, demonstrating that the upregulation of these gene sets is not due to contaminated CAF CM but due to genes activated by IL-6. IL-6 signalling is involved in normal cell response to infection, inflammation and the acute phase response²⁷³.

The ‘migration neutrophil chemotaxis’ cluster contains 28 gene sets. This cluster consists of gene sets involved in migration, related to both immune cells (granulocyte, neutrophil and leukocyte migration/chemotaxis gene sets) but also more generic cell migration gene sets such as ‘cell chemotaxis’ and ‘cell migration’. This cluster is primarily driven by high expression of the S100 family of genes (*S100A7*, 8 and 9, part of 20, 11 and 11 gene sets respectively) and *LBP* (*Lipopolysaccharide binding protein*), in 19/28 gene sets. This suggests that IL-6 signalling may be driving a migratory phenotype in SUM44PE cells or potentially drive expression of chemokines that can drive immune cell chemotaxis.

The ‘embryonic development morphogenesis’ cluster primarily consists of gene sets involved in blood vessel development, such as ‘blood vessel morphogenesis’ and ‘vasculature development’ and so could be considered an angiogenesis cluster. *LRG1*, which is significantly upregulated by IL-6 in CAF CM, is included in 12/25 of these gene sets and is known to drive angiogenesis in a wide range of diseases and cancers^{274–278}, often via STAT3 activation. Although the genes encoding for vascular endothelial growth factors (VEGF) are not modulated by IL-6 stimulation of the SUM44PE cells, a cluster of 5 gene sets annotated as ‘vascular endothelial growth’ is also enriched in CM stimulated cells, suggesting other components of this pro-angiogenic signalling pathway are activated.

This analysis also identified a cluster of 6 gene sets annotated as ‘signalling jak stat’, containing gene sets such as ‘receptor signalling pathway via STAT’, and two gene sets in the ‘mapk cascade erk1’ cluster (‘positive regulation of MAPK cascade’ and ‘positive regulation of ERK1 and ERK2 cascade’), confirming that the changes seen in activation of the pSTAT3 and MAPK pathway at the protein level (Figure 4-1 and Figure 4-8A) are also seen at the gene expression level too.

Taken together, these data suggest that IL-6 secreted by CAFs drives an, inflammatory, pro-angiogenic and migratory transcriptomic phenotype in ILC tumour cells.

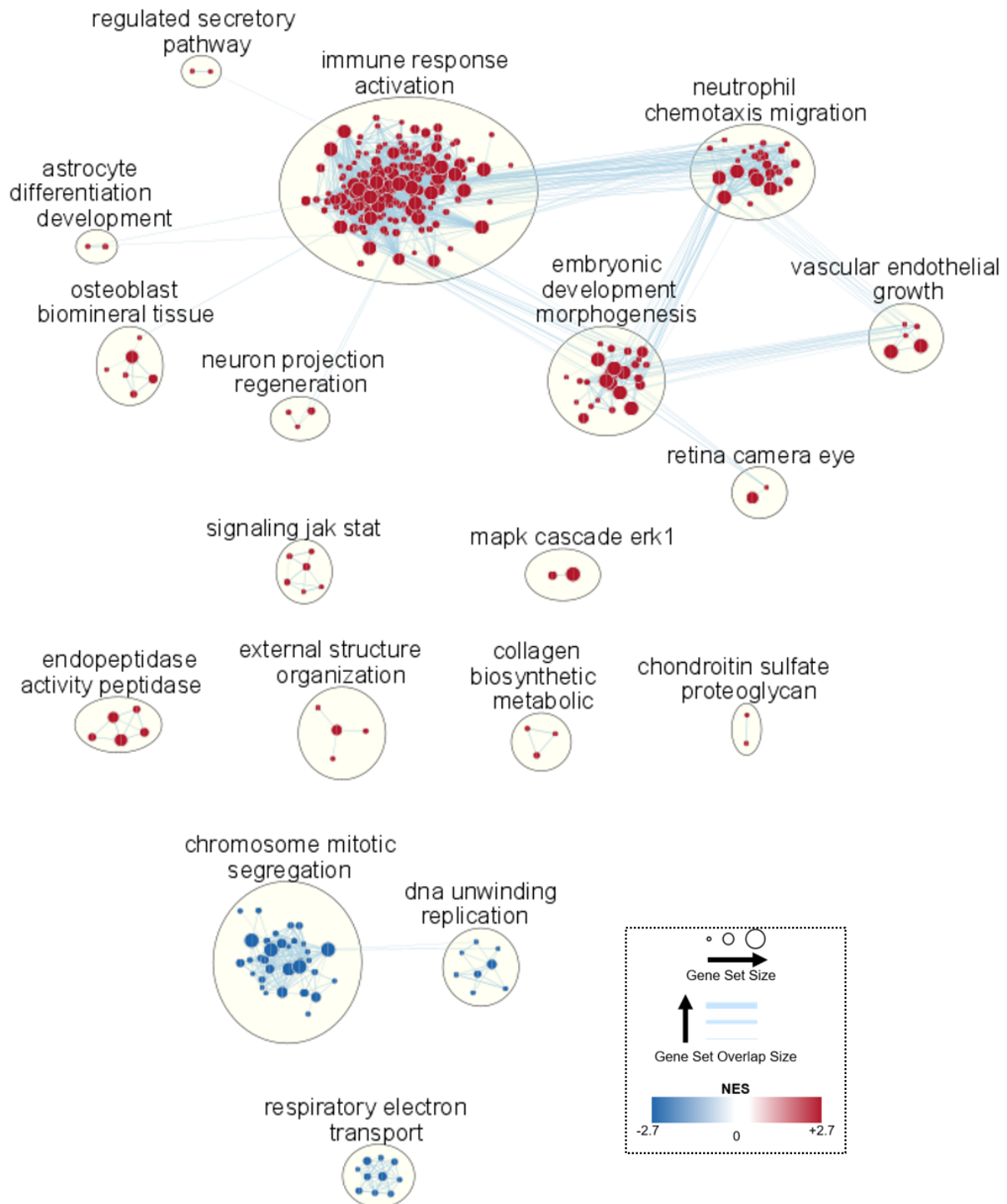


Figure 4-10. Conditioned media treated cells are enriched for immune signalling, migration and angiogenesis gene sets. Network of significantly enriched ($FDR < 0.01$) GO Biological Process gene sets comparing CM treated cells (red nodes, $NES > 0$) to control and CM + anti-IL-6 samples (blue nodes, $NES < 0$). The network was made using GSEA and Cytoscape EnrichmentMap plugin, edges connect gene sets with overlapping genes and thickness of the line indicates how much overlap there is between gene sets. Encircled clusters of nodes indicate interconnected nodes with similar major biological themes determined by Cytoscape AutoAnnotate plugin. NES=normalised enrichment score. Details of the gene sets in each cluster can be found in [Esme Bullock Thesis Documents](#).

4.2.6 An IL-6 gene signature is enriched in ILC tumours

To determine which genes induced by CAF CM were dependent on IL-6 signalling, a 43-gene “IL-6 gene signature” (IL6GS) was defined. These are genes with a more than two-fold increase in expression on CAF CM stimulation which was then reduced to levels seen in control cells in cells treated with CM + anti-IL-6 (Figure 4-11A and Table 4-2). Details of each IL6GS member and their reported role in breast cancer and in IL-6/STAT3 signalling can be found in Appendix 4. 38 of these genes have previously been annotated to have STAT3 bound to their promoters in ChIP-seq experiments in breast cancer cell lines (chip-atlas.org)²⁷⁹, and three of these genes (*CEBPD*, *S100A9* and *TMC5*) have been identified as part of an ER+ BC cell line consensus IL-6 signature¹⁵⁸, lending further support to their IL-6 dependence.

34 of the IL6GS genes are significantly differentially expressed between ER+ ILC and ER+ IDC tumours in the TCGA dataset, with 27 genes significantly more highly expressed in ILC tumours (Figure 4-11B and Table 4-2). Expression of these genes in the stroma could also contribute to high expression in ILC tumours in the TCGA RNAseq dataset, which analysed bulk tumour samples. To account for this, the IL6GS was also investigated in the ILC laser-capture microdissection dataset¹³ and primary CAF RNAseq dataset discussed in Chapter 3. Twenty of the IL6GS genes had no detected expression in ILC CAFs and 24 genes had a higher expression in the tumour epithelium than the tumour stroma of ILC tumours¹³ (Table 4-2). 23 genes also significantly and positively correlate with *IL6* expression in ER+ ILC tumours in the TCGA dataset and 12 genes significantly and positively correlate with pSTAT3. 10 genes significantly positively correlate with both pSTAT3 and *IL6*.

Overall, 12 genes in the IL6GS (*BBOX1*, *BCL3*, *CLCA2*, *FGD3*, *LRG1*, *LRP2*, *LRRC4C*, *LTF*, *MUCL1*, *PPP1R1B*, *THRSP* and *TMC5*) are significantly more highly expressed in ILC than IDC tumours, have higher expression in the tumour epithelium of ILC and are not expressed by primary ILC CAFs. Of these *BBOX1*, *LRRC4C*, *LTF* and *MUCL1* also significantly and positively correlate with *IL6* and STAT3 pY705 in ER+ ILC tumours in the TCGA dataset.

An IL6GS score was calculated for each ER+ ILC and ER+ IDC tumour in the TCGA dataset based on the sum of the z-scores for each gene within the signature. ILC tumours had significantly higher IL6GS scores compared to IDC tumours (Figure

4-11C), supporting an enrichment of these IL-6 driven genes in ILC. IL6GS scores significantly and positively correlate with *IL6* expression (Figure 4-11D) and significantly positively correlate with stromal content in ER+ ILC tumours determined by ESTIMATE (Figure 4-11E) and EDEc⁹⁶ (data not shown, Pearson's $r = 0.24$, p -value = 0.01, $n = 109$). This implies that paracrine IL-6 signalling from the stroma of ILC tumours may promote the expression of IL6GS genes as was seen *in vitro*.

The IL6GS score also significantly and positively correlates with *IL6* expression in ER+ IDC tumours (Pearson's $r = 0.2$, p -value < 0.0001, $n = 555$), though the correlation is weaker than in ILC tumours. However, the IL6GS score does not significantly correlate with stromal content in ER+ IDC tumours (data not shown, ESTIMATE Pearson's $r = 0.1$, EDEc Pearson's $r = 0.002$, $n = 315$). This implies that, although expression of the genes within the IL6GS may be driven by IL-6 signalling in both ILC and IDC tumours, only in ILC tumours does this signature correlate with stromal content, again supporting the CAF-driven paracrine nature of this signalling pathway in ILC.

Together, these data support an enrichment of the paracrine IL-6/STAT3 pathway and resultant gene expression in ILC tumours.

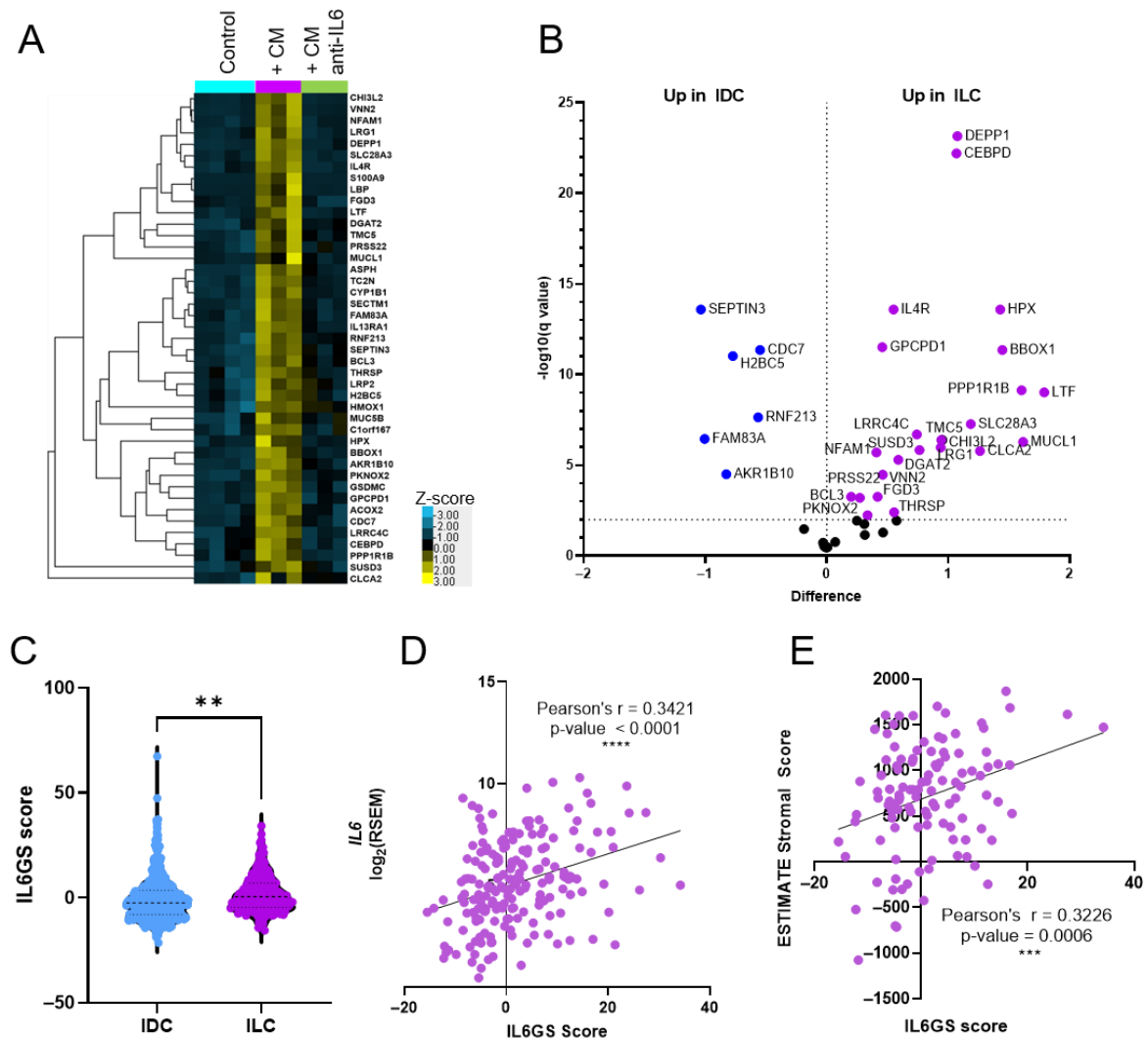


Figure 4-12. An IL-6 dependent gene signature is enriched in ILC tumours in TCGA dataset. (A) Heatmap of IL6GS genes in SUM44PE RNAseq data. (B) Volcano plot of expression of IL6GS genes in ER+ IDC (n=563) and ER+ ILC (n=191) tumours from the TCGA RNAseq dataset, unpaired two-tailed multiple t-tests with BKY correction, significant genes $q\text{-value} < 0.05$, blue dots significantly higher in IDC, purple dots significantly higher expression in ER+ ILC tumours. (C) IL6GS score in ER+ ILC and ER+ IDC tumours, two-tailed unpaired t-test, $p = 0.0015$. Pearson's correlation of IL6GS score with (D) *IL6* expression (n=191) and (E) with ESTIMATE stromal score (n=109) in ER+ ILC tumours in the TCGA dataset.

	SUM44PE RNAseq		TCGA ER+				CAF RNAseq ¹³	ILC LCM ¹ ₃
Gene	Log FC (CM/ control)	Log FC (CM + anti-IL6/ control)	Log FC (ILC/ IDC)	Signif in ILC	Correlation with <i>IL6</i> in ILC	Correlation with pSTAT3 in ILC	Expressed by CAFs	Log FC (TS/ TE)
<i>ACOX2</i>	1.95	0.30	0.31	YES	ns	ns	YES	-0.19
<i>AKR1B10</i>	2.75	0.99	-0.83	NO	ns	ns	YES	0.28
<i>ASPH</i>	1.61	0.46	0	NO	ns	ns	YES	-0.24
<i>BBOX1</i>	3.24	0.90	1.45	YES	****	**	NO	-0.32
<i>BCL3</i>	1.00	0.40	0.20	YES	***	ns	NO	-0.24
<i>C1orf167</i>	1.95	0.67	NA	NA	NA	NA	NA	NA
<i>CDC7</i>	1.22	-0.01	-0.55	NO	ns	ns	YES	-0.54
<i>CEBPD</i>	1.22	-0.31	1.07	YES	****	***	YES	0.00
<i>CHI3L2</i>	3.85	0.18	0.95	YES	***	ns	YES	-0.51
<i>CLCA2</i>	1.65	0.88	1.26	YES	ns	ns	NO	-0.07
<i>CYP1B1</i>	1.12	0.24	0.25	YES	****	*	YES	1.67
<i>DEPP1</i>	1.73	0.25	1.08	YES	****	ns	YES	NA
<i>DGAT2</i>	1.66	0.67	0.59	YES	****	ns	YES	-0.01
<i>FAM83A</i>	2.13	0.36	-1.01	NO	ns	ns	NO	0.03
<i>FGD3</i>	1.02	-0.25	0.42	YES	ns	ns	NO	-0.26
<i>GPCPD1</i>	1.09	0.11	0.46	YES	***	ns	YES	0.32
<i>GSDMC</i>	2.22	0.02	-0.01	NO	*	ns	NO	-0.02
<i>H2BC5</i>	1.08	0.35	-0.77	NO	ns	ns	YES	NA
<i>HMOX1</i>	1.48	0.99	-0.02	NO	*	ns	YES	0.90
<i>HPX</i>	2.67	0.51	1.43	YES	* (neg)	ns	YES	-0.72
<i>IL13RA1</i>	1.04	0.22	-0.03	NO	ns	ns	YES	-0.06
<i>IL4R</i>	1.65	-0.13	0.55	YES	****	**	YES	0.22
<i>LBP</i>	5.72	-0.09	0.31	NO	****	ns	YES	0.08
<i>LRG1</i>	2.10	-0.09	0.94	YES	ns	ns	NO	-1.07
<i>LRP2</i>	1.34	0.69	0.58	YES	ns	ns	NO	-0.45
<i>LRRC4C</i>	1.38	-0.10	0.74	YES	**	**	NO	-0.66
<i>LTF</i>	2.46	0.29	1.79	YES	**	*	NO	-0.56
<i>MUC5B</i>	1.55	0.74	0.46	NO	ns	ns	NO	0.38
<i>MUCL1</i>	2.34	0.46	1.62	YES	****	*	NO	-1.16
<i>NFAM1</i>	2.66	0.38	0.41	YES	ns	ns	NO	1.08
<i>PKNOX2</i>	1.90	0.68	0.34	YES	****	*	YES	0.14
<i>PPP1R1B</i>	1.32	-0.29	1.60	YES	**	ns	NO	-0.55
<i>PRSS22</i>	1.14	0.34	0.27	YES	*	*	YES	-0.67
<i>RNF213</i>	1.17	0.45	-0.57	NO	ns	ns	YES	0.00
<i>S100A9</i>	6.23	0.46	0.01	NO	****	ns	NO	0.06
<i>SECTM1</i>	1.45	0.23	-0.19	NO	*	ns	YES	0.44
<i>SEPTIN3</i>	1.91	0.90	-1.04	NO	ns	ns	NO	NA
<i>SLC28A3</i>	3.26	-0.31	1.19	YES	**	***	YES	0.11

SUSD3	1.40	0.29	0.76	YES	ns	ns	YES	-0.81
TC2N	1.35	0.35	0.07	NO	ns	ns	YES	-1.45
THRSP	1.28	0.39	0.56	YES	***	ns	NO	-0.03
TMC5	1.33	0.48	0.94	YES	ns	**	NO	-1.00
VNN2	3.65	0.71	0.46	YES	****	ns	NO	0.64

Table 4-2. Expression of IL6 gene signature genes in SUM44PE RNAseq data, ER+ tumours in the TCGA, ILC CAF RNAseq13 and ILC laser capture microdissection dataset13. Genes highlighted in orange are part of the ER+ IL-6 dependent gene signature from Casneuf *et al*, 2016¹⁵⁸, and those in yellow are not reported to have STAT3 bound to their promoters in breast cancer cell lines²⁷⁹. SUM44PE RNAseq data, log fold change (logFC) between CM treated cells vs control and CM + anti-IL-6 treated cells vs control. TCGA RNAseq dataset - ER+ IDC (n=563) and ER+ ILC (n=191) tumours, logFC of expression ILC/IDC, unpaired two-tailed multiple t-tests with BKY correction in Graphpad Prism. Pearson's correlation of each gene with *IL6* (RNAseq n=191) and STAT3 pY705 (RPPA n=147) in ER+ ILC tumours, p<0.05 *, p<0.01 **, p<0.001 ***, p<0.0001 ****, (neg) = negative correlation. If a gene is expressed by ILC CAFs in the CAF RNAseq dataset from Chapter 3, 'yes' is included, 'no' if not detected. Laser capture microdissection data (LCM) from 17 ILC tumours separated into the stroma (TS) and epithelial (TE) compartments¹³. logFC<0 = higher expression in the epithelium. A literature review of the IL6GS genes can be found in Appendix 4 and the results of differential gene expression analysis can be found in [Esme Bullock Thesis Documents](#).

Chapter 4.3 Discussion

These data identify a CAF-driven paracrine IL-6/STAT3 signalling pathway and downstream gene expression that is enriched in Invasive Lobular Carcinoma. Paracrine IL-6/STAT3 signalling has been well characterised previously in other subtypes of breast cancer, having been shown to promote therapy resistance, tumour cell metastasis and cancer stem cell (CSC) formation, but has not previously been interrogated in ILC.

4.3.1 ILC CAF cell lines used for experiments

Three different ILC CAF lines were used for experiments in this Chapter – ED38 for the RPPA (Figure 4-1), ED2334 for Western blots (Figure 4-5) and functional assays in Chapter 6, and ED26 for the RNAseq experiment (Figure 4-8). All three cell lines were demonstrated by ELISA (Figure 4-6) to secrete IL-6 and induce pSTAT3 activation dependent on IL-6 in their conditioned media, as shown in Figure 4-5D-E and Figure 4-8A.

ED38 CAFs were initially used for the RPPA as previous members of the lab had optimised their growth and had also determined that these had high expression of IGF-1, which was the initial focus of this project. However, ED38 CAFs were isolated from an ER-low HER2+ ILC patient (Table 3-1), so were not representative of the majority of ILC patients, and had lower IL-6 expression than the other CAFs as determined by ELISA and in the CAF RNAseq dataset. We decided to use ED2334 and ED26 for future experiments as these were both isolated from ER+ ILC patients, we had plenty of low passage cells available from both cell lines and both had been included in the FPPA dataset (Figure 3-6 and Table 3-5). In designing the RNAseq experiment, we expected that other components of the CAF CM besides IL-6 could also influence gene expression. Conditioned media from the ED26 CAF line was chosen for the RNAseq experiment as we had the FPPA dataset available, which could have indicated any other CAF secreted factors driving gene expression in SUM44PE cells.

4.3.2 CAF CM driven signalling pathway protein changes

4.3.2.1 IGF-1/PI3K pathway activation

As well as activating STAT3, CAF CM treatment of ILC cells at 30 min induced significant changes in the PI3K and MAPK pathways (Figure 4-1).

The PI3K pathway is hyper-activated in ILC via activating mutations in pathway genes³ and through autocrine signalling driven by E-cadherin inactivation^{52,53}, making it a promising therapeutic target in ILC. The RPPA showed significant activation of a number of proteins in the PI3K pathway, such as p-mTOR, following CM stimulation of SUM44PE, demonstrating that the pathway can also be activated by paracrine signalling from the tumour microenvironment.

ILC CAFs also secrete high levels of IGF1 (Figure 3-6) and the RPPA demonstrated increased pIGF1R on CAF CM stimulation of SUM44PE cells, which can promote activation of the downstream PI3K pathway. Previously our group has shown that CAFs express a high level of *PAPPA*, a secreted protease that cleaves IGF binding protein 4 (IGFBP4) and -5, leading to increased local bio-availability of IGF1 in the ILC tumour microenvironment¹³. Interestingly, IL-6 in CAF CM significantly downregulated *IGFBP5* in the tumour cells (logFC (CM/control) = -1.31, FDR=0.019). This suggests that ILC CAFs secrete high levels of two pro-tumorigenic factors (IL-6 and IGF1) and that IL-6 may increase signalling through the IGF1/PI3K pathway by reducing the amount of IGFBP5 expressed by the tumour cells. IL-6 can also directly activate the PI3K pathway, so potentially targeting IL-6 alongside the IGF1/PI3K pathway may improve the efficacy of PI3K pathway drugs in ILC.

4.3.2.2 STAT3 activation

The RPPA also showed CAF CM driven STAT1 Y701 and STAT3 Y705 phosphorylation at 30 minutes with only STAT1 pY701 returning to control levels at 24 hours (Figure 4-1). This was IL-6 dependent and it is known that in comparison to STAT3 activation, IL-6 activation of STAT1 is fairly transient^{148,257}. The interferon (IFN) alpha and gamma response pathways are two of the most significantly enriched Hallmark gene sets in CAF CM stimulated cells (Figure 4-9A and B) and primarily signal through STAT1 activation. IL-6 dependent STAT1 activation may be driving activation of IFN gene sets. However, STAT1 and STAT3 recognise very similar DNA sequences so it is more likely both STAT1 and STAT3 contribute to the enrichment of

these pathways. The Hallmarks IL6-JAK-STAT3 and both IFN pathways also contain many of the same genes, most notably *IL4R* as the most enriched gene in all 3 pathways.

Unlike pSTAT1, STAT3 pY705 remained high at 24 hours. Although pSTAT3 is known to drive IL-6 expression in a feed forward mechanism, the SUM44PE cells do not express IL-6, even after 24 hours of CAF CM stimulation. This suggests other mechanisms downstream of IL-6 may sustain the STAT3 activation. A number of genes in the IL6GS have been shown to also promote STAT3 activation, potentially promoting a feed-forward loop maintaining STAT3 activation.

CDC7, a gene in the IL6GS, encodes for a cyclin-dependent kinase that promotes the activation of the MCM complex, leading to activation of DNA replication. A siRNA-screen of kinases in HEK293T cells showed that knockdown of *CDC7* lead to a significant reduction of STAT3 transcriptional activity, but not STAT1 activity²⁸⁰. This could account for the sustained activation of STAT3 at 24 hours of CAF CM stimulation whilst STAT1 pY701 levels are returned to baseline in Figure 4-1. However, drugs targeting *CDC7* very slowly decreased STAT3 activity, which suggests that *CDC7* is not directly phosphorylating STAT3 and may support its sustained activation via transcriptional mechanisms. Similarly, the IL6GS gene *LRG1*, a secreted glycoprotein most well-known to drive angiogenesis in cancer²⁸¹, has been reported to both be activated by IL-6/STAT3 signalling^{274,282} and to activate STAT3 phosphorylation^{275,283}, again accounting for sustained STAT3 activation via a positive feedback loop. Expression of *AKR1C1*, an aldo/keto reductase family enzyme, has also been shown to be driven by IL-6 and is part of an ER+ breast cancer IL-6 signature¹⁵⁸. *AKR1C1* has been demonstrated to promote STAT3 activation and tumour cell migration in non-small cell lung carcinoma cell lines²⁸⁴. However, *AKR1C1* was excluded from the IL6GS as its expression remains high with the addition of anti-IL-6 to the CAF CM ($\log_{FC}(\text{CM} + \text{anti-IL6}/\text{control}) = 1.88$), suggesting other factors within the CAF CM can drive its expression alongside IL-6. Interestingly, there was no significant induction of *SOCS3*, the negative regulator of the pathway and classical STAT3 transcriptional target¹⁴⁷.

4.3.3 IL-6 dependent gene expression

RNAseq analysis of SUM44PE cells stimulated with ILC CAF CM identified 110 differentially expressed genes (upregulated = 76 DEGs, downregulated = 34 DEGs, Figure 4-8C). The vast majority of changes in gene expression and significantly enriched pathways were driven by IL-6 within the CAF conditioned media, as demonstrated by the clustering of control and CM + anti-IL-6 treated samples together (Figure 4-8B and C) and overlapping significantly enriched pathways relative to CM treated cells (Figure 4-9A and B).

The IL6GS consists of 43 genes with a more than 2-fold significant increase in expression driven solely by IL-6 in the CAF CM (Figure 4-11 and Table 4-2). Expression of many of these genes and the IL6GS score is significantly higher in ILC tumours compared to IDC tumours and the score correlates with stromal content in ILC but not IDC. This suggests that this paracrine IL-6/STAT3 pathway and downstream gene expression identified *in vitro* is enriched in, and relevant to, ILC patient tumour biology. STAT3 has been shown to at least weakly bind to the promoters in of most (38/42) IL6GS genes in breast cell line ChIPseq experiments (chip-atlas.org). However, IL-6 can activate other pathways, such as the PI3K and MAPK pathways (as seen in Figure 4-5), leading to activation of a wide range of different transcription factors, including the IL6GS members *PKNX2*, *BCL3* and *CEBPD*, and resultant gene expression. Details of all the genes in the IL6GS, their roles in breast cancer and relationship with IL-6/STAT3 signalling can be found in Appendix 4.

Only five genes (*S100A8*, *AKR1C1*, *S100A7*, *AKR1C2* and *CLEC7A*) were differentially expressed between control cells and CAF CM + anti-IL-6 treated cells (Figure 4-8D), though expression of these was reduced in comparison to CM only treated cells for all except *CLEC7A*. This suggests that although IL-6 is mostly responsible for driving their expression, other CAF secreted factors also contribute to their expression. *S100A7* has been shown to be induced in MCF7 and T47D cell lines by IGF1 signalling, also via STAT3²⁸⁵, and as ILC CAFs secrete high amounts of IGF1 as well as IL-6 (Figure 3-6), this could account for sustained high levels of expression.

4.3.3.1 IL-6 driven suppression of pro-proliferative gene sets

IL-6 has been shown to promote primary tumour growth in *in vivo* models of TNBC^{286,287}, HER2+ breast cancer²⁸⁸ and IDC^{158,165,287}. CAF CM stimulation of SUM44PE cells increased expression of pro-proliferative signalling proteins, such as phospho-ERK1/2 and phospho-Rb (Figure 4-1), a marker of cell cycle progression. However, IL-6 in CAF CM lead to decreased expression of genes involved in the cell-cycle, proliferation and oxidative phosphorylation in the SUM44PE cells (Figure 4-10 and Figure 4-11) and at 24 hours of CAF CM stimulation, pRb and pERK1/2 were reduced back to similar levels as control cells. This suggests that although CAF secreted factors may initially promote proliferation, IL-6 in CAF CM may reduce proliferation in the longer term. This paracrine IL-6 driven suppression of proliferation-related genes and signalling could account for research showing that high stromal content is associated with longer DFS in ILC^{3,143} and could potentially suggest that CAF secreted IL-6 may contribute to the long latency periods seen in ILC⁸.

IL6 expression was significantly decreased in the proliferative subtype of ILC tumours in the TCGA dataset, in comparison to the reactive-like and immune-related subtypes (Figure 4-9E-G). The proliferative subtype is the most cancer cell dense subtype of ILC, associated with significantly worse disease-specific survival³. A recent study has also shown high stromal content to be associated with better survival and lower Ki67 scores in ER+ ILC¹⁴¹. This suggests that CAFs in the ILC TME may suppress proliferation. CAF CM stimulated SUM44PE cells are enriched for reactive-like and immune-related ILC subtype signatures, supporting a downregulation of proliferative subtype related gene expression. The reactive-like subtype has a very high CAF content³ and the enrichment of the reactive-like gene signature in CAF CM treated cells supports the activity of the CAF-driven IL-6/STAT3 pathway identified *in vitro* in reactive-like ILC tumours.

CLCA2 is upregulated by IL-6 within the CAF CM (logFC=1.65, FDR=0.04) and has been shown to inhibit the proliferation and migration of breast cancer cells,^{289,290} which could account for the IL-6 dependent suppression of cell cycle dependent genes. In normal breast cells *CLCA2* is localised to adherens junctions and suppresses EMT through direct binding to β -catenin and decreasing expression of its EMT-promoting target genes²⁹¹. However, as ILC cells do not express E-cadherin, β -catenin

expression is already reduced and is localised to the cytoplasm. It would be interesting to determine the subcellular localisation of CLCA2 in ILC and whether loss of E-cadherin alters the tumour suppressive phenotype of CLCA2.

SERPINB5, also known as maspin, is significantly upregulated by CAF CM, although its expression is not solely IL-6 dependent ($\log_{2}FC$ (CM/control) = 4.43, $\log_{2}FC$ (CM+anti-IL6/control) = 1.35). Nuclear localisation of maspin is associated with improved survival in breast cancer and inhibition of proliferation whereas cytoplasmic maspin is associated with poorer outcomes^{292,293} and drives EMT and invasion in the TNBC MDA-MB-231 cell line²⁹⁴. Interestingly, in non-transformed MCF10A cells, maspin accumulated in the nucleus on EGF stimulation dependent on downstream STAT3 signalling, and also in sparsely seeded cells or those with disrupted adherens junctions²⁹⁵. This suggests that maspin is most likely to be localised to the nucleus in the E-cadherin negative ILC cell lines and could be contributing to the suppression of proliferative genes seen in the RNAseq data, although validation of the subcellular localisation would need to be assessed.

Although IL-6 appears to suppress gene sets involved in the cell cycle in SUM44PE cells, a number of the CAF CM induced genes have previously been shown to promote proliferation in breast cancer, such as *AKR1B10*²⁹⁶, *FAM83A*²⁹⁷, *MUC5B*²⁹⁸ and *MUCL1*²⁹⁹ (see Appendix 4). IL-6/STAT3 signalling is highly context dependent, and it would be interesting to see whether the IL-6 driven cell cycle gene suppression is ILC cell line specific or if the previously published phenotypes of genes such as *SERPINB5* and *CLCA2* differ in ILC cell lines due to E-cadherin inactivation.

Generally in breast cancer, IL-6/STAT3 signalling is considered to be pro-tumorigenic, driving increased primary tumour growth in *in vivo* models of TNBC and IDC^{121,165}. However, some studies have identified high expression of pSTAT3^{181,193,194} or IL-6 as positive prognostic indicators³⁰⁰, and many of the IL-6 upregulated DEGs in the SUM44PE cells have been associated with good prognosis in breast cancer. For example, *FGD3* and *SUSD3* are co-regulated in breast cancer and are positive prognostic indicators in ER+ and ER- breast cancer^{301–303}, with *SUSD3* being very significantly upregulated in AI responsive tumours³⁰⁴. Similarly, high protein expression of CEBP δ (*C/CAAT enhancer binding protein delta*, $\log_{2}FC$ = 1.22), a transcription factor that is well known to be induced by IL6/STAT3¹⁵⁸, has been shown

to be associated with positive prognosis when coupled to IL-6 expression in ER+ BC. *CEBPD* overexpression also suppressed MCF7 cell growth and invasion via downregulation of the Slug EMT transcription factor³⁰⁰. However, both *CEBPD* and *SUSD3* have also been shown to drive migration³⁰⁴, metastasis³⁰⁵ and CSC³⁰⁶ formation in breast cancer cells, suggesting a context dependent function for these genes. Furthermore, none of these survival analyses delineated ILC and IDC from the ER+ groups of tumours and many entirely excluded ILC tumours. Therefore it is possible that *CEBPD*, *FGD3* and *SUSD3*, which are significantly more highly expressed in ILC tumours (Figure 4-11B), may have a divergent role in ILC.

4.3.3.2 IL-6 regulation of metabolism related genes

ILC tumours are known to have lower rates of both glycolysis and oxidative phosphorylation³⁰⁷ than IDC tumours and STAT3 has been shown to drive the switch from aerobic respiration to glycolysis³⁰⁸, known as the 'Warburg effect', in breast cancer cell lines. GSEA identified an IL-6 dependent suppression of oxidative phosphorylation (Figure 4-9A and B) and 10 GO BP gene sets annotated as 'respiratory electron transport' (Figure 4-10) in ILC tumour cells, suggesting that IL-6/STAT3 signalling could potentially contribute to the low levels of oxidative phosphorylation seen in ILC tumours.

The IL6GS members *AKR1B10*, *ACOX2*, *BBOX1*, *DGAT2* and *THRSP* are involved in fatty acid metabolism (see Appendix 4). Recently, ILC cell lines and tumours have also been shown to be enriched for fatty acid metabolism³⁹, and JAK/STAT3 promotion of fatty acid oxidation has been shown to drive breast CSC formation³⁰⁹. Interestingly, long-term estrogen deprived ILC cell lines were shown to increase expression of fatty acid metabolism genes, which drove resistance to endocrine-targeted therapy *in vitro*⁶¹. These IL-6 driven changes in metabolism-related genes may represent another unique and potentially targetable feature of ILC tumours compared to IDC tumours that could be exploited in the clinic, although validation of the functional consequences of these gene expression changes is needed.

4.3.3.3 IL-6 and therapy resistance

IL-6/STAT3 has been implicated in resistance to all types of breast cancer therapy. CAF CM downregulated E2F gene sets (Figure 4-9) in an IL-6 dependent manner, a biomarker for pathologically complete response to chemotherapy in breast cancer³¹⁰.

ILC tumours are characterised by slow growth and poorer response to chemotherapy than IDC^{17,19}. This phenotype is primarily driven by ILC tumour cell intrinsic features⁴⁷. However, IL-6/STAT3 signalling has been shown to promote chemotherapy resistance in breast^{311,312} and other types of cancer³¹³, and these data suggest that paracrine IL-6 signalling from CAFs could potentially contribute to this slow growing chemotherapy resistant ILC phenotype.

CDK4/6 inhibitors (CDK4/6i) have recently been approved for treatment of metastatic breast cancer patients, and have improved disease free survival in ILC patients (reviewed in²²). However, as with all cancer therapies, resistance can develop, and mTOR³¹⁴, IFN³¹⁵ and STAT3¹⁷⁵ signalling have been implicated in driving this resistance. All three of these pathways are activated by CAF CM in the SUM44PE cells (Figure 4-1 and Figure 4-9A and B). This suggests that paracrine IL-6 signalling may contribute to resistance to CDK4/6 inhibitors in ILC, though further investigation is necessary. It would be interesting to see whether high levels of *IL6* and pSTAT3, or a downstream activated gene, correlate with worse outcomes for CDK4/6i treated ILC patients.

IL-6 also led to downregulation of the Estrogen Receptor (*ESR1/ERα*) (logFC (CM/control) = -0.73, p = 0.002, FDR=0.15) and downstream gene expression (Figure 4-9), potentially reducing the effectiveness of ER-targeted therapies commonly used in the treatment of ILC. IL-6 has been shown to induce endocrine resistance in the IDC cell line ZR-75-1 via ERα proteasomal degradation¹³¹ and to promote endocrine therapy resistance via STAT3 activation and formation of breast CSCs^{83,163}. Interestingly, the STAT3 pathway was also significantly enriched in long-term estrogen deprived (endocrine therapy resistant) SUM44PE and MDA-MB-134VI ILC cell lines compared to parental cells⁶¹, further supporting a role of the pathway in driving resistance in ILC. IL-6 in CAF CM also increases expression of a number of genes, such as *AKR1C1* and *BCL3*, that are known to drive tamoxifen resistance in ILC³¹⁶ and other breast cell lines⁷³. *IGFBP5* expression was also decreased in SUM44PE cells by IL-6, and has been shown to be suppressed in MCF7 cells by CAF secreted factors, leading to increased proliferation in the presence of fulvestrant dependent on BCL3 induction¹³³. Together, these studies suggest that IL-6 signalling may drive a slow growing chemo- and endocrine therapy-resistant phenotype in ILC. Further

analysis of the effect of IL-6 signalling on estrogen signalling in ILC can be found in Chapter 5.

4.3.3.4 Epithelial to mesenchymal transition and migration

IL-6/STAT3 signalling has been implicated in EMT and metastasis in a wide range of cancers^{250,318} and breast cancer subtypes^{160,165}. CAF CM treated SUM44PE cells were significantly enriched for the Hallmark EMT gene set (Figure 4-9A and B) and for a cluster of GO BP gene sets annotated as 'migration neutrophil chemotaxis', involving pro-migratory gene sets such as 'cell chemotaxis' (NES=2.2, FDR=7.69x10⁻⁴). The upregulation of these gene sets was dependent on IL-6 in the CAF CM, suggesting IL-6 drives a migratory phenotype in these ILC cells.

A number of the IL6SG genes have been shown to drive migration and metastasis in cancer. For example, STAT3 induced expression of LRG1 promotes invasion and metastasis in colorectal cancer models²⁸² and LRG1 also induces metastasis of melanoma cells via STAT3 activation³¹⁹. *AKR1B10* has been shown to behave as a metastasis enhancer in breast cancer via its ability to promote fatty acid oxidation³²⁰, although high expression was only associated with poor distant metastasis free survival in ER- and HER2+ BC, not in ER+ BC. Three genes encoding for mucins or mucin-like proteins, *MUC1*, *MUC5B* and *MUCL1*, were significantly upregulated by CAF CM stimulation. Mucins are glycoproteins that are often deregulated in cancer, with *MUC1*, *MUC5B* and *MUCL1* all reported to have higher expression in breast cancer than normal tissue³²¹. *MUC1* has been identified as a STAT3 target gene in breast cancer and correlates with poor BC specific survival⁸³. *MUC5B* has been shown to suppress *IL6* expression in MCF7 cells³²², so the expression of *MUC5B* could account for the lack of feed-forward IL-6 induction in SUM44PE cells. *MUC5B* and *MUCL1* (also known as Small Breast Epithelial Mucin/SBEM) have both been shown to drive invasion *in vitro*^{298,323,324}, with *MUCL1* also promoting proliferation^{299,324}, and being associated with poor survival and metastasis in breast cancer patients^{325–327}.

Further investigation of IL-6 driven EMT signalling and migration is assessed in Chapter 6.

4.3.3.5 Inflammation and immune signalling

IL-6/STAT3 plays an essential role in the immune system and inflammation^{271,272}. GSEA identified a very strong IL-6 dependent induction of inflammatory and immune-related signalling in the SUM44PE cells (Figure 4-9A and B and Figure 4-10), and a high expression of *IL6* in the immune related subtype of ILC.

The immune signalling signatures enriched in CAF CM treated SUM44PE cells are dominated by the IL-6 dependent induction of *IL4R* (logFC=1.66, FDR=7.86x10⁻¹⁰), the IL-4 receptor. *IL4R* is the most highly enriched gene in the IFN α and γ signatures in the Hallmarks gene sets (Figure 4-9A and B), and is present in 91/183 of the 'immune response activation' cluster of GO BP gene sets (Figure 4-10). *IL4R* is significantly more highly expressed in ILC than IDC tumours in the TCGA dataset (Figure 4-11). *IL4R* expression in the TME could also contribute to its enrichment in ILC tumours as *IL4R* is also very highly expressed by ILC CAFs (Table 4-2) and by pro-tumorigenic M2 macrophages³²⁸, which are present at greater densities in ILC than IDC tumours¹⁴. Many of the gene sets in the 'immune response activation' cluster that include *IL4R* were defined by analysing gene expression in immune cells rather than epithelial cells. *IL4R* expression on ILC tumour cells could impact immune cell phenotypes and inflammatory signalling, but further work is needed to validate this.

IL4R and *IL13RA1*, which is also induced by IL-6 in the CAF CM, form the heterodimeric type II receptor for the IL4 and IL13 cytokines, neither of which are expressed by SUM44PE cells or ILC CAFs but are primarily expressed by CD4⁺ T-cells and NK cells²⁷⁰. Therefore, in this *in vitro* system, there is no signalling through this receptor. Nevertheless, knockdown of *IL4R* on breast cancer cells has shown to reduce metastasis *in vivo* dependent on host IL4³²⁹. Targeting the receptor has also been shown to deplete the number of immunosuppressive cells in the breast TME³³⁰, with depletion of *IL4R*⁺ M2-macrophages inhibiting breast cancer bone metastasis growth³²⁸. This suggests that targeting *IL4R* may be a potential therapeutic option in ILC. However, *IL4R* is expressed by a wide range of immune cells and is essential for normal macrophage homeostasis, so therapies targeting this receptor may have wide ranging side effects.

CHI3L2 is a chitinase-like protein, exclusively expressed by human cells, and significantly induced by IL-6 (logFC = 3.85, FDR=2.95x10⁻²⁵). It has been reported to

be expressed solely by tumour associated macrophages (TAMs) in breast cancer³³¹. However, this study only looked at a limited number of IDC tumours, so expression in cancer cells in ILC tumours cannot be ruled out. CHI3L2 acts as a chemoattractant for monocytes, and could account for the strong enrichment for immune signatures and the immune-related subtype genes in CAF CM treated cells (Figure 4-9) as well as potentially contributing to the high infiltration of M2-type macrophages in ILC tumours¹⁴. Interestingly, in glioma, CHI3L2 is expressed by both cancer cells and macrophages, and correlates with poor outcome³³². In the same study, recombinant CHI3L2 treatment of CD8+ T-cells promoted their apoptosis, suggesting IL-6 induced CHI3L2 expression in ILC could contribute to an immunosuppressive microenvironment.

CAF CM stimulation led to a very significant increase in expression of a number of S100 family genes (*S100A7*, *S100A7A*, *S100A8* and *S100A9*) in SUM44PE cells. These genes encode for secreted calcium binding proteins and are part of a common ER+ IL-6 gene signature¹⁵⁸ with STAT3 directly binding to the promoters and driving transcription of *S100A8/A9* in breast cells²⁸⁸. Only *S100A9* was included in the IL6GS as the expression of the other S100 family genes remained high on addition of anti-IL6. In breast cancer, autocrine *S100A8* and *S100A9* signalling has been shown to drive proliferation, EMT and metastasis^{333–335}, in part through activation of FAK and YAP signalling³³⁶, and high tumour cell *S100A8* and *S100A9* expression is significantly associated with poorer outcomes across all breast cancer subtypes^{337,338}. Exogenous or macrophage secreted *S100A8* and *S100A9* have also been shown to promote fibroblast activation in models of TNBC³³⁹ and lung fibrosis³⁴⁰. It would be interesting to see whether reciprocal signalling via the S100 proteins from CAF CM stimulated ILC cells induces CAF activation and pro-tumorigenic properties.

S100A7, *S100A8* and *S100A9* play an essential role in promoting inflammation both in infections and cancer (reviewed in^{341,342}) and were included in 40, 36 and 43, respectively, of 183 gene sets in the 'immune activation response' cluster in Figure 4-10. Overexpression of *S100A7/8/9* *in vivo* has been shown to enhance primary tumour growth and metastasis both via tumour cell intrinsic signalling but also by increasing the recruitment of immunosuppressive immune cells^{335,343}. These proteins act as chemoattractants for monocytes and myeloid derived suppressor cells

(MDSCs), which also express high levels of S100A8 and S100A9³⁴¹. This suggests that STAT3 activation and subsequent S100 family expression could form part of crosstalk between tumour cells, CAFs and immune cells in ILC and may contribute to the relatively 'immune cold' ILC microenvironment^{34,344}.

IL-6/STAT3 signalling has been shown to promote both immunosuppression and immune activation in breast cancer, again highlighting the pleiotropic effects of this signalling pathway. In some studies, ILC tumours have been shown to have lower infiltrated lymphocytes than IDC tumours^{14,344} whereas others have identified higher expression of a number of immune signatures³⁰⁷ in the TCGA dataset and high infiltration of M2-macrophages¹⁴ in ILC tumours. A number of the IL-6 induced genes are known to behave as immune cell chemoattractants, such as *LRG1* and *S100A7* inducing neutrophil and macrophage chemotaxis, respectively^{283,335}. It would be useful to assess the correlation between some of the IL6GS genes or the IL6GS score, as well as overall stromal content, with immune cell infiltration by immunohistochemistry in ILC tumours.

This work has been done in the human cell line SUM44PE, and so *in vivo* work is only possible in immunocompromised mice, limiting our ability to assess the effects of genes induced by IL-6 in tumour cells on immune cell infiltration and activity in *in vivo* models. *In vitro* experiments could be carried out to stimulate immune cells with conditioned media from IL-6 stimulated tumour cells or directly by CAF CM, although this reductionist approach would over-simplify the complexity of the immune microenvironment. Any findings in terms of immune signalling need to be validated in ILC mouse model cell lines *in vitro*, such as the *K14-Cre/E-cadherin^{fl/fl}/Tp53^{fl/fl}* (KEP)⁶⁷ or *WAP-Cre/E-Cadherin^{fl/fl}/Pten^{fl/fl}* (WEPTn)^{37,101} models, and their impact on immune cells in the microenvironment assessed *in vivo*.

4.3.3.6 Angiogenesis

A key hallmark of cancer is angiogenesis, with the formation of blood vessels within the tumour providing nutrients, oxygen and an easier means of metastasis for tumour cells. In one study, ILC tumours were found to have a significantly higher microvessel density than IDC tumours and a positive correlation between CAF density and endothelial cell proliferation⁵⁵. In bulk analysis, ILC tumours have also been shown to

have higher expression of *PECAM1*³⁴⁵, the gene encoding for the endothelial marker CD31, than IDC tumours (also seen in TCGA and METABRIC).

GSEA detected significant enrichment for angiogenic and VEGF gene signatures dependent on IL-6 within the CAF CM (Figure 4-9B and Figure 4-10). 10 of the IL6GS genes (*S100A9*, *LRG1*, *IL4R*, *CYP1B1*, *RNF312*, *LRP2*, *LRRC4C*, *HMOX1*, *FGD3* and *BCL3*) were part of at least one gene set in the angiogenesis-related 'embryonic development morphogenesis' cluster (Figure 4-11). *LRG1*, a secreted glycoprotein (logFC = 2.1, FDR=1.62x10⁻¹³) known to be induced by IL-6/STAT3^{274,282}, was included in 12/25 of these gene sets. *LRG1* has been shown to promote angiogenesis *in vitro*, although it has not been assessed in breast cancer driven angiogenesis. Non-small cell lung carcinoma cell exosomes containing *LRG1* induce *VEGFA* expression in endothelial HUVECs via TGFβ signalling³⁴⁶. Autocrine *LRG1* signalling can also induce *VEGFA* expression in pancreatic cancer cells²⁷⁶ and in gastric cancer cells through *LRG1* mediated STAT3 transcriptional activity²⁷⁵. In the future, it would be worth assessing whether conditioned media from CAF CM stimulated SUM44PE cells promotes endothelial cell proliferation and tube formation.

The functional consequences of paracrine CAF signalling are investigated in Chapter 6.

Chapter 4.4 Conclusions

I have identified a paracrine CAF driven IL-6/STAT3 signalling pathway that is enriched in ILC compared to IDC. IL-6 is a pleiotropic cytokine and, as demonstrated by these data, can induce genes involved in both pro-tumorigenic features, such as migration and angiogenesis, and anti-tumorigenic features, such as downregulation of the cell cycle (Figure 4-13).

The consequences of STAT3 activation is highly context dependent, as is the activity of a number of the IL6 induced genes, with some reported to both suppress and promote tumorigenic features in breast cancer, dependent on subcellular localisation, post-transcriptional and translation modifications, breast cancer subtype and cell types. Whilst targeting IL-6 at the top of the pathway is an attractive therapeutic option, not least due to the number of clinically approved therapies targeting IL-6, IL6R and JAKs (see 1.3.3), IL-6 and pSTAT3 have been shown to be associated with both positive and negative outcomes in ER+ breast cancer^{165,181,193,194} so may not be the best biomarker or target. Identifying specific downstream genes driving the migration or angiogenic signatures may offer more targeted benefits for ILC patients, avoiding alleviation of IL-6 dependent inhibition of proliferative gene signatures.

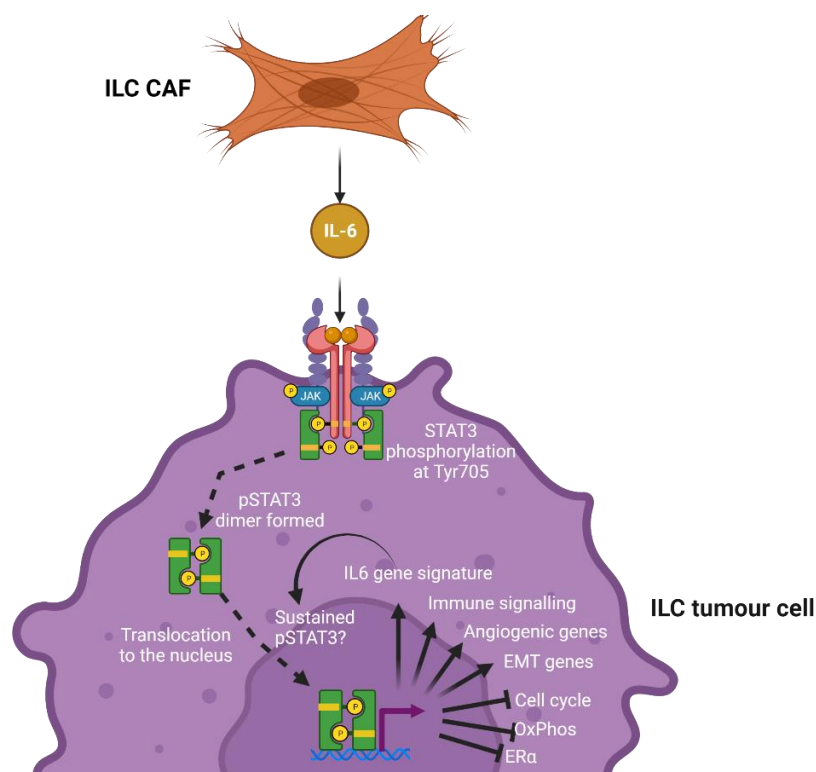


Figure 4-13. Schematic of the paracrine IL-6/STAT3 pathway in ILC. Made with Biorender.org

Chapter 5 Investigating the relationship between Estrogen Receptor signalling and IL-6/STAT3 in Invasive Lobular Carcinoma

Chapter 5.1 Introduction

As roughly 90% of Invasive Lobular Carcinoma (ILC) tumours are Estrogen Receptor Positive (ER+), the main therapeutic strategy for ILC patients is the use of estrogen-targeted therapies (ETs), such as tamoxifen pre-menopause and letrozole, an aromatase inhibitor, post-menopause³³. However, ILC patients have been shown to have a poorer response to tamoxifen than letrozole²⁵ and a poorer response in general to ETs compared to invasive ductal carcinoma (IDC) patients. Multiple studies have also shown that ILC patients and ILC cell lines respond differently to estrogen (E2) and ETs than IDC patients and cell lines^{59,252,347}. Due to the slow growing nature of ILC tumours, patients do not respond well to chemotherapy and have a poorer response than IDC patients,^{19,72} and so understanding mechanisms of resistance to ETs is important for improving ILC patient care and outcomes.

Work in the previous chapter identified a paracrine IL-6/STAT3 pathway enriched in ILC tumours. The IL-6/STAT3 pathway has been implicated in resistance to chemotherapy^{167,182,313}, immune checkpoint inhibitors³⁴⁸, HER2-targeted therapies^{181,183–185}, CDK4/6 inhibitors¹⁷⁵ and endocrine targeted therapies^{83,131,154} in breast cancer cell lines and patients, induced both by signalling from the tumour microenvironment (TME) as well as via tumour cell autonomous mechanisms. Importantly, active STAT3 (pY705) has been shown to be elevated in ET resistant breast cancer cell lines and patient samples^{83,154} and to directly promote resistance via multiple mechanisms, such as driving cancer stem cell (CSC) formation^{83,163,167,306}, and down regulating Estrogen Receptor (ER α) expression^{131,135,136}. Interestingly, in some cell studies the IL-6/STAT3 pathway has been shown to have a direct effect on ER α signalling, whereas Siersbæk et al described the pathways as ‘functionally uncoupled’ in the T47D cell line, with pSTAT3 forming a complex with ER α and FOXA1 to regulate a distinct subset of genes from the ER α -FOXA1 complex alone¹⁶⁵.

The TME and cancer-associated fibroblasts (CAFs) are instrumental in promoting therapy resistance in many cancers, and in breast cancer have been shown to promote ET resistance via a multitude of mechanisms. For example, paracrine IL-6 signalling^{131,169,185} and exosomal miRNA induced by IL-6/STAT3 activity in CAFs^{135,136} have been demonstrated as mechanisms of ET resistance in models of IDC. The tumour microenvironment of ILC tumours is very CAF-dense⁵ and although ILC tumour cell autonomous mechanisms of estrogen signalling and therapy resistance have previously been investigated^{59,61,62,252,347,349,350}, no studies have specifically explored the influence of CAFs and paracrine signalling in ILC E2 signalling and ET resistance. As the paracrine IL-6/STAT3 pathway identified in the previous chapter has been reported to promote resistance to ETs in breast cancer and is enriched in ILC, CAF derived IL-6 signalling may contribute to the differences in ET response seen between ILC and IDC patients.

The RNAseq data analysed in the previous chapter identified an IL-6 dependent suppression of *ESR1* expression and early estrogen signalling gene sets in SUM44PE cells stimulated with CAF conditioned media (CM). This chapter begins to interrogate those transcriptional changes and investigate the relevance of this signalling in ILC biology.

Chapter 5.2 Results

5.2.1 IL-6 stimulation of SUM44PE cells leads to downregulation of *ESR1* and ER α signalling

The RNAseq experiment analysed in Chapter 4 showed that CAF conditioned media (CM) stimulation of SUM44PE cells led to an IL-6 dependent decrease in expression of *ESR1* (Figure 5-1A), the gene encoding the Estrogen Receptor (ER α). There is a weaker correlation between *ESR1* and ER α expression in ILC tumours than IDC tumours³⁴⁷ and ILC cell lines regulate ER α expression differently from IDC cell lines post-translationally, so it is important to determine if these transcriptional changes are carried over to the protein level. The ILC cell lines, SUM44PE and MDA-MB-134VI, were stimulated with IL-6 in complete media for 1 week and ER α expression was detected by immunofluorescence (Figure 5-1B-E). This showed a significant reduction in nuclear ER α expression in both cell lines when treated with IL-6 compared to those in basal conditions. Interestingly, in unstimulated SUM44PE cells, ER α was almost entirely localised to the nucleus in almost all cells whereas MDA-MB-134VI cells showed a more diffuse expression pattern.

Work by other members of the lab (not shown) has demonstrated that IL-6 stimulation causes a significant reduction in *ESR1* expression in both ILC cell lines by RT-PCR as well as a decrease in ER α protein levels after 48 hours of IL-6 stimulation by western blot. Although IL-6/STAT3 signalling has been shown to be interlinked with estrogen signalling in breast cancer¹⁶⁵ and contribute to endocrine resistance in multiple IDC models, IL-6 treatment does not lead to decreased ER α expression in the IDC cell lines, MCF7 and T47D^{131,158}. This data demonstrates that IL-6 promotes downregulation of the Estrogen Receptor at the RNA and protein level in two ILC cell lines, supporting a relationship between these two pathways in ILC.

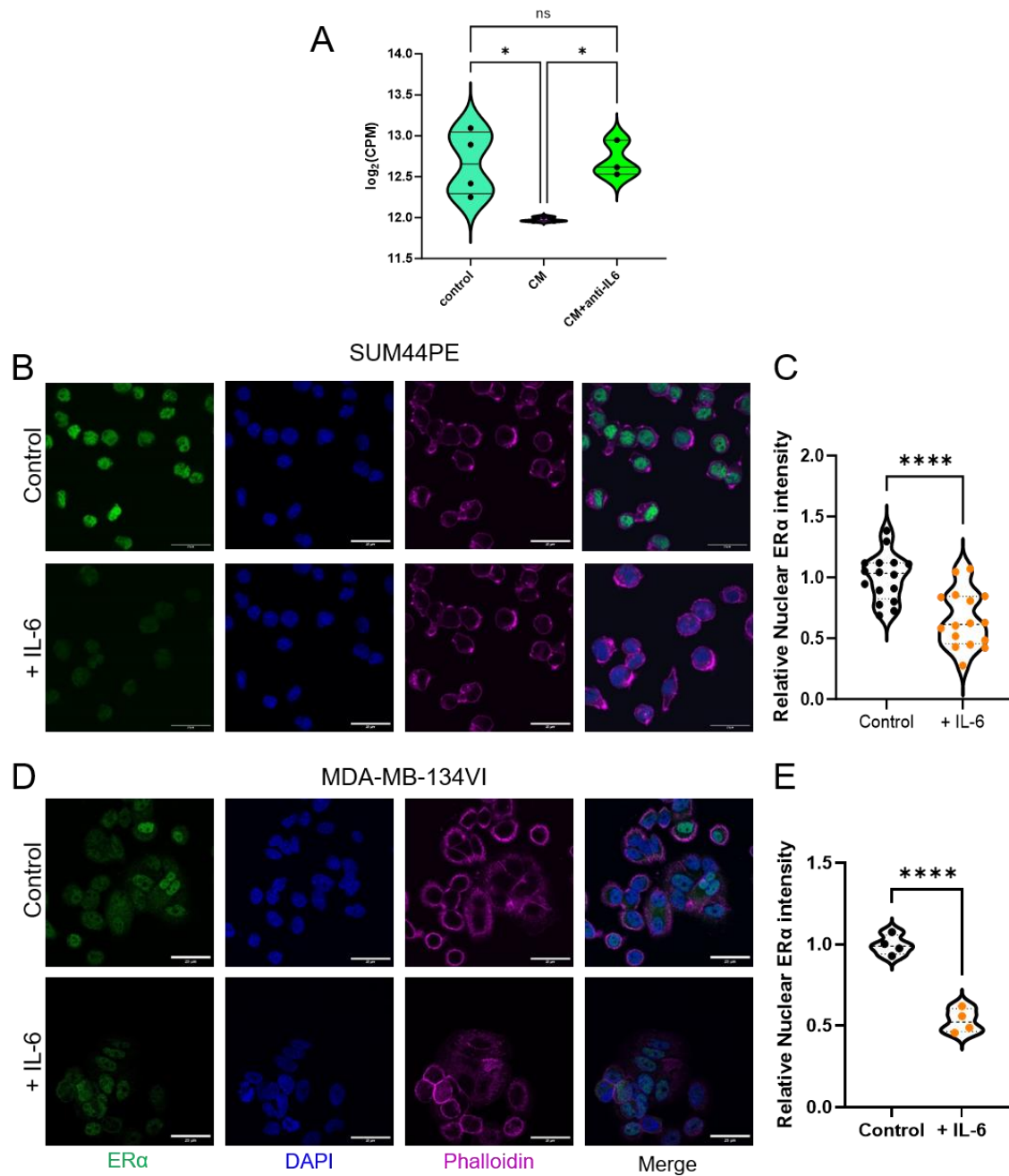


Figure 5-1. IL-6 stimulation of ILC cell lines leads to downregulation of the Estrogen Receptor at the RNA and protein level. (A) Expression of *ESR1* from the RNAseq data analysed in Chapter 4, one-way ANOVA with Tukey's multiple comparison test, adjusted $p < 0.05$ *. Representative images of (B) SUM44PE cells and (D) MDA-MB-134VI cells in complete media (top) or in complete media + IL-6 (10 ng/ml, bottom) for 1 week (media replenished every 2 days), showing ERα expression (green), nuclei labelled with DAPI (blue) and actin with phalloidin (magenta). Images at 60X magnification, scale bar showing 25 μ m. Quantification of relative nuclear localisation (mean gray value normalised to mean of control cells) of ERα in (C) $n=3$ with 5-6 random FoVs per condition and (E) one experiment with 4 random FoVs, each point in (C) and (E) representing the mean relative intensity of one FoV normalised to the mean of the control. Unpaired t-test in Graphpad Prism, $p < 0.0001$ ****.

5.2.2 IL-6 and ESR1 negatively correlate with each other in ER+ breast cancer tumours and cell lines

To determine if the findings *in vitro* were relevant to ILC tumours, the relationship between *IL6* and *ESR1* pathway components was examined in ER+ ILC and ER+ IDC tumours in the TCGA dataset (Figure 5-2). In both ILC and IDC tumours, there was a significant negative correlation between *IL6* and *ESR1* expression as well as with *FOXA1*, a pioneer transcription factor that supports ER α transcriptional activity⁵⁹. However, these correlations were stronger in ILC tumours despite fewer ILC than IDC samples in the TCGA dataset. *IL6* also negatively correlates with ER α protein in ER+ ILC and IDC tumours in the TCGA RPPA dataset, suggesting this relationship is maintained at the protein level, as we saw *in vitro*. However, this correlation and the correlation between *ESR1* and ER α is weaker in ILC tumours than in IDC tumours. This may reflect the reported discord between *ESR1* mRNA expression and ER α protein levels in ILC demonstrated by Sreekumar *et al*³⁴⁷.

The RNAseq analysis in Chapter 4 identified an IL-6 dependent gene signature (IL6GS) that was significantly more highly expressed in lobular tumours (see Figure 4-10 and Appendix 4). When a score was calculated for each ER+ tumours based on IL6GS gene expression in the TCGA RNAseq dataset, this significantly and positively correlated with *IL6* expression (Figure 4-12 and 5-2). Interestingly, the IL6GS score also significantly negatively correlates with *ESR1* and *FOXA1* expression in both ILC and IDC tumours, with these negative correlations being stronger in ILC tumours. However, there was no correlation between ER α expression and the IL6GS score in ILC tumours, whereas a strong negative correlation was seen between the IL6GS and ER α in IDC tumours.

This data suggests a relationship between *IL6* and *ESR1* in ER+ ILC tumours as well as in cell lines. The strong relationship at the transcriptional level may indicate that IL-6/STAT3 signalling is modifying *ESR1* transcription in ILC but having a limited effect on post-translational regulation. Further work is needed to understand how IL-6/STAT3 signalling is modifying expression of the Estrogen Receptor mechanistically in ILC and if this differs from IDC.

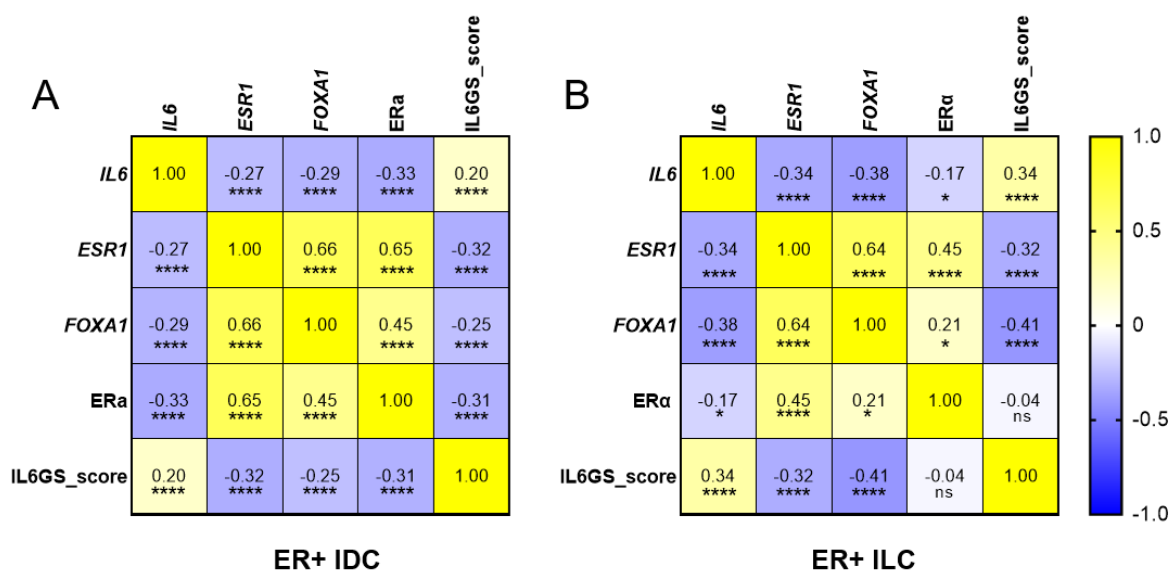


Figure 5-2. IL6 and the IL6 gene signature negatively correlate with ESR1 and FOXA1 in ER+ tumours. Correlation matrices of *IL6*, *ESR1* and *FOXA1* expression from the TCGA RNAseq dataset, *ERα* expression from the TCGA RPPA dataset and the IL6 gene signature (IL6GS) score calculated for each tumour based on expression of IL6GS genes in the TCGA RNAseq dataset (see Chapter4) in (A) ER+ IDC tumours (RNAseq n = 554, RPPA n = 420) and (B) ER+ ILC tumours (RNAseq n = 191, RPPA n = 143). Pearson's r shown, calculated in Graphpad Prism, p<0.0001 ****, p<0.05 *, ns = not significant, positive correlations indicated by yellow coloured box and negative correlations by blue boxes.

5.2.3 CAF CM suppresses estrogen signalling and upregulates endocrine therapy resistance signatures

Multiple studies have shown that response to estrogen signalling and estrogen-targeted therapies is unique in ILC^{59,62,252,349,350}, with the Oesterreich group identifying lobular-specific estrogen response genes²⁵². Downregulation of the Estrogen Receptor and subsequent downstream signalling has been shown to be a mechanism of ET resistance¹³¹. Gene set enrichment analysis (GSEA) of our SUM44PE RNAseq dataset (analysed in Chapter 4) was carried out comparing expression of estrogen signalling and ET resistance signatures between CAF CM treated cells and Control or CAF CM + anti-IL-6 treated SUM44PE cells (Figure 5-3 and descriptions in Table 5-1).

Nardone *et al* recently identified the pioneer factor and ER α binding partner FOXA1 as regulating an ILC-specific chromatin state that contributes to the unique estrogen signalling seen in ILC cells⁵⁹. The ILC-specific FOXA1 target gene set is very significantly enriched in control compared to CAF CM treated SUM44PE cells (NES = -2.35, FDR = 0.00). This suppression of ILC-specific FOXA1 target genes appears to be IL-6 dependent as CAF CM + anti-IL-6 treated cells were also significantly enriched for this gene set compared to CAF CM treated cells (NES = -2.78, FDR = 0.00).

CAF CM treatment also significantly repressed the generic Hallmark 'estrogen response early' gene set²⁰⁹ and the ILC-specific E2-induced genes²⁵² relative to control cells. Whilst the Hallmark early estrogen gene set was also significantly enriched in CAF CM + anti-IL-6 treated cells, there was no significant difference in enrichment for the ILC-specific E2 gene set. This suggests that other factors in the CAF CM may contribute to the suppression of these genes.

ILC tumours have been shown to be more responsive to letrozole treatment than tamoxifen treatment and this benefit is more pronounced than in IDC tumours²⁵. Selli *et al* investigated the transcriptional differences between ER+ tumours that responded to or were resistant to letrozole treatment, including 12 ILC tumours (responsive n = 8, resistant n=4)⁶⁹. GSEA showed that control SUM44PE cells were significantly enriched for genes downregulated in letrozole treated resistant tumours compared to responsive tumours (Figure 5-3) and CAF CM treated cells were enriched for upregulated genes. CAF CM treated cells were also significantly enriched for letrozole

resistance genes compared to cells treated with CAF CM and anti-IL-6, demonstrating that IL-6 within the CAF CM is driving expression of these letrozole resistance related genes.

This analysis demonstrates that IL-6 and potentially other ILC CAF secreted factors can influence generic and ILC specific estrogen related gene expression and may promote ET resistance. This implies that paracrine IL-6/STAT3 signalling could potentially contribute to the differences in ET response for ILC and IDC patients. Further work is needed to validate the functional consequences of these gene expression changes.

Details of the gene sets used in this analysis can be found in Table 5-1 and Appendix 3.

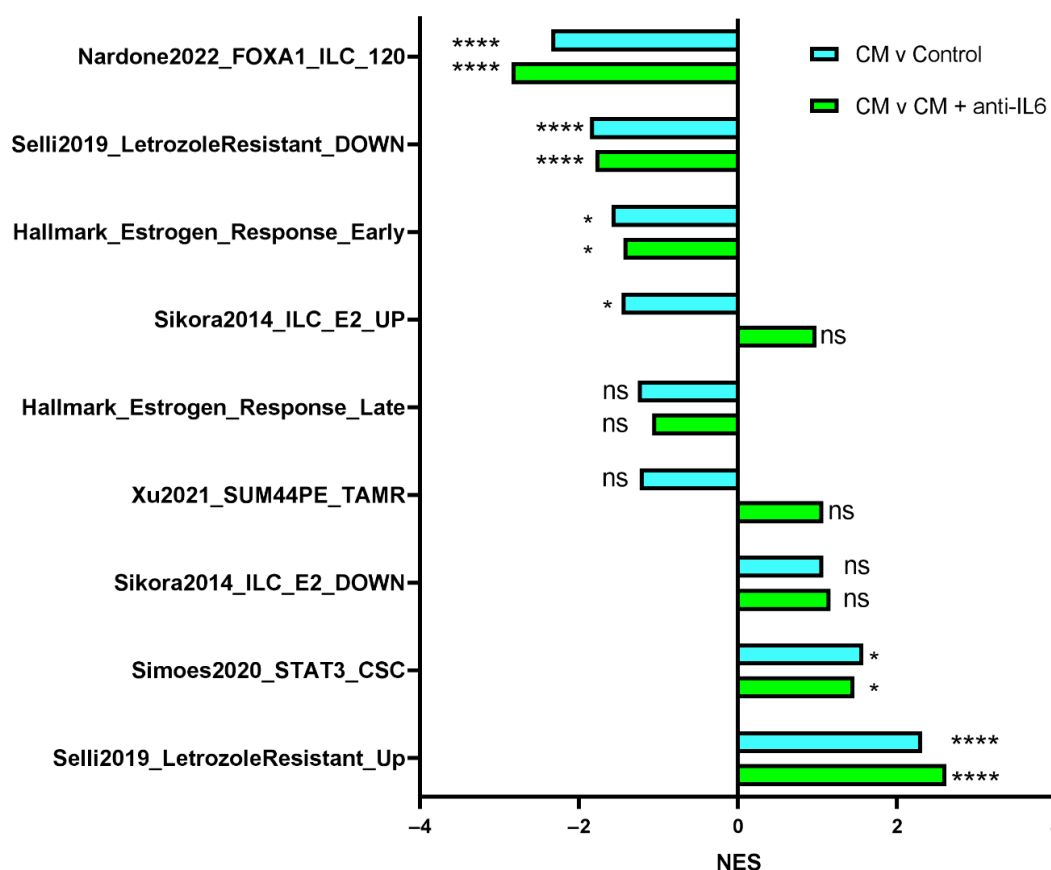


Figure 5-3. CAF CM suppresses generic and ILC-specific estrogen signalling and upregulates gene sets involved in endocrine therapy resistance. Gene set enrichment analysis for estrogen signalling^{59,209,252} and endocrine therapy resistance^{69,316} gene sets, showing normalised enrichment score (NES) for CAF CM treated cells compared to control samples (blue bars) or CAF CM + anti-IL-6 samples (green bars). NES>0 indicates enriched in CAF CM treated cells. Analysis carried out using GSEA_4.2.2 software, FDR<0.05 was considered significant – FDR<0.05 *, <0.01 **, <0.0001 ****, ns = not significant. Details of gene sets in Table 5-1 and Appendix 3.

		CM vs Control		CM vs CM + anti-IL6		
Gene Set Name	Description	NES	FDR	NES	FDR	Reference
Hallmark Estrogen Response Early	Hallmark gene set for early estrogen response	-1.586	0.015	-1.437	0.031	Liberzon <i>et al</i> , 2011 ³²
Hallmark Estrogen Response Late	Hallmark gene set for late estrogen response	-1.259	0.117	-1.073	0.294	Liberzon <i>et al</i> , 2011 ³²
Nardone 2022 FOXA1 ILC 120	ILC-specific FOXA1 target genes	-2.344	0.000	-2.846	0.000	Nardone <i>et al</i> , 2022 ⁵
Selli 2019 Letrozole Resistant DOWN	Genes downregulated in letrozole resistant tumours vs dormant tumours after treatment	-1.859	0.001	-1.789	0.000	Selli <i>et al</i> , 2019 ³³
Selli 2019 Letrozole Resistant UP	Genes upregulated in letrozole resistant tumours vs dormant tumours after treatment	2.318	0.000	2.620	0.000	Selli <i>et al</i> , 2019 ³³
Sikora 2014 ILC E2 DOWN	ILC-specific E2 downregulated genes	1.074	0.300	1.167	0.297	Sikora <i>et al</i> , 2014 ³
Sikora 2014 ILC E2 UP	ILC-specific E2 upregulated genes	-1.459	0.034	0.992	0.479	Sikora <i>et al</i> , 2014 ³
Simoes 2020 STAT3 CSC	STAT3 related genes upregulated in ALDH+ ET treated cells compared to ALDH- cells from metastatic breast cancer patients	1.578	0.014	1.467	0.039	Simões <i>et al</i> , 2020 ¹⁸
Xu 2021 SUM44PE TAMR	Genes upregulated in tamoxifen resistant and treated SUM44PE cells	-1.236	0.116	1.075	0.395	Xu <i>et al</i> , 2021 ³⁴

Table 5-1. Descriptions of gene sets used in Figure 5-3. Descriptions of gene sets and the results of the GSEA comparing gene expression in CAF CM treated cells to control and CAF CM + anti-IL6 treated cells. NES = normalised enrichment score, FDR<0.05 was considered significant. Genes in these gene sets can be found in Appendix 3.

Xu *et al* used previously published datasets to identify genes associated with tamoxifen resistance in ILC cell lines³¹⁶. This analysis identified 184 genes that were commonly differentially expressed in tamoxifen treated or tamoxifen resistant clones of SUM44PE cells relative to untreated or parental SUM44PE cells, respectively. Of these, 74 were also regulated by CAF CM treatment of SUM44PE cells in our RNAseq dataset. This gene set of SUM44PE tamoxifen resistance related genes is not enriched in CAF CM treated SUM44PE cells relative to control or CAF CM + anti-IL-6 treated cells (Figure 5-3 and Table 5-1). However, five of the genes in this gene set (*S100A8*, *AKR1C1*, *CYP1B1*, *CEBPD*, and *MUC1*) were also significantly upregulated by CAF CM treatment of SUM44PE cells (Figure 5-4A and B).

Expression of *CEBPD*, *CYP1B1* and *MUC1* was significantly increased by CAF CM, dependent on IL-6 (Figure 5-4B). Expression of *AKR1C1* and *S100A8* was also significantly increased by CAF CM but this was only partially IL-6 dependent, as expression of these genes was still significantly higher in CAF CM + anti-IL6 samples compared to control cells. All these genes, except for *MUC1*, positively and significantly correlate with *IL6* expression and negatively and significantly correlate with *ESR1* expression in ER+ ILC tumours in TCGA data (Figure 5-4C). *AKR1C1*, *CEBPD*, *CYP1B1* and *MUC1* are also significantly more highly expressed in ER+ ILC than ER+ IDC tumours (Figure 5-4D). This suggests that there is a relationship between these tamoxifen resistance related genes and *IL6/ESR1* in tumours and that they are enriched in ILC tumours.

Overall, this analysis shows that CAF CM treatment of SUM44PE cells suppresses downstream Estrogen Receptor and related signalling and upregulates genes involved with endocrine therapy resistance in tumours and ILC cell lines.

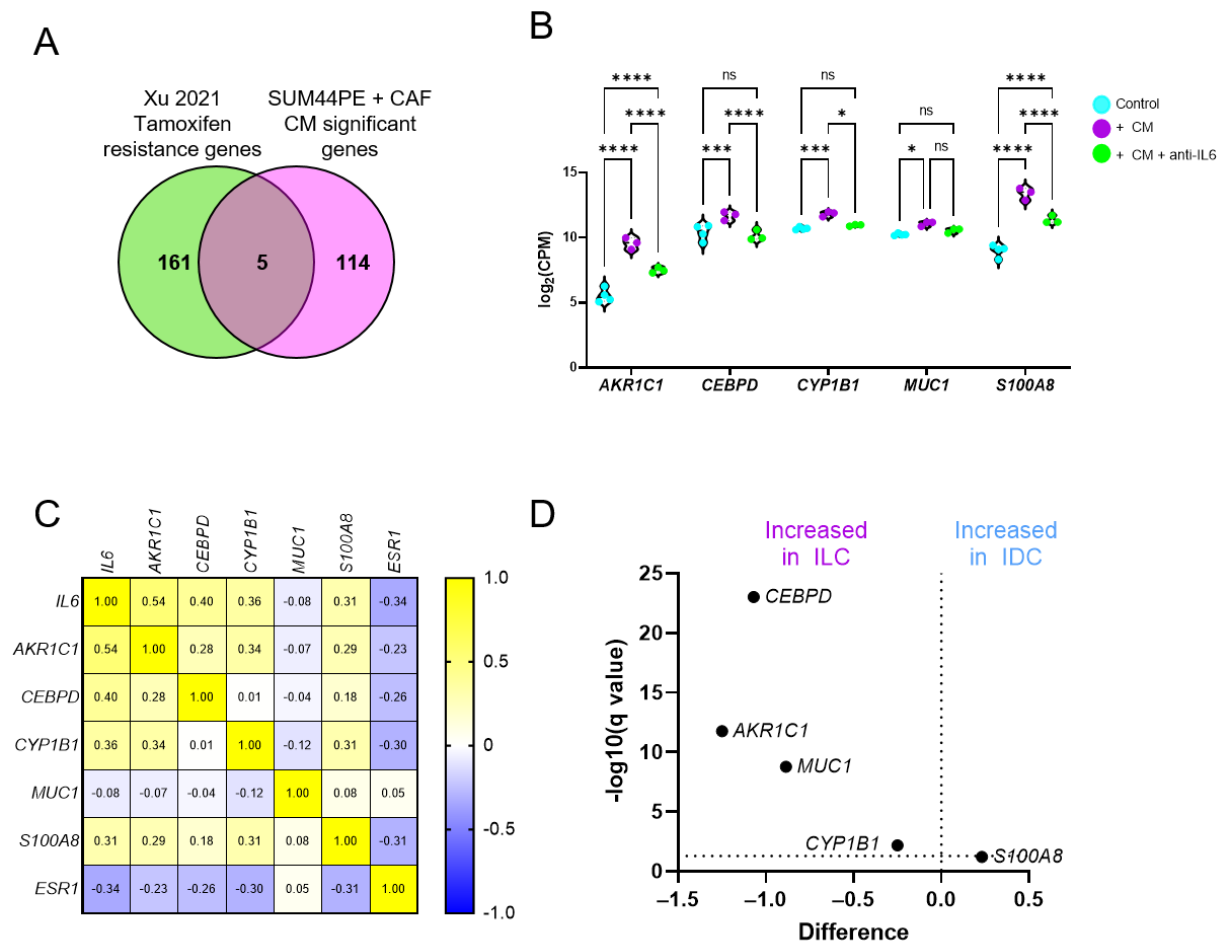


Figure 5-4. CAF CM induces expression of genes implicated in Tamoxifen resistance. (A) 5/119 genes differentially expressed between control and CAF CM treated SUM44PE cells are upregulated in tamoxifen resistant SUM44PE cells from Xu et al 2021³¹⁶ analysis. (B) Expression of the 5 tamoxifen resistance related genes in the SUM44PE RNAseq dataset, expression in control samples in blue, + CAF CM in purple and CAF CM + anti-IL-6 in green. (C) Correlation matrix of the 5 CAF CM induced tamoxifen resistance genes in TCGA ER+ ILC tumours (n=191), Pearson's correlation r values shown, all values significant (p<0.05) except for *CEBPD* v *CYP1B1* and *MUC1* with all genes. (D) Volcano plot showing differential expression of these genes in the TCGA RNAseq dataset between ER+ ILC tumours (n=191) and ER+ IDC tumours (n=563), multiple unpaired t-tests with BKY correction in Graphpad Prism, points above the horizontal dotted line are significantly differentially expressed, FDR<0.05.

5.2.4 IL-6 does not drive ALDH+ cancer stem cell formation in SUM44PE cells

STAT3 has been shown to cause endocrine therapy resistance by promoting breast cancer stem cell (CSC) formation in a number of models^{83,133,163}, and inhibiting STAT3 expression or targeting it with drugs such as SFX-01 reduces the stem-like activity of breast cancer cell lines and patient derived mammospheres⁸³. One marker of breast cancer stem cells is expression of aldehyde dehydrogenase (ALDH) enzymes³⁵¹, with *ALDH1A1* and *ALDH1A3* isoforms driving most ALDH activity in breast cancer³⁵². In the TCGA dataset, *IL6* expression significantly and positively correlates with *ALDH1A3* in ER+ ILC patients (Pearson's $r=0.48$, $p<0.0001$, $n=191$) suggesting a possible link between IL-6/STAT3 signalling and ALDH+ CSC formation in ILC.

GSEA showed a significant enrichment for STAT3-related genes upregulated in ALDH+ cells from metastatic patient samples in CAF CM treated SUM44PE cells (Figure 5-5A)⁸³. To assess whether IL-6/STAT3 signalling is driving CSC formation in the SUM44PE cells, I carried out an ALDEFLUOR assay, in which ALDH+ cells are fluorescently labelled. Stimulation of SUM44PE cells with IL-6 for 6 days did not change the percentage of ALDH+ cells (Figure 5-5B and C). However, IL-6 stimulated cells did have a higher mean intensity of the ALDEFLUOR fluorescence (Figure 5-5D), potentially suggesting that although IL-6 does not increase the number of CSCs, it may increase the expression of ALDH enzymes. However, a western blot showed no increase in ALDH1A3 protein expression (Figure 5-5E and F). This suggests that, at least at in these experimental conditions, IL-6/STAT3 does not promote ALDH+ SUM44PE CSC formation. The influence of IL-6/STAT3 signalling on mammosphere or colony formation has not been assessed but would be interesting to see.

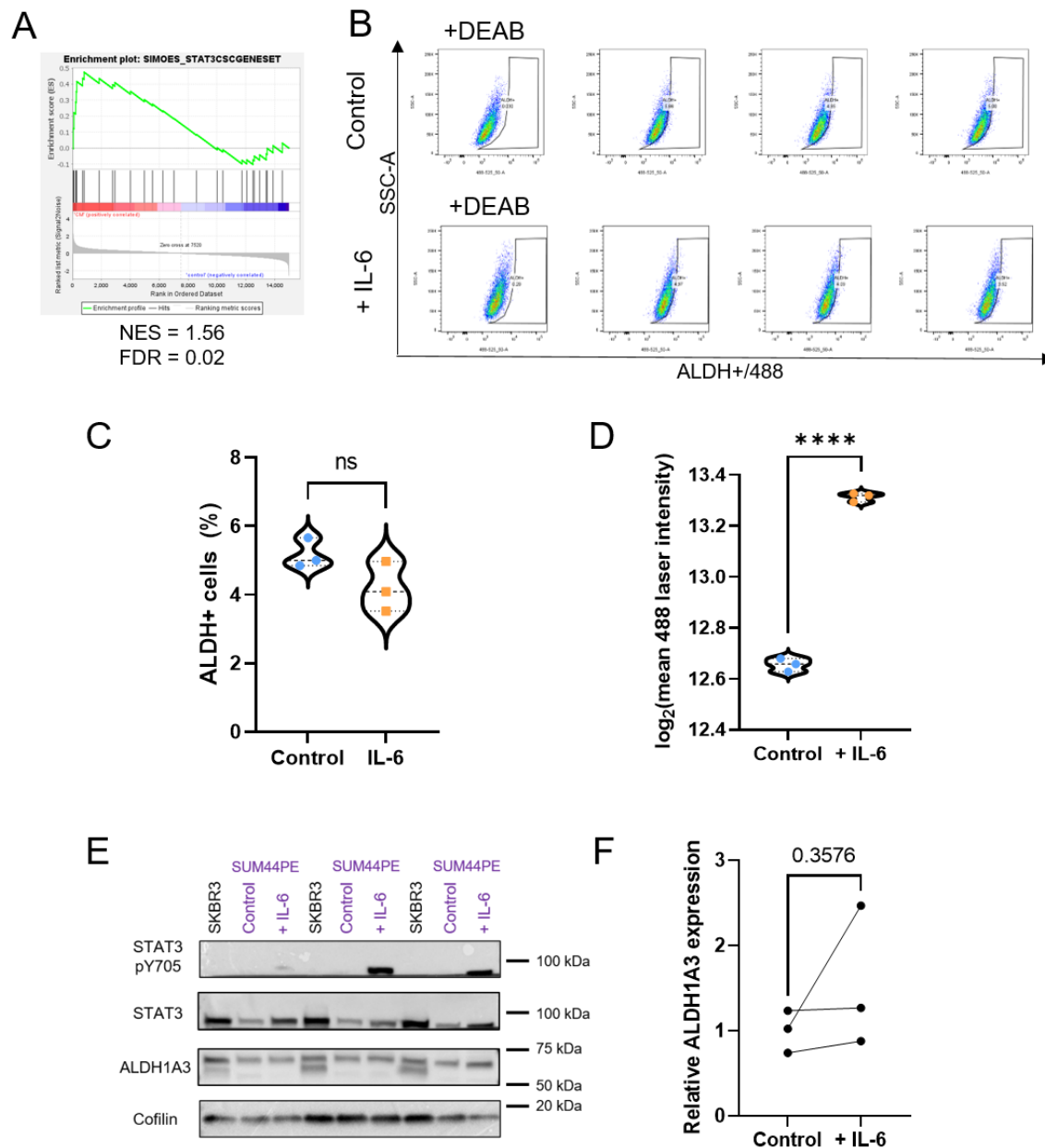


Figure 5-5. IL-6 does not drive the formation of ALDH+ breast cancer stem cells. (A) GSEA plot of STAT3 related genes upregulated in ALDH+ cells⁸³ in CAF CM treated SUM44PE cells (left, red) compared to both control cells (right). (B) Flow cytometry plots of control (top row) SUM44PE cells or cells stimulated with IL-6 (10 ng/ml) for 6 days (bottom) stained with ALDEFLUOR reagent, ALDH+ cells determined relative to match samples treated with the ALDH inhibitor DEAB. (C) Quantification of percent of ALDH positive cells. (D) Quantification of ALDEFLUOR reagent fluorescence intensity. (E) Western blot of the HER2+ IDC cell line SKBR3 as a positive control for ALDH expression and SUM44PE cells +/- IL-6 for 6 day, stained for STAT3 pY705, total STAT3, ALDH1A3 and cofilin as loading control, (F) quantification of ALDH1A3 band in (E). Unpaired (B and D) and paired (F) two-tailed t-test in GraphPad Prism, **** p<0.0001, n=3.

Chapter 5.3 Discussion

This data suggests that CAF derived IL-6 stimulation modulates estrogen receptor expression and signalling in ILC cell lines. There is a well-established link between these two signalling pathways although this has not been extensively explored in ILC cell lines and tumours. Further work is needed to determine the mechanism of this relationship.

5.3.1 IL-6/STAT3 and ER α signalling in breast cancer

Multiple studies have shown the inter-linking of IL-6/STAT3 and ER α but little work has interrogated the relationship between these two pathways in ILC. Interestingly, IL-6 stimulation of SUM44PE cells lead to a downregulation of Hallmark estrogen early and late signalling which contrasts with the enrichment of these gene sets in IL-6 stimulated T47D cells, although this analysis was after 1 hour rather than 24 hours of stimulation¹⁶⁵. This supports a potentially distinct mechanism and interaction of the IL-6/STAT3 pathway with ER α signalling in ILC compared to IDC.

Many ER α targets common to both IDC and ILC cell lines, such as *MYC* and *GREB1*²⁵², were not significantly altered by IL-6 in CAF CM. In addition, expression of the ILC-specific E2 responsive gene *SNAI1* on CAF CM stimulation was increased^{252,350}. This potentially implies that IL-6/STAT3 regulates a subset of ER α target genes rather than completely inhibiting its activity in SUM44PE cells. The IL6GS also contains some genes known to be upregulated by ER α (*CYP1B1*³⁵³, *THRSP*⁵⁹, *S100A9*, *LTF* and *ACOX2*²⁰⁹) and STAT3 and ER α have been shown to bind to many of the same promoters and enhancers in breast cancer cell lines¹⁶⁵. This suggests that although IL-6 stimulation mostly suppresses ER α and downstream signalling in SUM44PE cells, IL-6 may also support expression of some ER α target genes.

One of the many ways estrogen signalling differs in ILC compared to IDC is that ILC cell lines have a more open chromatin state dependent on FOXA1 activity that allows ER α to remain bound to these sites upon tamoxifen treatment⁵⁹. The fact that ILC-specific FOXA1 target genes were so significantly suppressed by CAF CM treatment of SUM44PE cells suggests that IL-6 regulates both generic E2 responsive genes as well as genes that are unique to ILC ER α -FOXA1 signalling. This is also supported by CAF CM driven downregulation of ILC-specific E2 induced genes²⁵², though this may

be also dependent on other CAF CM secreted factors besides IL-6 (Figure 5-3). Siersbæk *et al* demonstrated that IL-6 induces the formation of a pSTAT3-ER α -FOXA1 complex in T47D cells that regulates a distinct set of genes from the ER α -FOXA1 complex alone¹⁶⁵. Further work is needed to interrogate the effect of IL-6 signalling on the FOXA1-dependent chromatin state of ILC and whether the same pSTAT3-ER-FOXA1 complex forms and regulates the same genes in the presence of E2 and IL-6 in ILC cell lines as was seen in IDC cell lines.

It would be interesting to untangle the mechanism by which IL-6/STAT3 signalling leads to downregulation of ER α and downstream signalling. Direct binding of STAT3 to the *ESR1* promoter and recruitment of transcriptional repressors could be driving this repression or it could be via more indirect methods, such as co-opting ER α co-factors or occupying estrogen-response elements. Co-culture of the ZR-75-1 ER+ IDC cell line with metastatic bone marrow stromal cells lead to an IL-6 dependent proteasomal degradation of ER α , and an anti-IL6 antibody restored the fulvestrant sensitivity of this cell line¹³¹. A similar mechanism may be occurring in SUM44PE cells and should be investigated. However, the weaker correlation between *IL6* and ER α compared to *ESR1* expression in ILC tumours (Figure 5-2) suggests that a transcriptional mechanism of regulation may be more likely. Another possibility is that IL-6 stimulation promotes the outgrowth of an ER-/ER low sub-population of ILC tumour cells and suppresses growth or promotes apoptosis of ER+ cells rather than directly suppressing expression of ER α . As all these experiments were carried out in estrogen-free conditions, except for the immunofluorescence experiments in Figure 5-1 in which cells were cultured in media with phenol-red and FBS, it would be useful to follow up these experiments with the addition of E2 to see how this influences IL-6 induced gene expression and ER α expression.

IL-6 has been shown to promote ET resistance in MCF7 and T47D IDC cell lines¹³¹ via upregulation of the PI3K and ERK pathways rather than downregulation of ER α expression. In future, it may be worth doing comparative studies investigating the impact of IL-6/STAT3 signalling on ET resistance dependent on E-cadherin expression, either comparing ILC and IDC cell lines or using cell lines with either inducible expression or knockdown of E-cadherin. However, the FOXA1-dependent chromatin landscape that contributes to ILC tamoxifen resistance was shown to be

independent of E-cadherin expression⁵⁹, suggesting other mechanisms may be driving the divergence between ILC and IDC response to E2 and ETs.

5.3.2 CAF CM induced endocrine therapy resistance genes

Analysis of the RNAseq data of CAF CM stimulated SUM44PE cells also identified an enrichment for genes implicated in tamoxifen and letrozole resistance (Figure 5-3 and Figure 5-4). The mechanism of these drugs differs as tamoxifen is a selective ER modulator, which competes with E2 for ER α binding, whereas letrozole is an aromatase inhibitor and inhibits the conversion of androgens into E2 in post-menopausal women³⁵⁴. IL-6/STAT3 signalling may mediate resistance to each drug via different mechanisms. Interestingly, in long-term estrogen deprived (LTED) clones of SUM44PE and MDA-MB-134VI cells, the STAT3 pathway and genes in the IL6GS (*IL4R*, *BCL3* and *FAM83A*) were upregulated relative to parental cells in both cell lines, even though E2 independence was mediated via distinct mechanisms in LTED SUM44PEs and MDA-MB-134VI cells⁶¹. Upregulation of the STAT3 pathway has been found in many ET resistant breast cancer cell lines and patient samples^{61,83,355}, suggesting this is an important and central pathway for supporting ET resistance in breast cancer in general.

5.3.2.1 IL-6 driven upregulation of letrozole resistance related genes

Response to E2 and ETs has been shown to differ *in vitro* between ILC and IDC cell lines^{59,252,347}, with many of these studies validated in patient datasets. However, one study found that letrozole responsive ILC and IDC patient tumours respond very similarly to letrozole treatment, with a very large overlap in differentially expressed genes relative to matched pre-treatment samples⁶⁸. However, this work exclusively analysed letrozole responsive tumours, and so the mechanisms of resistance may differ between the subtypes. CAF CM treated SUM44PE cells were significantly enriched for genes upregulated in letrozole resistant ER+ tumours compared to responsive tumours⁶⁹ dependent on IL-6 within CAF CM (NES = 2.32 and 2.62 compared to control and CAF CM + anti-IL-6 respectively, FDR = 0.00, Figure 5-3). This implies that IL-6 in CAF CM may promote letrozole resistance, though this needs to be further validated *in vitro*. Whilst this dataset did include 8 letrozole responsive and 4 letrozole resistant ILC tumours, it would be useful to determine if there are any differences in gene expression between letrozole resistant ILC and IDC tumours in

larger datasets or to identify ILC-specific or enriched letrozole resistant gene expression changes.

5.3.2.2 CAF CM driven SUM44PE Tamoxifen resistance genes

CAF CM treatment lead to significant upregulation of 5 genes shown to be upregulated in Tamoxifen resistant SUM44PE cells (Figure 5-4)³¹⁶. All of these genes (*AKR1C1*, *CEBPD*, *CYP1B1*, *MUC1* and *S100A8*) have been identified as IL-6/STAT3 target genes previously^{83,158,356}, demonstrating that these are not ILC-specific IL-6 induced genes. However, addition of anti-IL6 to CAF CM did not reduce *AKR1C1* and *S100A8* expression back to control cell levels (Figure 5-4B), suggesting that other components of the CAF conditioned media besides IL-6 are also contributing to their expression and their potential abrogation of tamoxifen effectiveness. It would be useful to determine the other CAF CM factors that promote the expression of these genes.

CYP1B1 and *AKR1C1* are enzymes involved in the metabolism of hormones and so could potentially be promoting tamoxifen resistance by affecting the metabolism of E2 and tamoxifen. Although overexpression of the *AKR1C1*, 2 and 3 enzymes was demonstrated to increase tamoxifen resistance in SUM44PE cells and knockdown of these genes increased response to tamoxifen, the mechanism was not further interrogated by the authors³¹⁶. *AKR1C1* overexpression has been shown to drive metastasis of non-small cell lung carcinoma by enhancing STAT3 phosphorylation by direct binding and promoting the STAT3-JAK2 interaction independent of its aldoketo-reductase activity²⁸⁴. *AKR1C1* may be supporting tamoxifen resistance via its enzymatic activity or it could potentially be dependent on its ability to promote sustained STAT3 phosphorylation. The fact that *AKR1C2* and 3 can also promote tamoxifen resistance³¹⁶, though to a lesser extent than *AKR1C1*, but do not promote STAT3 phosphorylation²⁸⁴ suggests that tamoxifen resistance could be a combination of both enzymatic and non-enzymatic activity of *AKR1C1*.

CYP1B1 is a P450 cytochrome enzyme and is involved in the metabolism of active estradiol (E2) to 4-hydroxylestradiol (4OHE2) which increases free-radical activity and promotes DNA-adduct formation³⁵⁷. Its expression is elevated in many cancers and it has been implicated in therapy resistance. For example, in a prostate cancer model, increased *CYP1B1* and 4OHE2 promoted increased ER α activity and subsequent induction of IL-6 expression, driving CSC formation and therapy resistance³⁵³.

CEBPD encodes for the transcription factor and known IL-6/STAT3 target¹⁵⁸ CCAAT enhancer binding protein delta (CEBPδ) that has been shown to drive both pro- and anti-tumorigenic phenotypes in breast cancer^{300,358,359}. Interestingly, *in vivo* data has shown that knockdown of *CEBPD* leads to suppression of the stem-like abilities of multiple breast cancer cell lines, as well as decreasing metastases and circulating tumour cells³⁰⁶. Importantly, *CEBPD* only promoted EMT markers and CSC formation in ER+ cell lines in hypoxic conditions or in the presence of exogenous IL-6, underlining the context dependency of this signalling. The upregulation of *CEBPD* in the analysis by Xu *et al* implies that it may play a role in tamoxifen resistance, potentially via supporting CSC formation. ERα promotes stabilisation of CEBPδ at the protein level by inhibiting its proteasomal degradation³⁰⁰ and so it will be important to assess whether the IL-6 induced increase in *CEBPD* mRNA is carried over to the protein level as IL-6 stimulation is also driving ERα downregulation in our data.

MUC1 and *S100A8* have both previously been identified as STAT3 induced genes in breast cancer^{83,158}. *MUC1* is highly expressed in ALDH+ ET resistant CSCs and high expression in combination with another STAT3 target gene *OSMR* significantly predicts decreased recurrence free survival in ER+ patients. However, neither of these genes have been directly implicated in driving ET resistance or CSC formation. Although AKR1C1 and CYP1B1 have been shown to directly influence ET response, the upregulation of these five genes in tamoxifen resistant SUM44PE cells may only serve as a marker of increased STAT3 activity rather than actively influencing resistance.

The ILC-specific FOXA1 gene set, which was significantly down regulated by CAF CM in SUM44PE cells, was shown to promote resistance to tamoxifen in ILC cells *in vitro* and be associated with significantly poorer outcome in Luminal A ILC patients⁵⁹. This contradicts with the significant increase in genes associated with Tamoxifen resistance in SUM44PE cells and highlights the need for further interrogation and experimental validation of CAF CM driven ET resistance in ILC.

5.3.2.3 CDK4/6 inhibitor resistance

Cyclin dependent kinase 4/6 inhibitors (CDK4/6i) have recently been approved for treatment of metastatic breast cancer, which has shown to increase ILC patient progression free survival^{70,360,361}. However, some patients are intrinsically resistant

and acquired resistance frequently occurs. CDK4/6i resistant cell lines have been shown to have reduced ER α expression^{175,315} and increased pSTAT3 via induction of autocrine IL-6 signalling¹⁷⁵, with inhibition of STAT3 increasing cell response to CDK4/6 inhibitors. Interestingly, increased interferon signalling is also associated with CDK4/6 resistance and immunosuppression in ER+ breast cancer³¹⁵ and GSEA identified the Hallmark interferon α and γ gene sets as significantly upregulated by IL-6 in CAF CM in SUM44PE cells (Figure 4-8), suggesting paracrine signalling from ILC CAFs could potentially also predispose cells to CDK4/6i resistance.

Taken together, this data demonstrates that CAF CM induces expression of genes associated with multiple mechanisms of ET resistance and CDK4/6i resistance in ILC tumour cells, consistent with the previously reported importance of IL-6/STAT3 signalling in breast cancer therapy resistance. However, this is some of the first work to explicitly investigate this pathway in ILC. *In vitro* experiments and analysis of patient data is necessary to validate the phenotypes implied by the gene expression changes.

5.3.3 Breast cancer stem cells

The IL-6/STAT3 pathway has been shown to promote CSC formation in many breast cancer models, leading to increased endocrine therapy resistance^{83,163,167}. There was a significant enrichment for STAT3-related genes identified in ALDH+ metastatic patient cells compared to ALDH- cells in CAF CM treated SUM44PE cells⁸³. However, this gene set is not an unbiased ALDH+ CSC gene set and therefore this enrichment may simply reflect increased STAT3 activity in CAF CM treated cells rather than increased ALDH+ CSC related gene expression. IL-6 treatment did not increase the percentage of ALDH+ SUM44PE cells (Figure 5-5A), suggesting that even though IL-6/STAT3 signalling is altering ER α signalling and ET resistance genes, it is not promoting ALDH+ CSC formation.

However, IL-6 treated SUM44PE cells did have a higher mean laser intensity than untreated cells and this higher laser intensity was also seen in the DEAB negative control (Figure 5-5B). This may be due to the assay conditions, in which the ALDEFLUOR substrate is added to samples and then a small proportion of each sample is removed and the ALDEFLUOR inhibitor DEAB is added as a negative reference sample. As the ALDEFLUOR reagent reacts immediately, it may be that

there are a number of ALDH⁺ cells in the negative control and this could obfuscate our results. This experiment should be repeated with the inhibitor added before the ALDEFLUOR reagent to ensure a proper negative control.

A number of genes in the IL6GS have been shown to be upregulated in CSCs or to promote CSC formation^{83,306,353,362,363}. However, as IL-6/STAT3 signalling has been so consistently shown to promote CSC formation, these genes may just be downstream of this pathway rather than directly responsible for the phenotype. A number of other breast CSC markers have been described, including CD44⁺/CD24⁻¹⁶⁷ and CD133^{hi}¹⁶³ cells, which STAT3 has been associated with. Of note is that a single pilot experiment with SUM44PE cells stimulated with IL-6 for 6 days found no change in the numbers of CD44⁺/CD24⁻ cells, with virtually all cells expressing CD24 (data not shown) and no significant change in CD133 expression was seen in the RNAseq dataset. However, the RNAseq analysis was carried out after 24 hours of CAF CM stimulation and so further analysis of CSC markers should be carried out after long term incubation and, as these are rare populations, through single cell analysis by flow cytometry or scRNAseq.

Experiments to investigate the stem-like abilities of SUM44PE cells and whether they are influenced by paracrine signalling from CAFs should also be carried out, such as mammosphere formation efficiency and colony formation. Tamoxifen and fulvestrant treatment of patient derived cells from IDC and ILC patients has been shown to increase pSTAT3 expression and numbers of ALDH⁺ cells^{83,364} via cell autonomous methods – it would be interesting to see whether paracrine IL-6 signalling predisposes cells to ET resistance or whether the combination of ET and IL-6 treatment promotes CSC formation.

Chapter 5.4 Conclusions

These data show that IL-6/STAT3 signalling negatively regulates expression of the Estrogen Receptor at the mRNA and protein level and induces expression of genes associated with resistance to multiple endocrine targeted therapies.

Currently this work is speculative, although analysis of the literature and gene expression changes strongly supports a paracrine IL-6/STAT3 driven ET resistance phenotype in these cells. Further work is needed to fully elucidate the role of IL-6/STAT3 in regulating ER α in ILC and to determine whether IL-6 stimulation does in fact promote resistance to endocrine targeted therapies *in vitro* and in patient data as is suggested by this analysis. It would be useful to expand these experiments into PDX models and patient derived cell lines as well as patient data to see if our findings in the SUM44PE cell line extend into other ILC models.

Chapter 6 Functional consequences of IL-6/STAT3 signalling in Invasive Lobular Carcinoma

The siRNA targeting STAT3 used in this chapter was designed and formulated by the Poli group at the University of Turin. The concept of the Zebrafish embryo metastasis assay was suggested by Prof Valeria Poli and initial exploratory experiments (not included here) were carried out at the Molecular Biotechnology Centre at the University of Turin, Italy with microinjections done by Lorenzo Prever, funded by the Royal Society of Edinburgh Saltire Early Career Fellowship. Dr Adelaide Young carried out Zebrafish embryo maintenance/husbandry and did the microinjections for the zebrafish embryo xenograft assay carried out in Edinburgh, as well as any handling and experimentation of embryos after 4-days post fertilisation. Aleksandra Rozyczko carried out the whole-mount zebrafish sample preparation and imaging, and assisted with tissue culture.

Chapter 6.1 Introduction

The Invasive Lobular Carcinoma (ILC) enriched IL-6/STAT3 pathway has been shown to promote a wide range of pro-tumorigenic phenotypes in breast cancer, although its role in ILC has not previously been assessed. Work in Chapter 4 identified an IL-6 dependent enrichment for genes involved in inflammation, migration and angiogenesis, and a downregulation in Estrogen signalling pathways and proliferation (Figure 4-9 and 4-10).

IL-6/STAT3 signalling has been shown to drive epithelial-to-mesenchymal transition (EMT) and metastasis in a number of breast cancer models^{167,365}, with combined high expression of STAT3 and IL-6¹⁶⁵ and expression of STAT3 gene signatures⁸³ correlating with poor recurrence free survival in ER+ breast cancer patients. A number of the genes in the IL-6 gene signature defined in Chapter 4 have also been shown to drive EMT and migration in breast cancer, such as *BCL3*^{366,367} and *MUCL1*^{323,326}. However, the majority of these studies have used IDC cell lines, such as MCF7 or T47D, and no studies so far have looked specifically at the impact of the IL-6/STAT3 pathway in ILC cell lines, which deviate from IDC cell lines in growth⁴⁷, gene expression^{48,251,253}, mutations⁷⁹, invasion³⁶⁸, response to estrogen and endocrine targeted therapies^{59,252,350}, metabolism^{39,61} and behaviour *in vivo*^{35,176}.

As IL-6/STAT3 signalling appears to be driving epithelial-to-mesenchymal (EMT) and migratory gene signatures in SUM44PE cells, we wanted to assess the effect IL-6 signalling has on cell migration and metastasis. One of the issues with studying invasive lobular carcinoma is the lack of models and the speed at which they grow⁷⁹. *In vitro*, there are two validated ER+ human ILC cell models, SUM44PE and MDA-MB-134VI, although these cells are very slow growing. There are also mouse cell lines available, namely the KEP (*K14Cre;Cdh1^{ff};Tp53^{ff}*) and WEPTn (*Wcre;Cdh1^{ff};Pten^{ff}*) models developed by the Jonkers and Derksen groups^{37,67}. The gold-standard for *in vivo* models of ILC is the mouse intraductal xenograft model, which closely recapitulates the progression and features of ILC³⁵. However, these models take up to a year to reach end point, are expensive and relatively low throughput.

An alternative *in vivo* model for metastasis is the Zebrafish (*Danio rerio*) embryo xenograft model (ZFEX)⁹², which provides a useful ‘in between’ for *in vitro* and mouse *in vivo* experiments. The ZFEX assay involves the micro-injection of fluorescently

labelled human (or mouse) cancer cells into the perivitelline space (PVS) of 2-day post fertilisation (2dpf) zebrafish embryos to form a 'primary tumour' and then assessing the spread of tumour cells from the primary site throughout the embryo. This assay only takes two days, as embryos are fixed at 4dpf/2-days post injection (dpi), and is relatively high throughput compared to mouse models, allowing high numbers of embryos to be assessed in each experiment. Using Tg(Fli:eGFP) zebrafish embryos (transparent embryos with GFP-labelled endothelial cells) allows easy imaging and assessment of angiogenesis as well as tumour cell dissemination. It is also possible to carry out whole-mount immunofluorescence, taking z-stack confocal microscopy images of the entire embryo, rather than sections or dissected out organs, allowing imaging of the entire embryo and detection of single cells disseminated throughout the embryo.

The work in this chapter demonstrates that IL-6/STAT3 signalling drives a mesenchymal-like phenotype and increased SUM44PE cell motility *in vitro* and that IL-6 stimulation leads to increased cell dissemination in ZFEX model of metastasis, suggesting IL-6/STAT3 signalling may be a useful therapeutic target in ILC.

Chapter 6.2 Results

6.2.1 CAF conditioned media drives ILC cell proliferation

The RPPA and RNAseq experiment in Chapter 4 determined that CAF conditioned media induces activation of a number of signalling pathways and the biological consequences of this needed to be determined. Paracrine signalling from CAFs is well known to support tumour cell growth and proliferation both *in vitro* and *in vivo*^{122,128} and IL-6/STAT3 signalling can also promote proliferation in breast cancer^{124,128,165}.

SUM44PE cells are very slow growing and have low metabolic activity³⁰⁷ so metabolic proliferation assays, such as the MTT assay, are not appropriate for these cells. A 5'-ethynyl-2'-deoxyuridine (EdU) based proliferation assay was established, in which the thymidine analogue EdU is added to the media of cells, and is incorporated into the DNA of cells that pass through S-phase. EdU+ cells are then fluorescently labelled with TAMRA-azide and the percentage of labelled cells can be assessed by microscopy or flow cytometry. The higher the percentage of EdU+ cells, the faster the rate of proliferation and the minimum length of time for 100% of cells to be labelled is the length of the cell cycle.

The RPPA data in Figure 4-1 identified a significant induction of proteins involved in IGF-1 signalling and downstream activation of the PI3K pathway (see Chapter 4). The IGF-1 and PI3K pathways are often hyper-activated in ILC due to mutations in PI3K pathway genes³ and via both autocrine^{52,222} and paracrine¹³ mechanisms. The impact of these signalling pathways on proliferation of SUM44PE cells was assessed. SUM44PE media contains insulin to support cell survival and growth (see Table 2-1). An EdU proliferation assay was carried out in SUM44PE cells, with EdU incorporation measured at 24, 48 or 72 hours. Treatment with the dual insulin receptor (InsR)/IGF-1 receptor (IGF1R) inhibitor OSI-906 significantly reduced proliferation at each time point (Figure 6-1A and B), suggesting that insulin in the media is promoting SUM44PE proliferation.

Stimulation with IGF-1 or CAF CM from ED38 and ED2334, two independent ILC CAF cell lines, increased proliferation of the SUM44PE cells relative to cells grown in growth factor depleted media (Figure 6-1C). The addition of OSI-906 significantly reduced the proliferation in all conditions (Figure 6-1D and E), decreasing the percentage of EdU+ cells to below the levels of unstimulated cells. This suggests that inhibiting IGF1R/InsR

may block autocrine signalling as well as paracrine or exogenous sources of IGF-1/insulin.

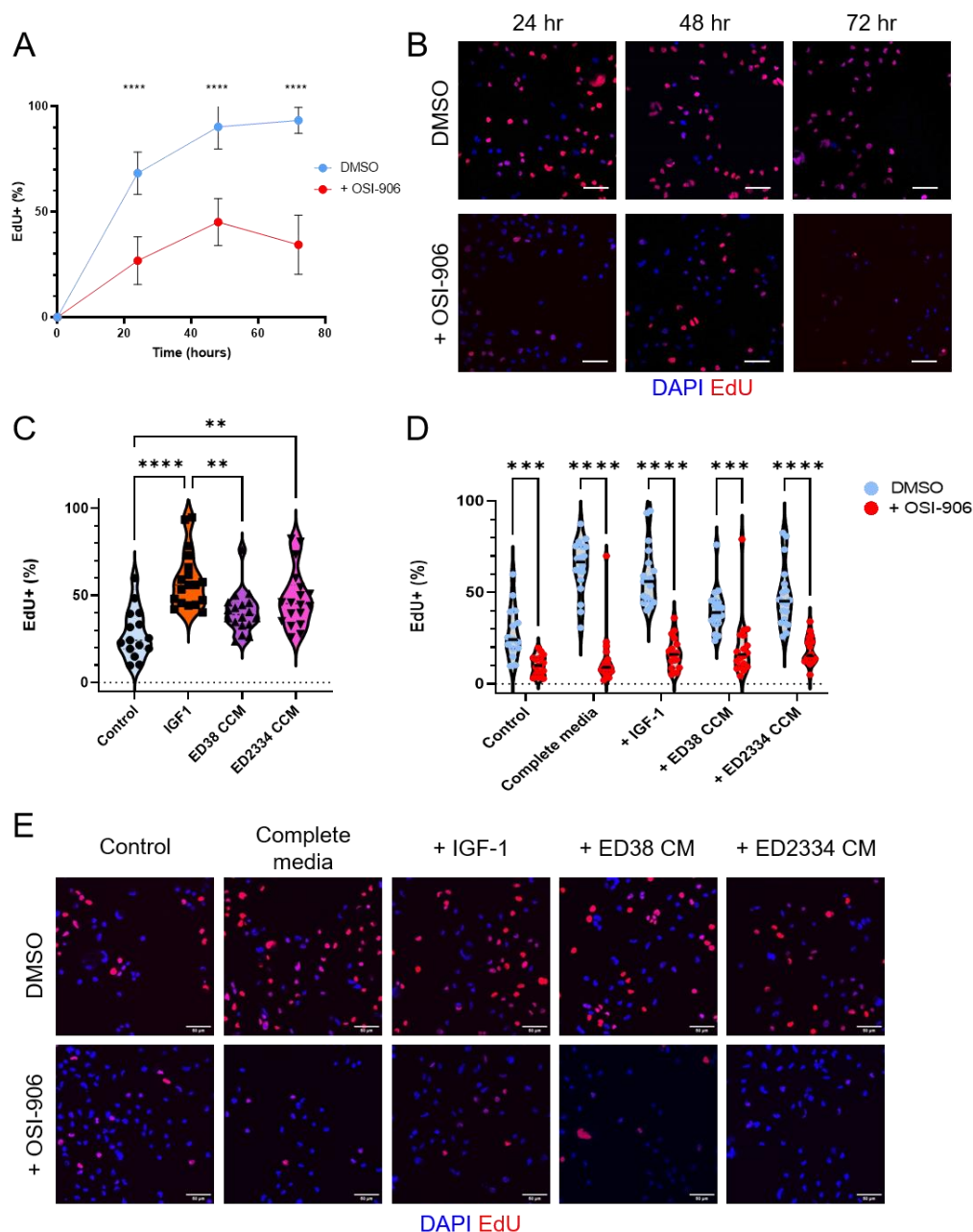


Figure 6-1. An IGF1R/InsR inhibitor reduces SUM44PE cell proliferation. (A) Graph showing mean and standard deviation (sd) of percentage of EdU+ cells at each time point for cells in complete media with DMSO (blue) and in cells in complete media treated with OSI-906 (50 mM) (red), 2-way ANOVA in GraphPad Prism, **** $p < 0.0001$. (B) Representative images of each time point. (C) Graphs showing percentage of EdU+ SUM44PE cells in growth factor depleted media (control) or stimulated for 24 hours with IGF-1 and CAF CM, and (D) with the addition of OSI-906. (C) and (D) Each point represents the percentage of EdU+ cells in a single field of view (FoV). (C) One-way ANOVA with Tukey's multiple comparison test, ** adjusted $p < 0.01$, **** adjusted $p < 0.0001$ and (D) multiple unpaired t-tests with FDR correction in GraphPad Prism, FDR < 0.001 ***, FDR < 0.0001 ****. (E) Representative images of each condition, nuclei labelled with DAPI (blue) and EdU+ cells labelled with TAMRA-Azide (red). In (A), (C) and (D) 3 repeats per condition, 6 random FoVs per repeat, images at 40X on Olympus FV3000, scale bar 50 μm .

The RNAseq data in Chapter 4 showed an IL-6 dependent downregulation of genes related to the cell cycle in SUM44PE cells after 24 hours of CAF CM stimulation. To determine if these gene expression changes resulted in a suppression of proliferation, SUM44PE cells in GF-depleted media were stimulated with recombinant human IL-6 and EdU for 24 hours. This led to a decrease in the number of EdU+ SUM44PE cells, although this reduction in proliferation was not significant (Figure 6-2). This suggests that IL-6 may reduce SUM44PE proliferation as suggested by the IL-6 driven reduction in proliferative gene signatures seen in Chapter 4. Further experiments at later time points or a time course experiment may be needed to fully uncover the impact of IL-6 on SUM44PE cell proliferation.

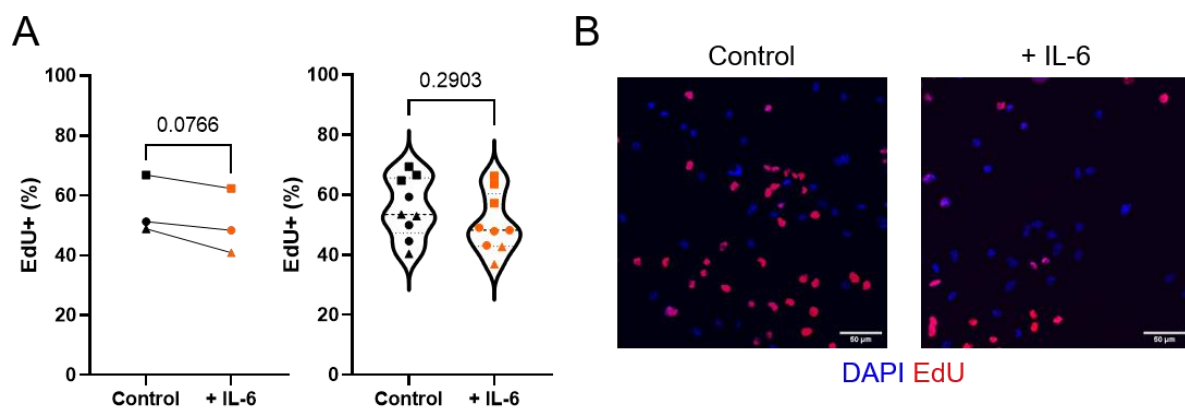


Figure 6-2. IL-6 treated cells have decreased rates of proliferation. (A) Graphs showing percentage of EdU+ SUM44PE cells in growth factor depleted media (control) or +IL-6 for 24 hours. Left – showing mean % EdU+ cells across all FoVs in three independent experiments and right – data from same three independent experiments, showing mean % of EdU+ cells in three technical replicates in each independent repeat, different shapes indicating different experiments. Paired (left) and unpaired (right) t-tests in Graphpad prism, numbers indicating p-value. (B) Representative images of each condition with nuclei labelled with DAPI (blue) and EdU+ nuclei labelled with TAMRA-azide (red), scale bar 50 µm, images taken at 40X on Olympus FV3000 microscope.

6.2.2 IL-6 drives epithelial-to-mesenchymal transition gene signatures and morphological changes in SUM44PE cells

The RNAseq data in Chapter 4 identified a significant enrichment for EMT gene signatures in SUM44PE cells stimulated with CAF CM, dependent on IL-6 (see Figures 4-9 and 4-10). EMT gene sets from the hallmarks dataset²⁰⁹ and from Emad 2020^{369,370} were both significantly enriched in CAF CM stimulated cells compared to both control and CAF CM + anti-IL-6 (Figure 6-3A and B), demonstrating that IL-6 within the CAF CM is driving expression of EMT-related genes. Although the expression of classical EMT transcription factor genes was not significantly altered by addition of CAF CM, the RNAseq data showed an IL-6 dependent increase in *SNAI1*, *SNAI2*, *VIM*, *ZEB1* and *ZEB2*, and a decrease in *TWIST1* expression with the addition of CAF CM (Figure 6-3C).

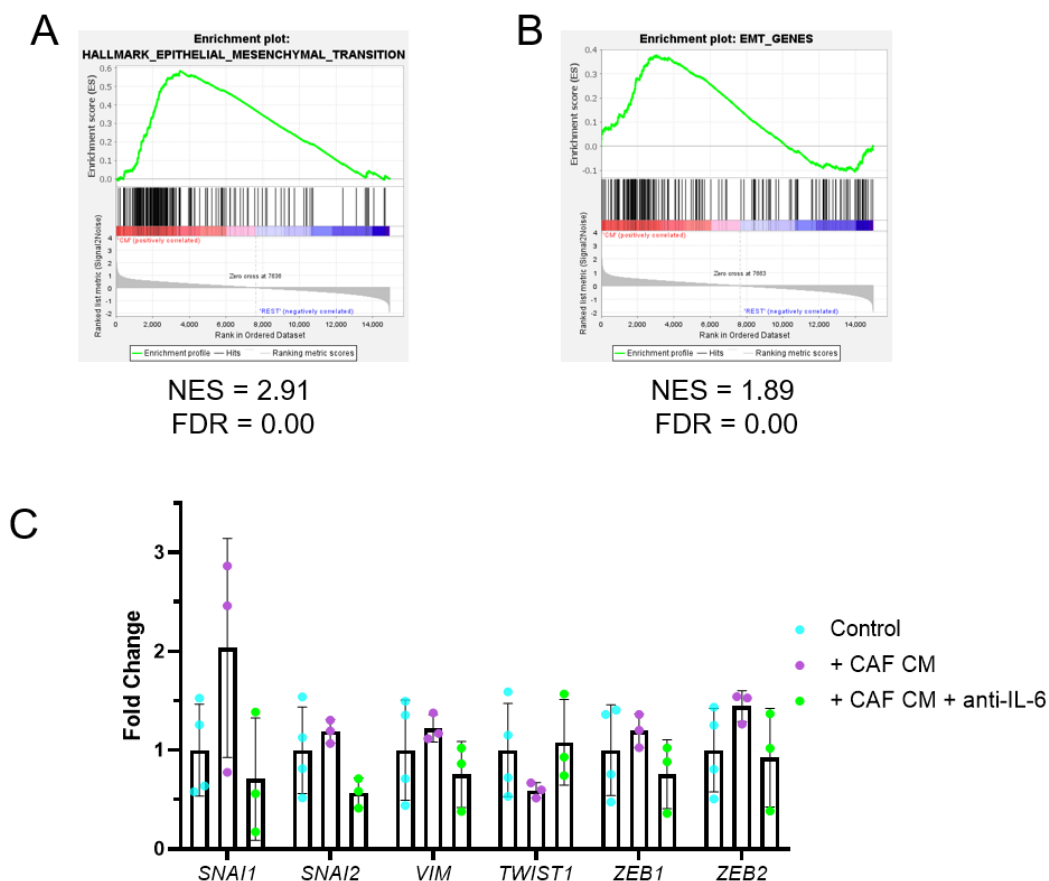


Figure 6-3. IL-6 in CAF CM drives EMT genes in SUM44PE cells. Gene set enrichment analysis plot of CAF CM treated SUM44PE cells (left, red) compared to both control and CAF CM + anti-IL6 cells (right, blue) for (A) Hallmarks EMT gene set and (B) Emad 2020^{369,370}. (C) Fold change in expression of EMT transcription factors in SUM44PE cells, no significant differences with glmFIT ANOVA-like test in EdgeR. All data from SUM44PE RNAseq experiment, described in detail in Chapter 4.

Cells that have gone through EMT display a more elongated mesenchymal-like morphology, compared to the more rounded epithelial cells. However, ILC cell lines in basal conditions display a very rounded epithelial-like morphology, with little cytoplasm, rather than a mesenchymal-like elongated phenotype, despite loss of E-cadherin expression, which is a key feature of EMT. However, when SUM44PE cells were seeded on ILC CAF derived extracellular matrices (CDMs, Figure 6-4A)²⁰⁰, an IL-6 dependent elongation and morphological change could be seen (Figure 6-4B). Untreated SUM44PE cells or cells pre-treated with IL-6 for 1 week were seeded onto ED2334 ILC CDMs. After 24 hours (0 hour time point, Figure 6-4), the media was replaced with complete media ('control' and 'pre-treated') or complete media + IL-6 ('+ IL-6' and 'pre-treated + IL-6'). Images were taken in using the Incucyte S3 and the changes in morphology were measured in ImageJ using aspect ratio (AR) as a readout, with more elongated cells having a larger AR.

Regardless of IL-6 stimulation, the mean AR significantly increased after 48 hours in all conditions (Figure 6-4C), suggesting adherence to CAF-secreted extracellular matrix proteins alone can promote a more mesenchymal-like phenotype. However, the addition of IL-6 to both the control and pre-treated cells increased the mean AR by far more than the increase in AR for unstimulated cells. The AR of pre-treated cells without additional IL-6 increased by 0.3 at 48 hours compared to an increase of 0.05 for unstimulated control cells. This suggests an IL-6 driven EMT-like transcriptional programme may be promoting this mesenchymal-like phenotype even 48 hours after IL-6 withdrawal.

Although at 0 hours there was no significant difference in AR between the conditions (Figure 6-4D), the mean AR for pre-treated cells (1.6 and 1.65) was larger than for untreated cells (1.44 and 1.42) demonstrating that IL-6 pre-treatment also contributes to a more mesenchymal-like morphology. At 48 hours, untreated control cells had a significantly lower AR (1.5) compared to all other conditions (+ IL-6 AR=1.9, pre-treated AR=1.89, pre-treated + IL-6 AR=2.4). There was no significant difference in AR for untreated cells + IL-6 compared to cells only pre-treated with IL-6, again supporting the maintenance of IL-6-activated transcriptional programmes after IL-6 withdrawal. However, there was a significant increase in AR between pre-treated cells

and pre-treated cells + IL-6, suggesting that continuous stimulation with IL-6 further promotes this mesenchymal-like morphological change.

The morphology of the SUM44PE cells on the CDMs is heterogeneous, with a mixture of rounded epithelial-like and elongated mesenchymal-like cells present in each condition, as can be seen in Figure 6-4B. A perfect circle has an AR of 1 and the more elongated a cell, the larger its AR, with mesenchymal-like cells previously having been defined as having an AR greater than 1.7¹⁷⁶. The percentage of mesenchymal-like cells (AR>1.7) was determined. Again, this showed a significant increase in cells with a mesenchymal-like morphology after 48 hours on CDMs in all conditions (Figure 6-4E), with the largest increase seen in cells treated with IL-6 short term, increasing from 13.8% to 35.9% of cells with an AR>1.7. At 0 hours, there were still some mesenchymal-like cells in the control cells (16%), although this was far lower than the pre-treated cells (23.5%), and there was a significant difference between pre-treated + IL-6 cells (29.5%) and both control and + IL-6 cells (Figure 6-4F). Control cells had a significantly lower percentage of cells with an AR>1.7 compared to all other conditions at 48 hours.

This demonstrates that IL-6 promotes an elongated mesenchymal-like phenotype in SUM44PE cells and suggests that this occurs via transcriptional mechanisms, as the pre-treated cells continued to elongate after IL-6 was removed. An RNA-seq experiment has since been carried out with SUM44PE cells stimulated with IL-6 for a week which will help understand which IL-6-dependent genes may be driving this mesenchymal-like phenotype.

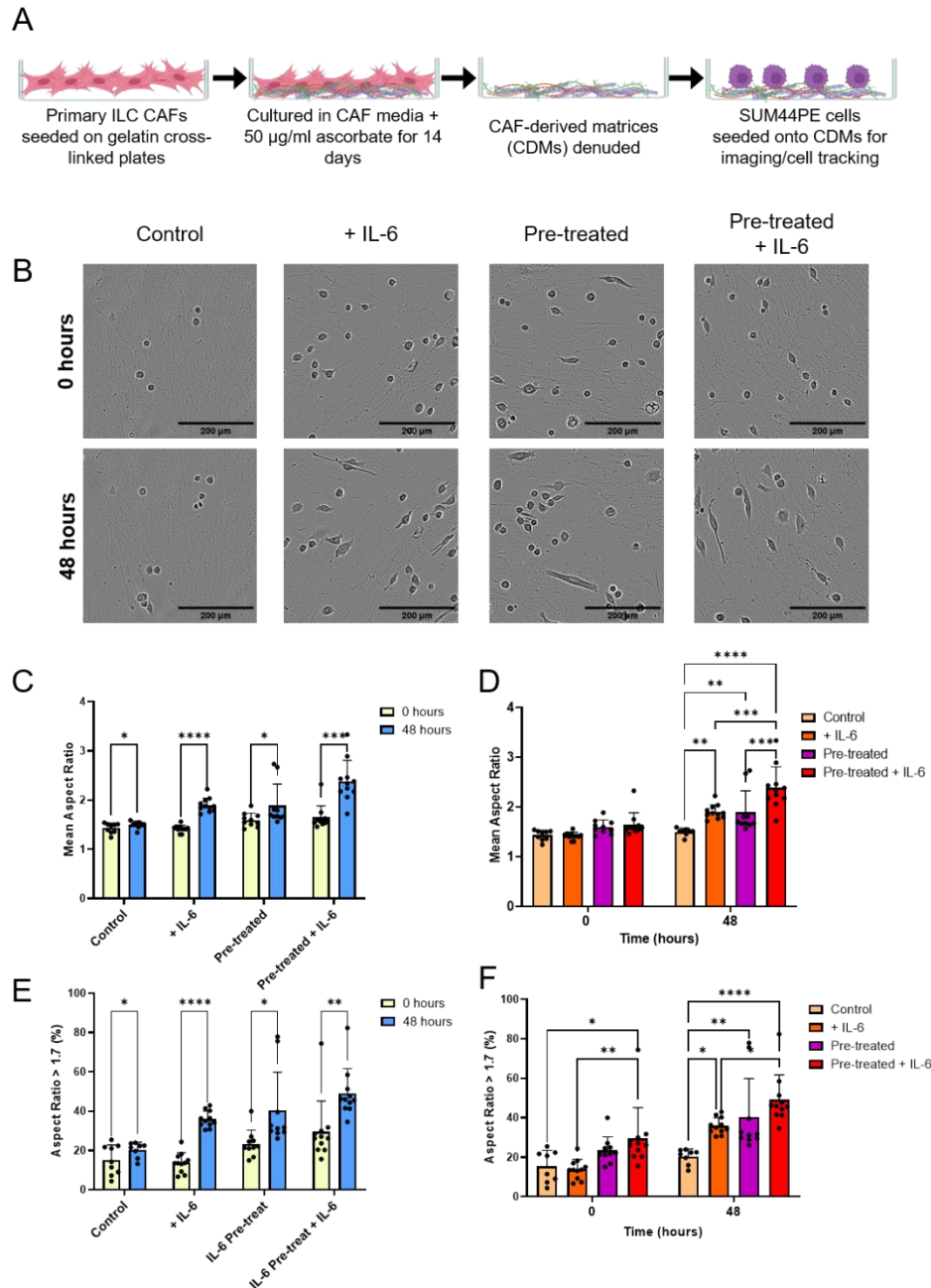


Figure 6-4. IL-6 drives a morphological change in SUM44PE cells on CAF-derived matrices. (A) Schematic showing how the CDMs are generated from primary ILC CAFs, made in Biorender. SUM44PE cells cultured in complete media or in complete media + IL-6 (10 ng/ml) for 1 week (Pre-treated) were seeded onto ED2334 ILC CAF derived matrices (CDMs). After 24 hours media was changed and replaced with complete media (control and Pre-treated) or complete media + IL-6 (+ IL-6 and Pre-treated + IL-6, 10 ng/ml). (B) Representative images of SUM44PE cells on CDMs at the time of stimulation (0 hours, top row of images) and 48 hours after stimulation (48 hours, bottom row). Showing cropped regions of images taken at 10x on Incucyte S3, scale bar showing 200 µm. Graphs show mean aspect ratio in each FoV (C and D) or percentage of cells with AR>1.7 (E and F), with points representing each FoV. (C and E) Comparing changes at 0 and 48 hours within each condition, and (D and F) comparing changes between each condition at 0 hours and 48 hours. Statistics in Graphpad Prism, (C and E) multiple paired t-test with BKY correction showing q-value and (D and F) two-way ANOVA with Sidak's correction, showing adjusted p-value. q-value/adjusted p<0.05 *, p<0.01 **, p<0.001 ***, p<0.0001 ****. N=3 for each condition, with 2-4 FoVs per repeat, an average of 89 cells per FoV, minimum 27 cells.

To determine if these morphological changes were dependent on STAT3 activity downstream of IL-6, SUM44PE cells were treated with siRNA against luciferase (siCTRL) or STAT3 (siSTAT3) encapsulated in C12-200 liposomes¹⁹⁹ for 48 hours (Figure 6-5A). Cells were then seeded onto ED26 ILC CAF CDMs, allowed to adhere for 24 hours then stimulated with or without IL-6 (Figure 6-5B).

IL-6 stimulation significantly increased the mean AR of control and siCTRL treated SUM44PE cells, as seen in Figure 6-4, but no significant changes were seen for siSTAT3 treated cells (Figure 6-5C). At 0 hours, there were no differences in mean AR between any of the groups (Figure 6-5D). However, at 72 hours, IL-6 treated control and siCTRL cells had an increased mean AR relative to untreated cells (although the difference in siCTRL cells was not significant, adjusted $p = 0.062$), whereas no difference was seen between IL-6 treated and untreated siSTAT3 cells.

A significant increase in the percentage of mesenchymal-like cells ($AR > 1.7$) could be seen for IL-6 stimulated control, siCTRL and siSTAT3 cells between 0 and 72 hours (Figure 6-5E). However, the increase for IL-6 treated siSTAT3 cells (increased by 11%) was far lower than for control and siCTRL cells (19% and 17%, respectively). At 72 hours, the percentage of cells with an $AR > 1.7$ was higher for IL-6 treated control and siCTRL cells than untreated cells but no difference was seen for siSTAT3 cells (Figure 6-5F). This shows that IL-6 driven morphological changes are dependent on STAT3 expression.

As in the experiment in Figure 6-4, adherence to CDMs caused an increase in the percentage of mesenchymal-like cells independent of IL-6 stimulation in control (8%, adjusted $p = 0.09$) and siCTRL (13%, adjusted $p = 0.0009$) cells (Figure 6-5E). However, no significant increase was seen for untreated siSTAT3 cells (4%, adjusted $p = 0.7$). This suggests that the CDM driven morphological changes may potentially act through STAT3 too. It would be useful to carry out immunofluorescence to determine if the mesenchymal-like cells have nuclear localised STAT3 and whether the mesenchymal-like cells in the siSTAT3 treated group retained or restored STAT3 expression.

Together, these data demonstrate that IL-6 drives an EMT-like transcriptional programme (Figure 6-3) and promotes elongation of SUM44PE cells (Figure 6-4), causing a more mesenchymal-like morphology, which is dependent on STAT3 activity downstream of IL-6 stimulation (Figure 6-5).

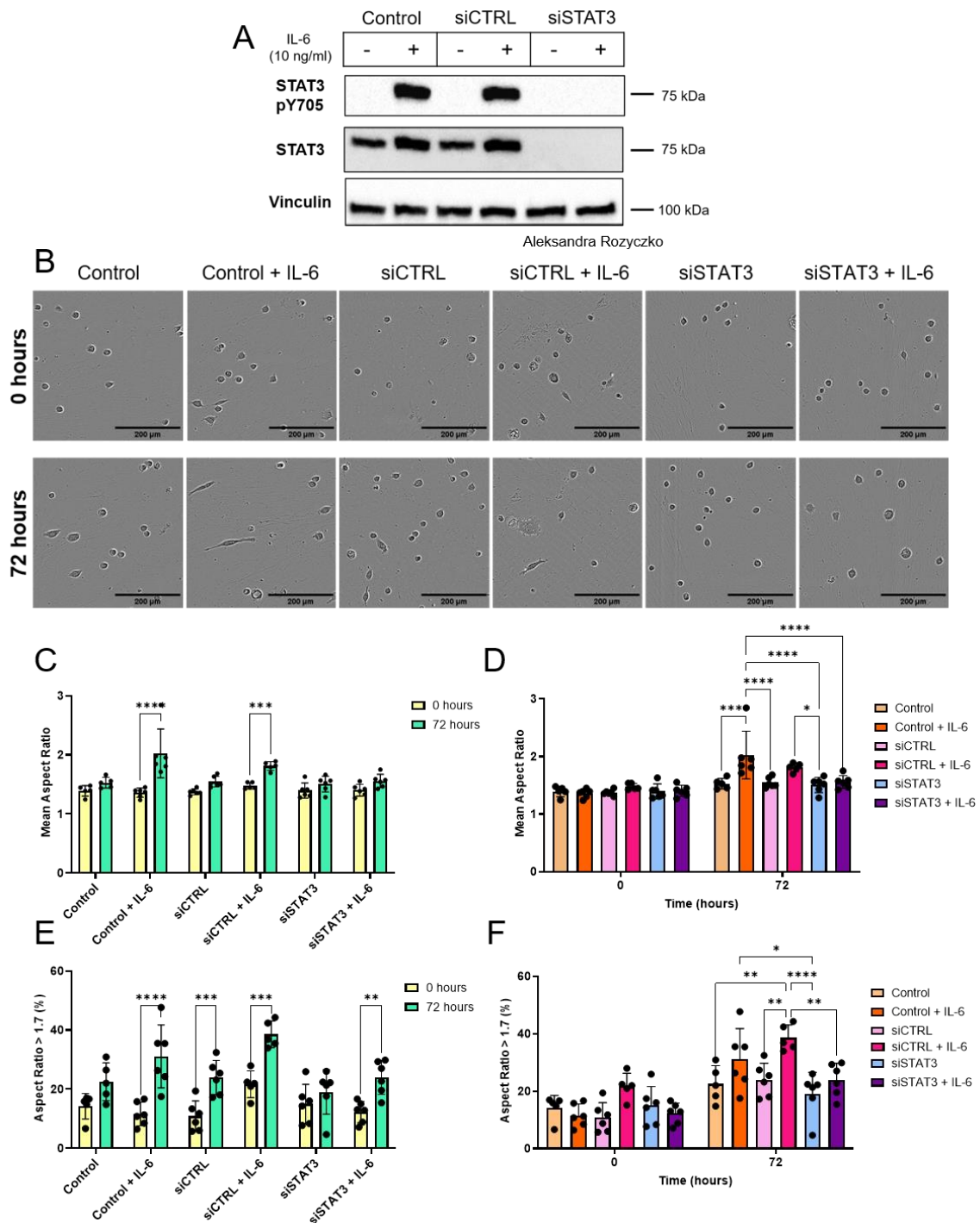


Figure 6-5. IL-6 driven changes in morphology are dependent on STAT3 expression. SUM44PE cells were cultured in complete media (control) or complete media with siCTRL and siSTAT3 lipoparticles (1 μ g/ml) for two days. For the western blot in (A), cells were then stimulated +/- IL-6 (10 ng/ml) for 24 hours and expression of STAT3 pY705, total STAT3 and vinculin as loading control was assessed, carried out by Aleksandra Rozyczko. (B) After siRNA treatment, cells were seeded onto CDMs and after 24 hours (0 hour time point), the media was replaced with complete media +/- IL-6 as indicated. Representative images of SUM44PE cells seeded on ED26 ILC CAF derived CDMs at 0 hours and 72 hours. Cropped images taken at 10X on Incucyte S3, scale bar showing 200 μ m. Graphs showing mean AR (C and D) and percentage of cells with AR>1.7 (E and F) with each point representing one FoV. Statistics in Graphpad Prism, two-way ANOVA with Sidak's correction, adjusted $p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***, $p < 0.0001$ ****. N=3, with 2 FoVs per repeat, an average of 60 cells per FoV, minimum 14 cells.

6.2.3 IL-6 drives random 2D migration of SUM44PE cells

IL-6/STAT3 signalling has been linked to increased cell migration and GSEA identified an IL-6 dependent enrichment of a number of migration related gene sets in SUM44PE cells (Figure 4-10). To assess whether IL-6 stimulation promotes migration of an ILC cell line, SUM44PE cells were seeded onto collagen coated plates and the migration of individual cells was tracked over 24 hours, with the speed of each cell calculated using the Trackmate²⁰¹ plugin in ImageJ.

Stimulation of cells with IL-6 significantly increased the mean speed of migration relative to control cells (Figure 6-6A-C). Although not significant, stimulation with ILC CAF CM increased the speed of migration compared to control cells (overall mean speed 0.39 $\mu\text{m}/\text{min}$ and 0.35 $\mu\text{m}/\text{min}$ respectively), and the addition of an IL-6 neutralising antibody reduced the speed (0.32 $\mu\text{m}/\text{min}$). IL-6 also significantly increased SUM44PE migration over 72 hours (Figure 6-6D and E), confirming that IL-6 promotes the migration of SUM44PE cells. Looking at the speeds of individual cells (Figure 6-6B and D), IL-6, and CAF CM to a lesser extent, increase the speed of a proportion of the total cells, with the majority having similar speeds to control cells. This shows the heterogeneous effect of IL-6 stimulation, similar to heterogeneous effects of IL-6 stimulation on the morphology of SUM44PE cells seen in Figures 6-4 and 6-5.

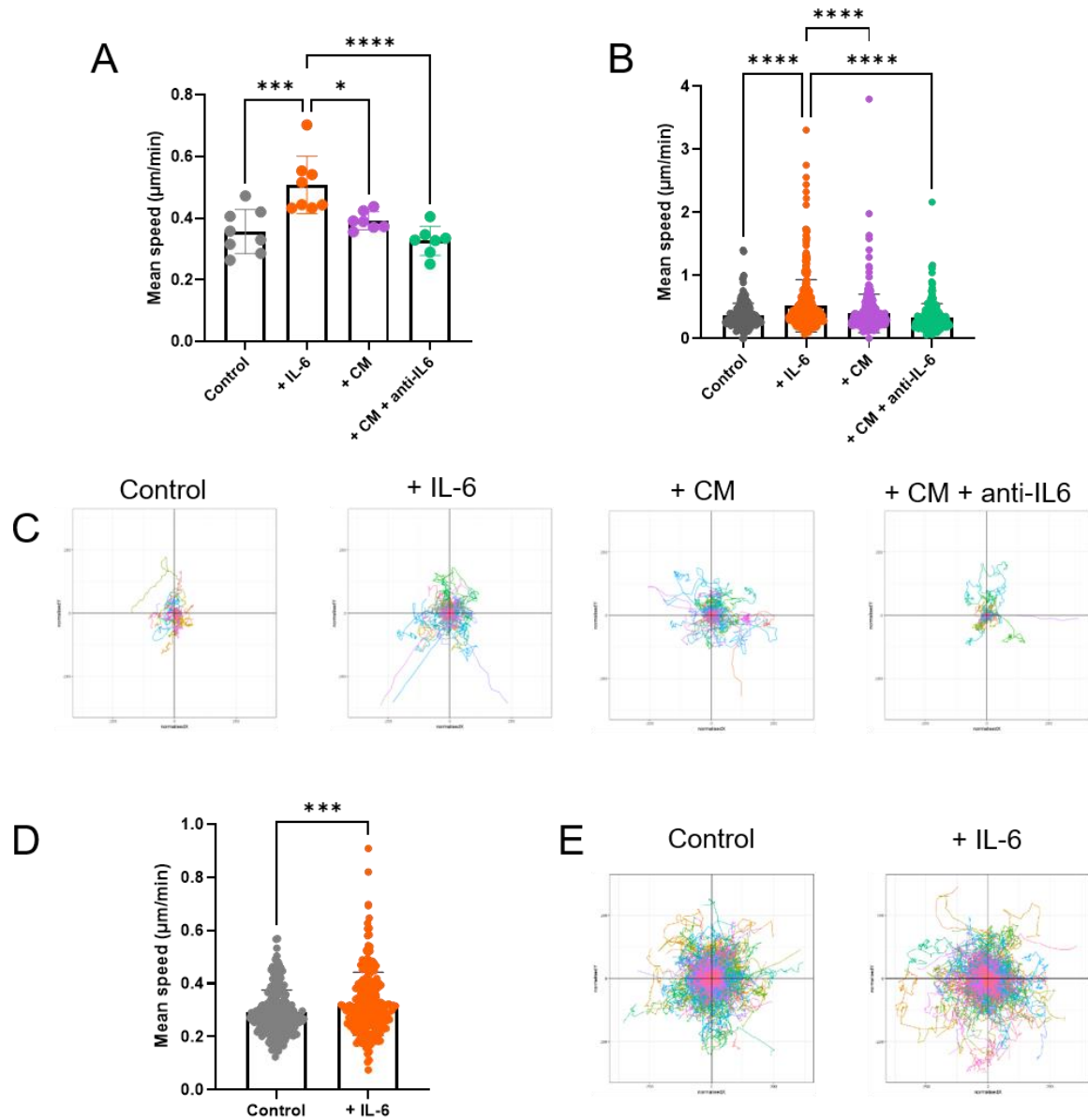


Figure 6-6. IL-6 increases SUM44PE 2D migration. (A-C) SUM44PE cells were seeded into collagen coated imagelock plates and after 24 hours, media was changed to growth factor depleted media (control) or growth factor depleted media + IL-6 (10 ng/ml), ED2334 CAF CM or CAF CM + anti-IL-6 (1 µg/ml). Images were taken every 15 min for 24 hours, n=3 repeats per condition with 3 FoVs per repeat. (A) Mean speed of each FoV and (B) mean speed with each point showing an individual cell, one-way ANOVA with Tukey's multiple comparison test – adjusted $p < 0.05$ *, $p < 0.001$ ***, $p < 0.0001$ ****. (C) Representative co-ordinate plots of cell tracks in each condition. (D) Speed of individual SUM44PE cells over 72 hours in complete media +/- IL-6 (10 ng/ml), images taken every 30 min, n=3 with one FoV per repeat, unpaired t-test $p < 0.0001$. (E) Representative co-ordinate plots of cell tracks. Analysed using Trackmate²⁰¹ plugin in ImageJ.

6.2.4 IL-6 drives SUM44PE cell dissemination in a Zebrafish embryo assay

To determine if IL-6 drives SUM44PE cell migration and dissemination *in vivo*, a zebrafish embryo xenograft (ZFEX) assay was used, following the protocol developed by the Fior lab⁸⁵. Ideally, *in vivo* mouse models would be used but as I had limited time and ILC mouse xenografts take up to a year to reach end point³⁵, not accounting for time to analyse tissues, the rapid and relatively high throughput zebrafish assay was used. SUM44PE cells were pre-treated for 1 week with recombinant human IL-6 in complete media. Pre-treated or control cells were then stained using the lipophilic Dil-dye, labelling the cells fluorescently red. Cells were then microinjected into the perivitelline space (PVS) of two-day post-fertilisation (2dpf) Tg(Fli:eGFP) zebrafish embryos, which are transparent with GFP-labelled vasculature, allowing easy visualisation of injected cells and angiogenesis (see anatomy and schematic, Figure 6-7).

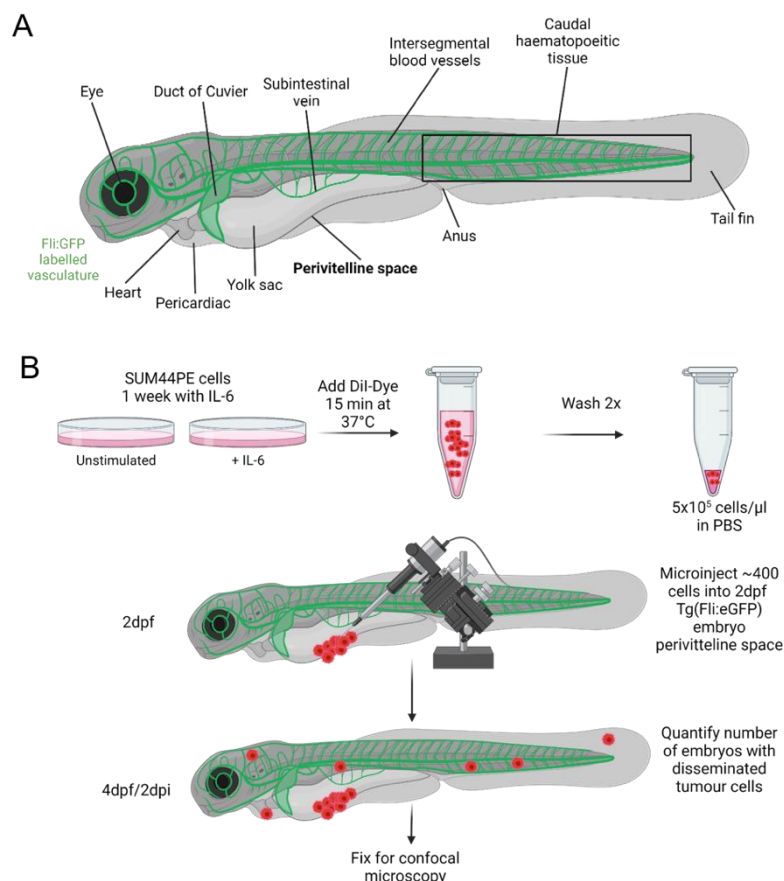


Figure 6-7. Zebrafish embryo xenograft model. (A) Diagram of the anatomy of a two days post fertilisation (2dpf) Tg(Fli:eGFP) zebrafish embryo. (B) Schematic of the assay, with Dil-stained SUM44PE cells injected into the perivitelline space of 2dpf embryos, with metastasis quantified 2 days post injection (2dpi)/4dpf. Made in Biorender.

Roughly 4 hours post injection, embryos were screened to select for live GFP+ embryos with SUM44PE cells only within the primary site (PVS) (Figure 6-8A). Embryos with deformed yolk sacs or cardiac oedema were excluded (Figure 6-8B), as were those with cells injected into the centre of the yolk sac (Figure 6-8C) instead of the PVS. Embryos with SUM44PE cells already outside the yolk sac were also excluded from the main experiment and referred to as 'CIRC' embryos, as the cells had been injected directly into the circulation (Figure 6-8D). These embryos were kept for analysis of later stages of metastasis, although due to time constraints, limited analysis has been carried out on these embryos so far. Embryos not injected with SUM44PE cells were kept as controls to monitor normal zebrafish development (Figure 6-8E).

After 24 hours (1dpi/3dpf), the embryos were then sorted into groups based on primary tumour size (Figure 6-9), as it is difficult to precisely control for the number of cells injected into the PVS. Embryos with tumour cell masses in the PVS smaller than the size of the eye were designated as (+) group, the same size as the eye as (++) and larger than the eye as (+++), although the vast majority of the embryos in the (+++) group do not survive, due to cardiac oedema or the yolk being split in half by the tumour (as can be seen in Figure 6-9, bottom row). At this time point, disseminated SUM44PE cells can already be detected outside of the primary site in some embryos in both the ++ and +++ groups. All screening and sorting of embryos was carried out on the same mesoscope on live anaesthetised embryos at 3.2x with a fixed exposure time of 150 ms for the Dil channel.

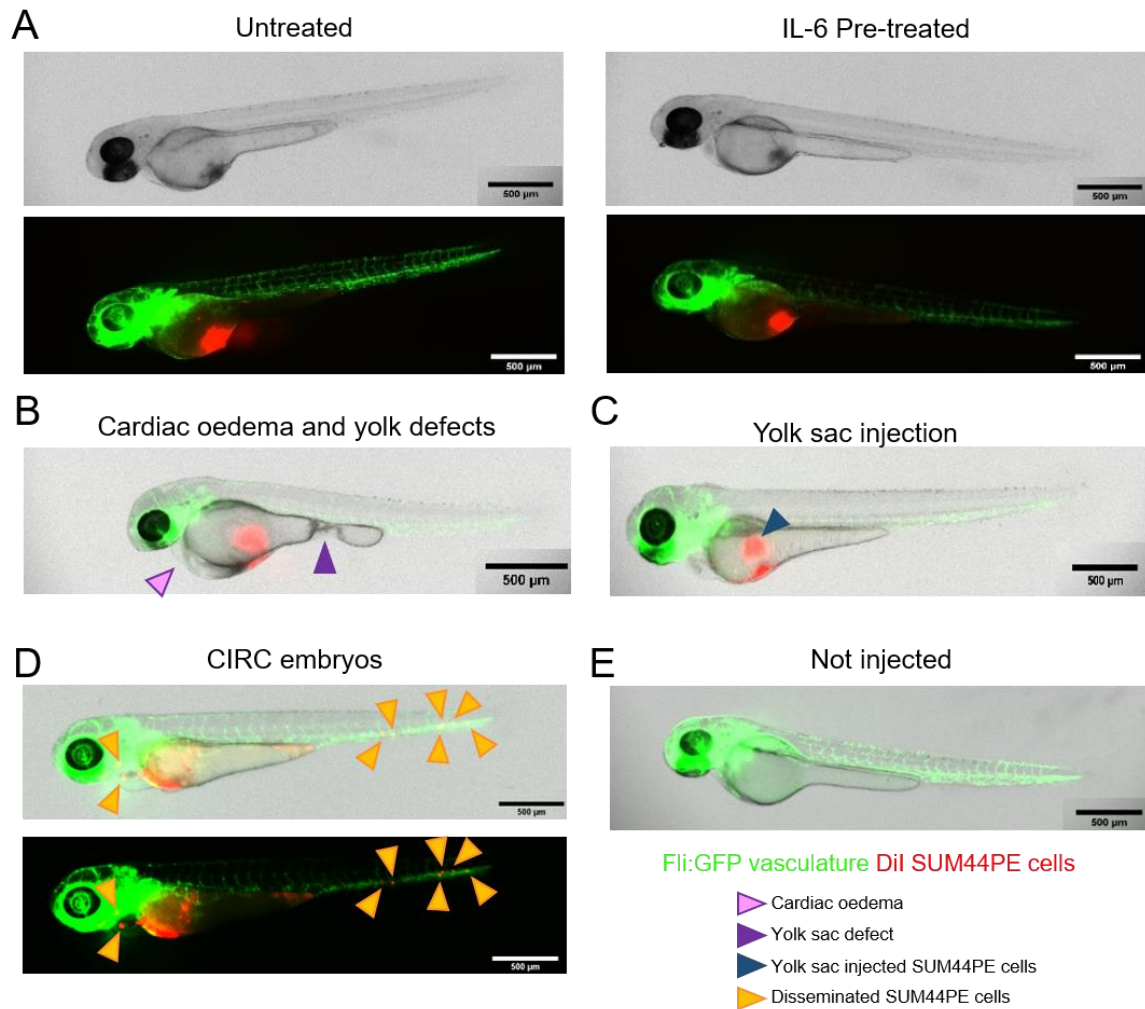


Figure 6-8. Screening embryos immediately after injection of SUM44PE cells. Representative images of 2dpf Tg(Fli:eGFP) embryos taken roughly 4 hours after injection. (A) Embryos with untreated (left) or IL-6 pre-treated (right) SUM44PE cells injected into the PVS of 2dpf embryos. Representative images of embryos excluded due to (B) cardiac oedema (purple outlined arrows) and yolk sac defects (purple filled arrow) or due to (C) SUM44PE cell injection into the yolk sac (blue arrow) instead of PVS. (D) Embryos with SUM44PE cells injected directly into the circulation (yellow arrows) were classed as 'CIRC' embryos and excluded from analysis. (E) A control embryo, not injected with SUM44PE cells. Images taken at 3.2x magnification on a mesoscope, showing brightfield images, GFP labelled vasculature in green and Dil-dyed SUM44PE cells in red. Scale bar, 500 µm.

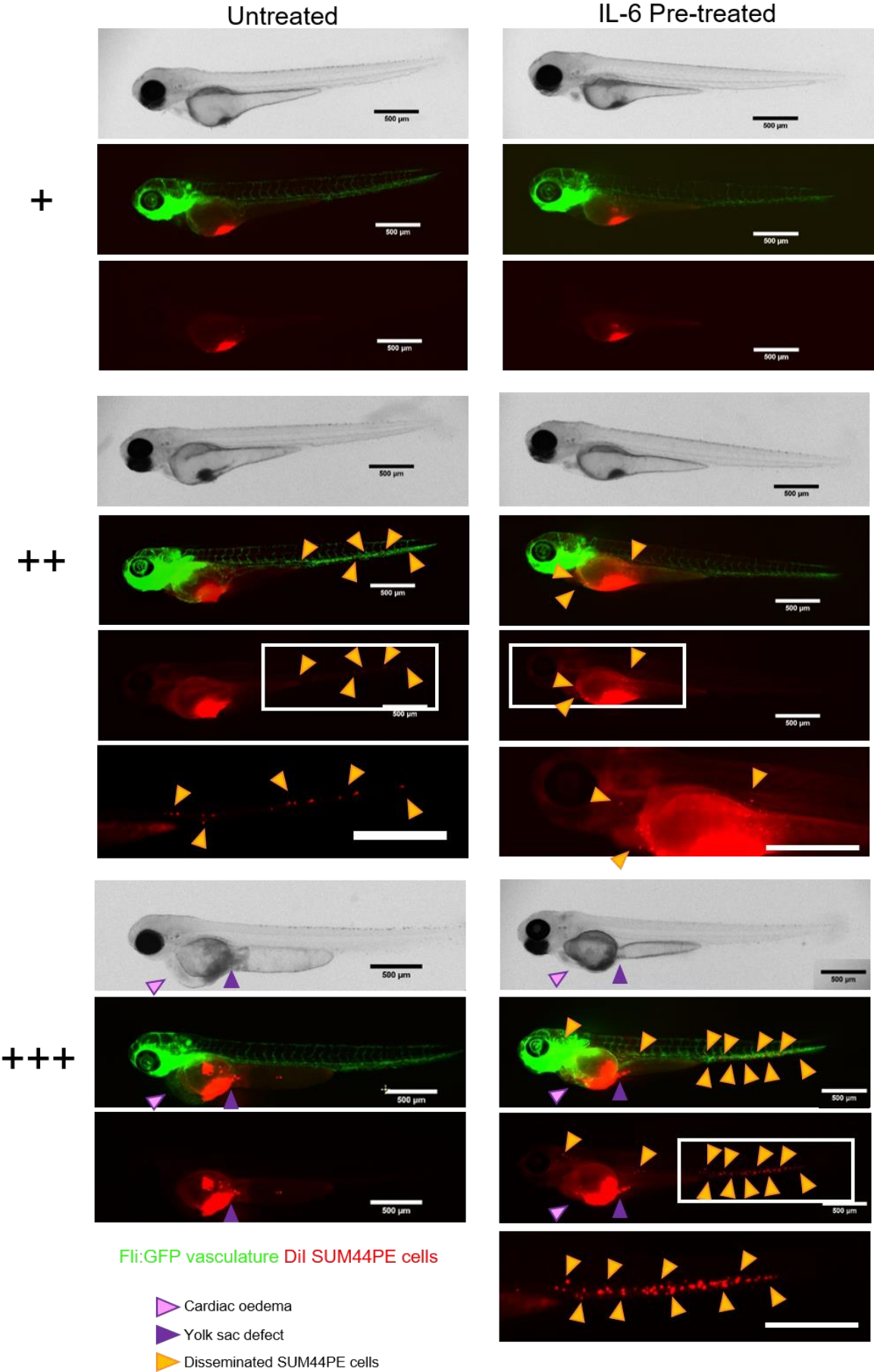


Figure 6-9. Representative images of 1dpi/3dpf zebrafish embryos. At 1dpi, embryos are sorted into groups based on tumour size, (+) for tumours smaller than the eye (top), (++) for tumours the size of the eye (middle), (+++) for embryos larger than the eye (bottom). Left column shows embryos injected with untreated SUM44PE cells, right column shows those injected with IL6 treated cells. Yellow arrows indicate disseminated cells, cardiac oedema and yolk sac defects indicated with purple outlined and filled arrows, respectively. White boxes indicate zoomed in regions in panel below, showing disseminated cells. Scale bars indicate 500 μ m, images taken on a mesoscope at 3.2x magnification, GFP labelled vasculature in green and Dil-dyed SUM44PE cells in red.

48-hours after injection (2dpi/4dpf), embryos with and without disseminated SUM44PE cells outside of the yolk sac detectable at 3.2x on the mesoscope were quantified (Figure 6-10A). The percentage of embryos with disseminated cells was calculated for each group (Figure 6-10C-E)

Embryos without any detectable disseminated cells (no MET, Figure 6-10A, left) and with disseminated cells (MET, Figure 6-10A, right) were found in embryos injected with both untreated and IL-6 pre-treated SUM44PE cells. In MET embryos, SUM44PE cells were disseminated to the heart region, within the subintestinal vein (SIV), which is the net-like vessel structure above the yolk sac, throughout the head region, the caudal haematopoietic tissue (CHT) and in the tail fin (Figure 6-10A right and B).

In the (+) group, there were significantly more embryos with disseminated tumour cells in those injected with IL-6 pre-treated cells than those injected with untreated cells (n=4, Figure 6-10B). This trend was replicated in the (++) group (Figure 6-10C, control=55.6%, IL-6=88.24%, n=3, p = 0.1358), though there were fewer (++) embryos than (+) embryos. There were insufficient replicates with enough (+++) embryos to assess. Overall, when assessing all embryos injected with either untreated SUM44PE cells or IL-6 pre-treated cells, regardless of primary tumour size (as in early experiments (n=2), tumour size was not classified), there were significantly more MET embryos in the IL-6 pre-treated group (Figure 6-10D, control=54.81%, IL-6=90.59%, n=6 independent experiments, paired t-test p-value=0.0071). This demonstrates that IL-6 pre-treated cells more readily disseminate from the primary site in the ZFEX model, showing that IL-6 promotes ILC SUM44PE cell migration *in vitro* and *in vivo*.

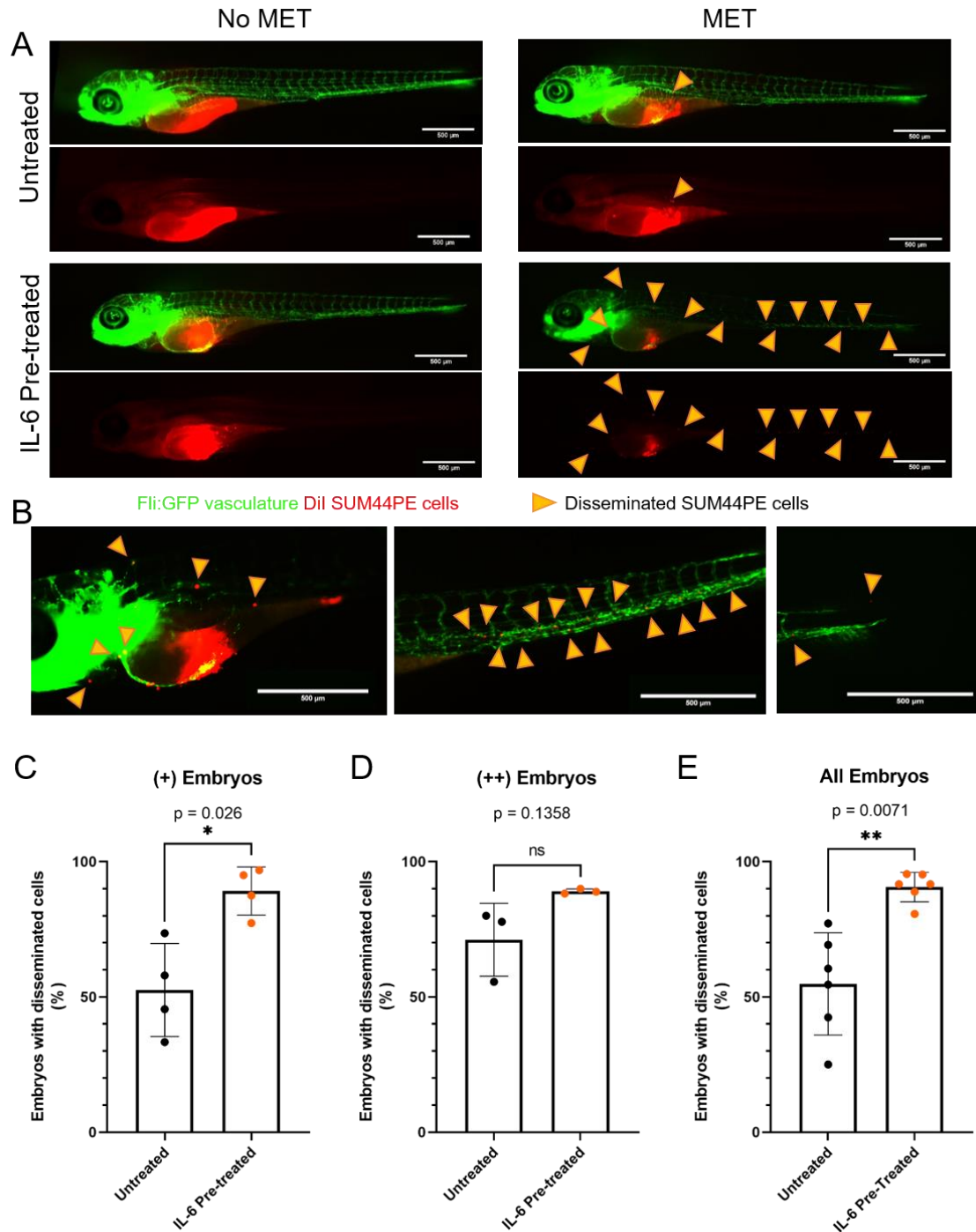


Figure 6-10. IL-6 pre-treated SUM44PE cells more readily disseminate in ZFEX at 2dpi. (A) Representative images of 2dpi/4dpf embryos injected with (top) untreated SUM44PE cells or (bottom) IL-6 pre-treated SUM44PE cells. Left - embryos without any disseminated cells (no MET), right - embryos with disseminated cells (MET), indicated by arrows. (B) Disseminated cells in the head region (left), CHT (middle) and tail fin (right). Images taken at 3.2x (A) and 8x (B) on a mesoscope, scale bars indicate 500 µm, vasculature in green (Fli:GFP) and SUM44PE cells in red (Dil-dye). Quantification of the percentage of (C) (+) embryos ($n=4$ independent repeats, >8 embryos in each group, mean no. embryos per group = 25), (D) (++) embryos ($n=3$ independent repeats, >9 embryos in each group, mean no. embryos per group = 14) and (E) all embryos ($n=6$ independent repeats, >8 embryos per repeat, mean no. of embryos per group = 37) with disseminated SUM44PE cells detected outside the yolk sac/PVS primary site. Bars show mean $\pm$ sd, two-way paired t-test in Graphpad Prism.

To quantify the number of disseminated SUM44PE cells, 2dpi/4dpf embryos were fixed and processed for whole-mount confocal microscopy (Figure 6-11A-C). The human and zebrafish nuclei were labelled with Hoechst, embryo vessels labelled by Fli:eGFP expression and SUM44PE cells labelled by the red Dil stain. Z-stacks of the entire embryo were taken at 20x with images stitched together in a montage to form an image of the entire embryo. Embryo processing and imaging was done by Aleksandra Rozyczko.

A mask was drawn around the PVS and yolk sac and Imaris software automatically generated masks for GFP+ vasculature and detected Dil-dyed SUM44PE cells (see 2.4.5). SUM44PE cells that were outside of the PVS/yolk sac region and had exited the vasculature were referred to as disseminated cells. The number of disseminated cells outside the yolk and vasculature was quantified in (+) embryos, and embryos injected with IL-6 pre-treated cells had more disseminated cells, although this was non-significant. When the percentage of disseminated cells relative to total SUM44PE cells within the embryo was calculated, there was a significantly higher proportion of disseminated IL-6 pre-treated cells than untreated cells (Figure 6-11D).

This demonstrates that, not only does IL-6 drive migration of ILC cell lines *in vitro* (Figure 6-6), but it promotes dissemination of tumour cells in the ZFEX model. More embryos injected with IL-6 pre-treated cells have disseminated cells and of those with tumour cells at distant sites, the embryos injected with IL-6 pre-treated cells have more disseminated cells.

To assess whether the SUM44PE cells were proliferating within the embryos, EdU was added to the media of CIRC embryos at 2dpi/4dpf for 24 hours and embryos were culled at 3dpi/5dpf. All handling and culling of these embryos after 4dpf was carried out by Adelaide Young. This labelled both the cells of the embryos and the SUM44PE cells with EdU if they passed through S-phase during this time point, with labelling of the rapidly developing embryo acting as a positive control. A small proportion of SUM44PE cells in both the untreated and IL-6 treated groups were labelled with EdU (Figure 6-12), with EdU+ cells detected in the primary site and in distant sites such as the forebrain region and CHT. There was a trend towards SUM44PE cells in the IL-6 pre-treated group having a lower rate of proliferation (Figure 6-12C, Untreated mean = 3.2%, IL-6 Pre-treated mean = 1.9%, n=4 embryos in each group, p-value = 0.25),

tallying with the suppression of cell cycle related genes seen in Chapter 4. However, this was only carried out in one experimental repeat, using embryos in the 'CIRC' group - those with SUM44PE cells injected directly into the circulation. More independent repeats are needed to verify this and this would be better carried out in embryos with defined primary tumour sizes and no immediate disseminated cells.

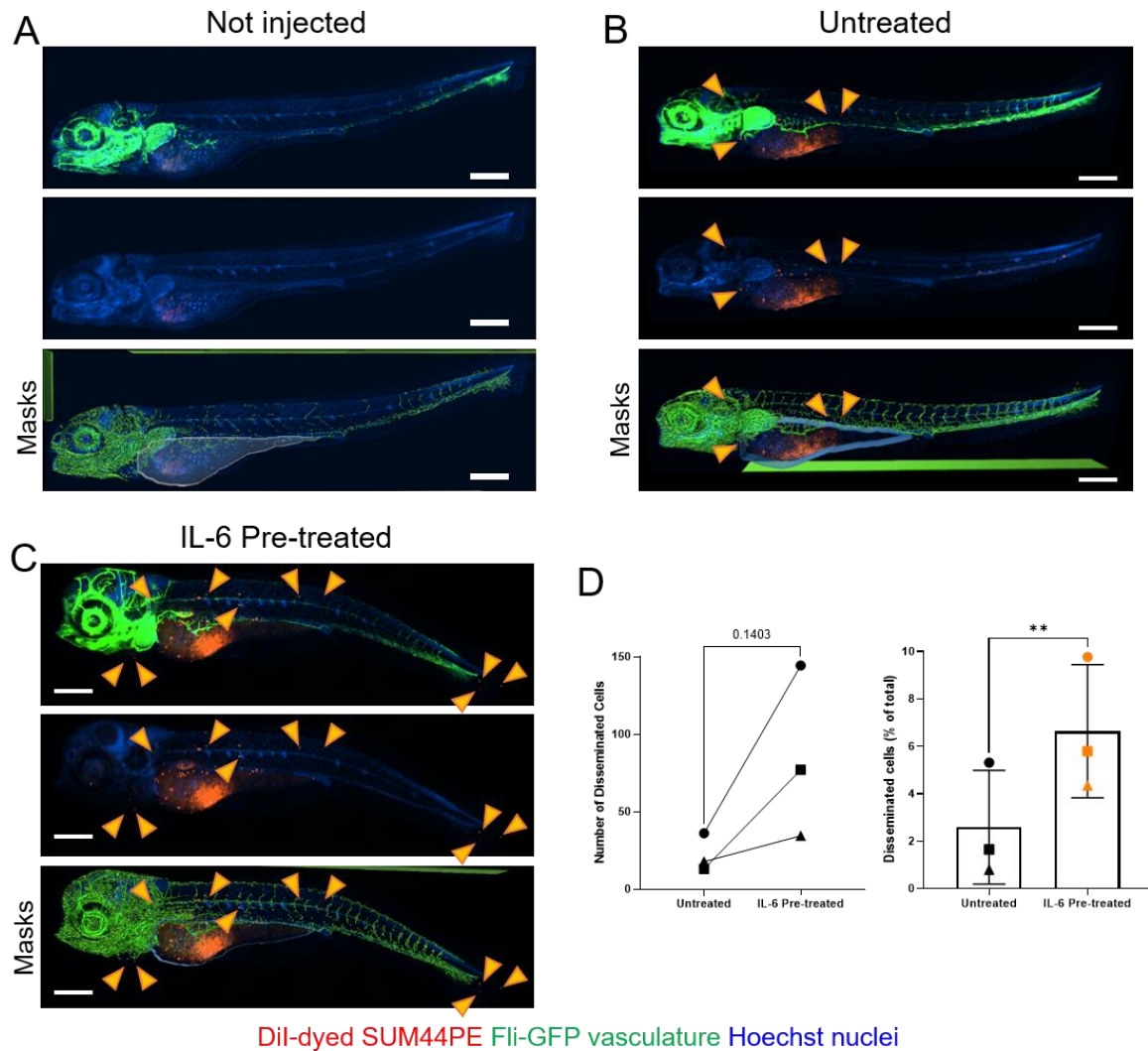


Figure 6-11. IL-6 pre-treatment increases the number of disseminated SUM44PE cells in the ZFEX model. Representative 3D projected stitched images of 2dpi/4dpf embryos taken at 20x magnification on the Dragonfly confocal, scale bars represent 300 μ m. (A) Embryo not injected with SUM44PE cells, (B) embryo injected with untreated cells and (C) embryo injected with IL-6 pre-treated cells. Bottom images show embryos with overlay of masks drawn around yolk sac and automated masks for GFP-labelled vasculature generated in Imaris software. Orange arrows indicate disseminated cells outside of the yolk and vessel masks. (D) Quantification of (left) the absolute number of disseminated cells in (+) embryos, and (right) the percentage of all SUM44PE cells in the embryo that are disseminated detected automatically in Imaris software, showing mean $\pm$ sd, cells. N=3 independent repeats with 3-4 embryos in each group for each repeat, each point represents the mean number or percentage of disseminated cells in each repeat, different repeats represented by different shapes. For left and right, paired t-test in Graphpad prism, ** $p=0.0043$.

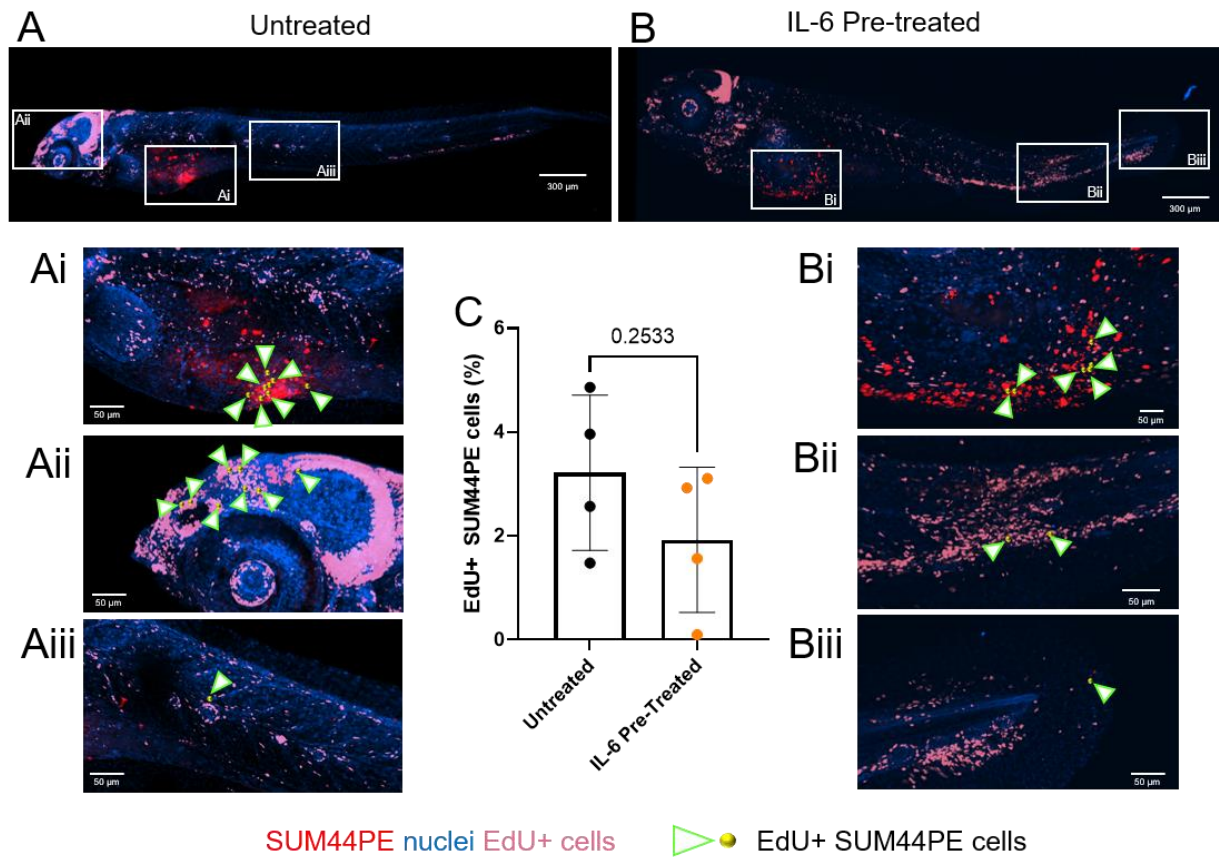


Figure 6-12. IL-6 pre-treated SUM44PE cells trend towards being less proliferative in the ZFEX model at 3dpi/5dpf. Representative images of 3dpi/5dpf CIRC embryos after 24 hours of EdU labelling, EdU labelled cells in pink, SUM44PE cells in red and nuclei stained with Hoechst in blue, EdU+ SUM44PE cells indicated with green arrows with yellow spots from Imaris software detection. (A) Embryo injected with untreated SUM44PE cells, with inserts showing EdU+ SUM44PE cells in (Ai) PVS primary site, (Aii) head (forebrain and olfactory bulb) region and (Aiii) CHT. (B) Embryo injected with IL-6 pre-treated SUM44PE cells with inserts showing EdU+ SUM44PE cells in (Bi) the PVS primary site, (Bii) the CHT and (Biii) the tail fin. Imaging on Dragonfly confocal at 20x with montage of z-stacks of the entire embryo by Aleksandra Rozyczko, scale bar in (A) and (B), 300 μ m, (A/Bi-iii) scale bar 50 μ m. (C) Quantification of the percentage of total SUM44PE cells throughout the embryo that are EdU+, SUM44PE and EdU spots automatically detected in Imaris software, with EdU+ SUM44PE cells determined by an overlap of SUM44PE and EdU spots. Data from one experiment with 4 embryos each in the untreated or pre-treated groups, each point represents one embryos, bars show mean $\pm$ sd, unpaired two-tailed t-test in Graphpad Prism.

6.2.5 IL-6 alters vasculature in the ZFEX model

From the RNAseq analysis in Chapter 4, CAF CM treated SUM44PE cells were enriched for gene sets involved in angiogenesis and expression of these gene sets were IL-6 dependent (Figure 4-10). Due to the transparent nature and the GFP-labelling of the vasculature of the Tg(Fli:eGFP) zebrafish embryos, this is an ideal model to begin to interrogate the effect of IL-6 on angiogenesis.

Evidence of angiogenesis could be seen in some embryos injected with both untreated and IL-6 pre-treated SUM44PE cells as early as 1 dpi (Figure 6-13A), with newly formed blood vessels beginning to extend towards the primary site from the subintestinal vein (SIV), the net-like structure over the top of the yolk sac, and from the duct of Cuvier (DoC) at the front of the yolk sac (see Figure 6-7A for anatomy). At 2dpi/4dpf, abnormal blood vessels could clearly be seen in many embryos injected with untreated and IL-6 pre-treated cells (Figure 6-13B), including embryos both with and without disseminated cells. However, it is important to note that no abnormal blood vessels were seen in any embryos that were not injected with cells and that in SUM44PE injected embryos there was wide variability in terms of vessel formation, structure and branching.

Unfortunately, at the time of collection, the percentage of all embryos with evidence of angiogenesis was not quantified on the mesoscope, but was quantified in a random sample of embryos from the confocal microscopy images (Figure 6-14). A total of 28 embryos (untreated $n=15$, +IL-6 $n=13$) were assessed for angiogenesis. From this sample, 53% and 47% of embryos injected with untreated and IL-6 treated SUM44PE cells had evidence of angiogenesis ($p>0.999$, Fisher's exact test), respectively. Of the embryos with angiogenesis, there was a significantly larger volume of blood vessels infiltrating into primary tumours formed of IL-6 pre-treated cells than untreated cells (Figure 6-14B and C). This suggests SUM44PE cells can induce abnormal blood vessel formation in the ZFEX model and that IL-6 pre-treatment may enhance this effect. There was no correlation between infiltrating vessel volume and disseminated cells (Pearson's $r = -0.08$, $n=11$), suggesting dissemination and angiogenesis are independent of each other. A larger sample size and scoring at later times following injection are required.

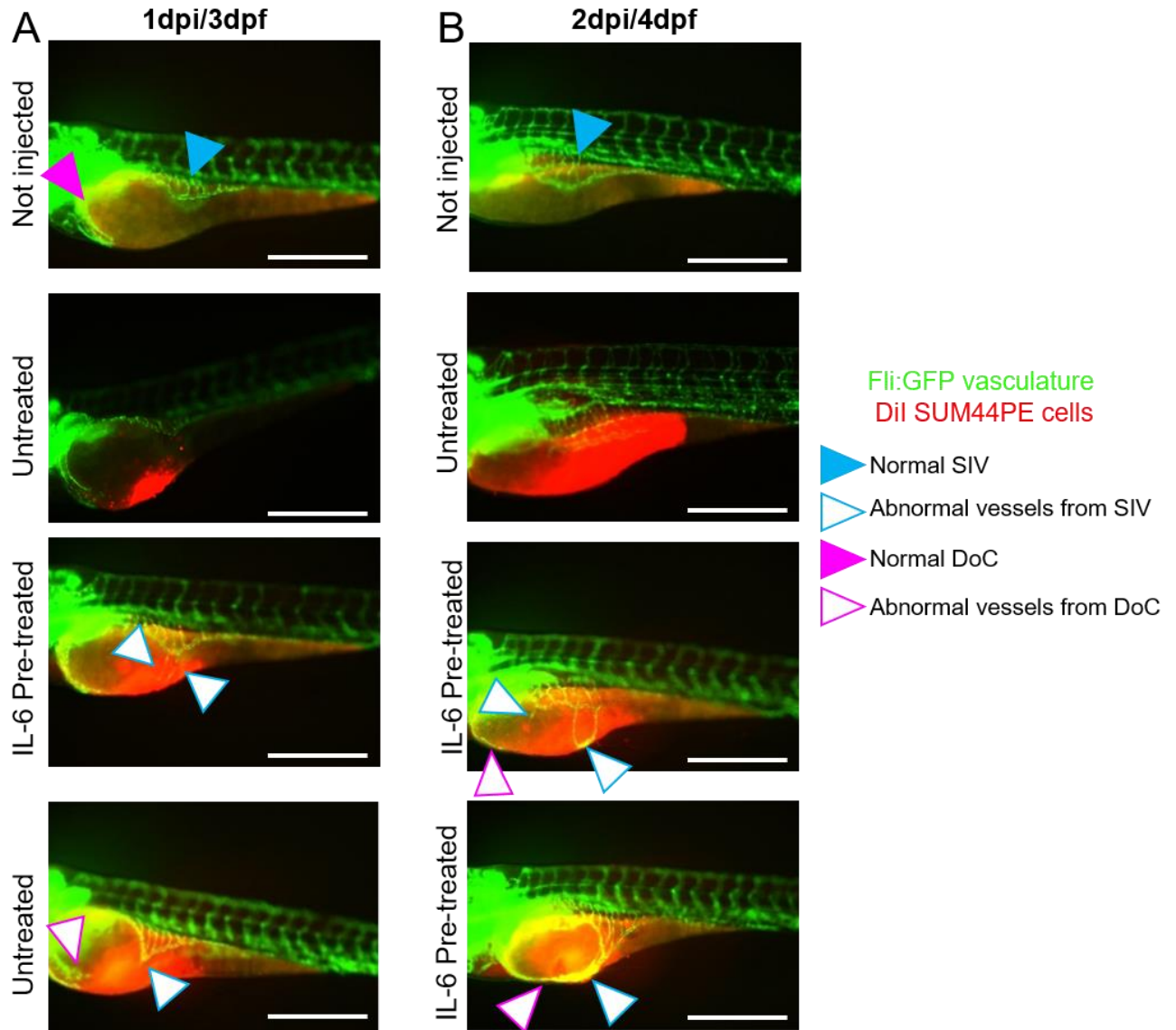


Figure 6-13. SUM44PE cells promote angiogenesis in ZFEX model. Representative images of the yolk sac region of control embryos (not injected, top) and embryos injected with SUM44PE cells at (A) 1dpi/3dpf and (B) 2dpi/4dpf, showing the range of abnormal blood vessel growth. Top images in (A) and (B) show embryos not injected with SUM44PE cells, with the normal SIV blood vessel structure indicated with blue arrows and normal DoC structure indicated with a pink arrow. Abnormal blood vessels extending from the SIV indicated with white arrows with blue outlines and from the DoC with pink outlines. Images taken at 8x on mesoscope, scale bars 500 μ m.

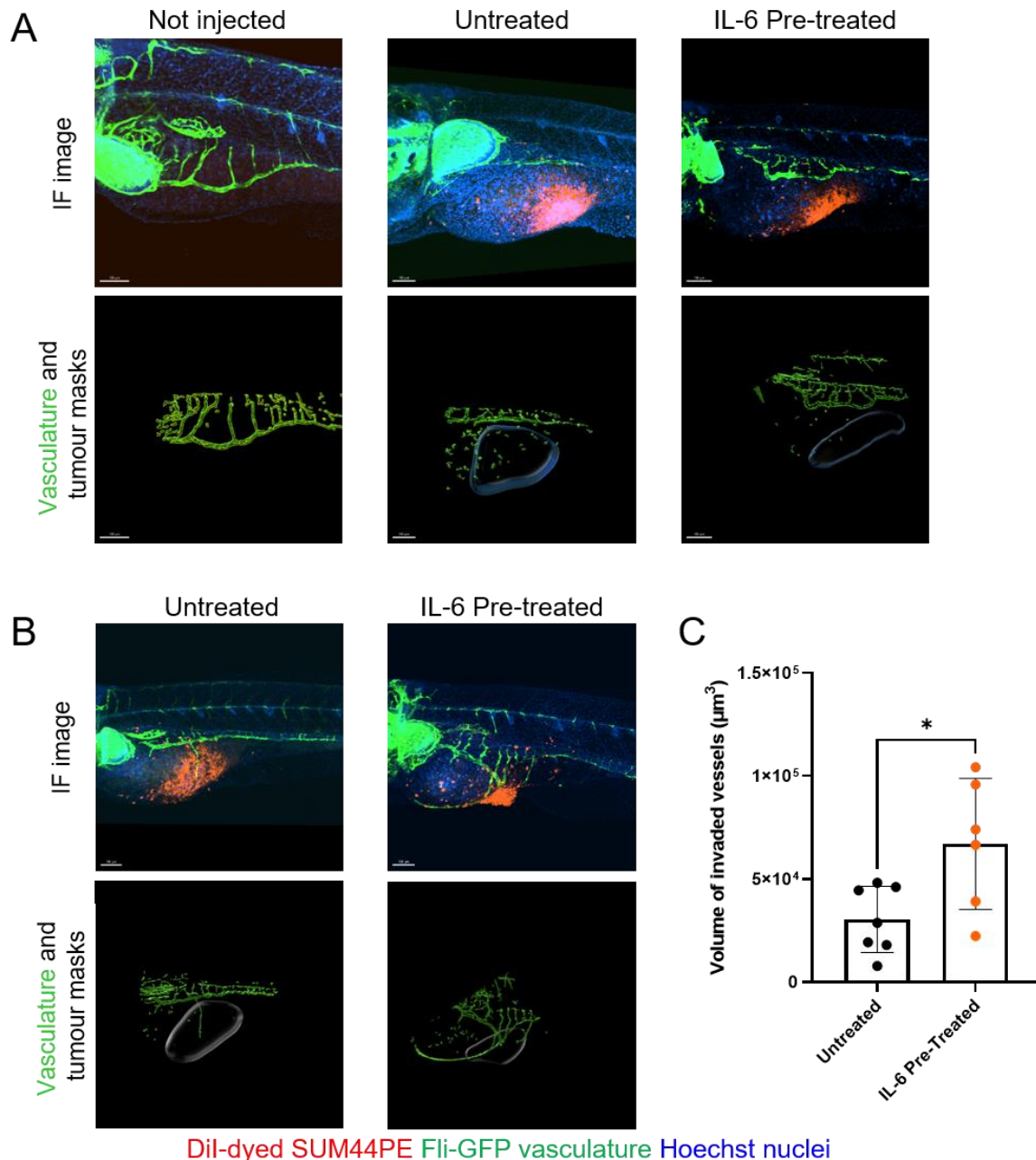


Figure 6-14. IL-6 pre-treated cells promote increased blood vessel density in the ZFEX assay. Representative 3D projections of 2dpi zebrafish embryos showing (A) no evidence of angiogenesis or (B) abnormal blood vessels invading into the primary tumour. Top images in (A) and (B) showing zoomed-in 3D projections of Z-stack of entire embryo taken at 20x, scale bar shows 100 μm . Nuclei labelled with Hoechst, Dil-dyed SUM44PE cells in red and blood vessels in green. Bottom images in (A) and (B) show manually drawn masks of the primary tumour (grey outline) and automatically detected masks of blood vessels in the SIV and yolk sac region (green), generated in Imaris. Imaris software was used to determine the volume of blood vessels within the primary tumour. (C) Graph showing the volume of blood vessels within the primary tumour excluding embryos with no evidence of angiogenesis. Each point represents one embryo, $n=7$ for each group, bar showing mean $\pm$ sd. Unpaired t-test in Graphpad prism, $p = 0.021$.

Chapter 6.3 Discussion

The work in the chapter demonstrates that IL-6 drives expression of EMT gene signatures and promotes a more mesenchymal-like phenotype in ILC cell lines. IL-6 is shown to increase the speed of SUM44PE cell migration *in vitro* and promote dissemination of cells in zebrafish embryo xenograft (ZFEX) model. This suggests that IL-6/STAT3 signalling promotes epithelial-to-mesenchymal plasticity and migration in ILC although further work is needed to elucidate the downstream transcriptional changes and mechanisms driving these phenotypes.

6.3.1 Proliferation

Proliferation of SUM44PE cells appears to be mostly dependent on IGF-1/Ins signalling rather than IL-6 (Figure 6-1 and Figure 6-2). E-cadherin loss has been shown to increase IGF1R activity, potentially by releasing the receptor from sequestration at adherens junctions⁵², with E-cadherin negative cell lines and ILC cell lines more responsive to IGF-1 stimulation²²², and more sensitive to IGF-1 pathway inhibitors, than E-cadherin+ IDC cell lines. ILC tumours have also been shown to have higher expression of IGF-1 and IGF1R than IDC tumours⁵², and work by our group has demonstrated that ILC tumours have higher stromal expression of *IGF1* and *PAPPA*¹³, a protease that degrades IGF-1 binding proteins, increasing IGF-1 local bioavailability. ILC CAF CM was found to contain a very high level of IGF-1 and other IGF-1 pathway proteins (Figure 3-6) and CAF CM stimulation increased proliferation of the SUM44PE cells. This appears to be dependent on IGF1R/InsR activity as it was reversed by treatment with the IGF1R/InsR inhibitor OSI-906 (Figure 6-1C-E). This further supports a paracrine IGF-1 pathway driving proliferation of ILC tumours.

Several lines of evidence support targeting the IGF-1 pathway in ILC^{13,52,53}. However, IGF-1 signalling acts through the Phospho-inositol-3-kinase (PI3K) pathway and activating PI3K pathway mutations are found in roughly 50% of ILC tumours³ which may hinder the efficacy of IGF-1 targeted therapies in this group of patients. For tumours without PI3K activating mutations, (SUM44PE cells do not have mutations in these pathways³⁷¹), there may be some benefit. However, there has been limited success in IGF-1 pathway targeted therapies in clinical trials⁵⁶, with development of OSI-906 discontinued³⁷², although some clinical benefit has been seen in ER+ breast cancer patients when an IGF-1-targeted monoclonal antibody was used in

combination with an mTOR inhibitor³⁷³. Greater success has been found in PI3K pathway targeted therapies in breast cancer, with a 30% response rate found in one trial using the *PIK3CA* targeted inhibitor alpelisib in ER+ breast cancers with PI3K pathway mutations⁵⁷. Although this trial included 4 ILC patients, their response was not recorded separately, highlighting the need for either increased ILC patient inclusion in trials or ILC-specific trials^{22,38,77}.

There was a trend towards decreased proliferation of ILC SUM44PE cells when stimulated with IL-6 (Figure 6-2) but more repeats are needed to confirm this. It may also be useful to assess proliferation over a longer period than just at 24 hours as an initial increase in pro-proliferative signalling was seen at 30 min of CAF CM stimulation in the RPPA (Figure 4-1), whereas there was a suppression of proliferation-related gene sets after 24 hours of stimulation (Figure 4-9 and 4-10). Furthermore, *IL6* expression is lowest in the proliferative subtype of ILC (Figure 4-9). However, IL-6/STAT3 activity has been shown to promote primary tumour growth in *in vivo* models of IDC^{165,176}. This may be dependent on STAT3 activity directly promoting cell cycle progression or potentially via other indirect mechanisms, such as immunosuppression¹⁵².

Interestingly, in IDC models, E-cadherin loss has been shown to reduce metastasis^{176,374} and metastatic outgrowth of dormant disseminated mesenchymal IDC cells is dependent on E-cadherin re-expression and reestablishment of a more proliferative epithelial cell state¹⁷⁶. pSTAT3 directly binding to the promoter region of *CDH1* in MCF7 cells has been suggested as the mechanism of reactivation of these dormant cells¹⁷⁶. However, inactivating mutations in *CDH1* are a defining feature of ILC and ILC GEMMs and intraductal xenografts still demonstrate high rates of metastasis^{35,37,80}, as is seen in ILC patients, though with long latency periods. The inability of IL-6/STAT3 signalling to reverse EMT by reactivation of E-cadherin expression in ILC cells in the metastatic site may contribute to the frequency of late recurrences seen in ILC. This suggests paracrine IL-6 signalling may promote both the low rate of proliferation seen in primary ILC tumours and characteristic late recurrences^{10,11}. This also implies a potential divergence in the effect of IL-6/STAT3 signalling in ILC and IDC cell lines and tumours, possibly dependent on E-cadherin, and warrants further investigation.

High stromal content, as assessed by H&E staining¹⁴¹ or by high stroma-related gene expression in the ‘reactive-like’ subtype of ILC³, has been associated with better outcomes in ER+ ILC. Paracrine IL-6/STAT3 dependent suppression of proliferation and related gene expression in ILC cells could potentially contribute to this, particularly as *IL6* is highly expressed in the reactive-like subtype of ILC (Figure 4-9E) and is highly secreted by ILC CAFs *in vitro* (Figure 4-6). However, these analyses of stromal content have not looked at specific stroma-expressed genes, specific subsets of stromal cell types or the ECM, and did not assess the immune composition of these tumours. Further analysis incorporating the composition of the stroma and stroma-specific gene expression, as well as assessing how immune and non-immune stromal cells influence each other, is needed to determine exactly how high stromal content contributes to outcomes for ILC patients and whether CAFs or even IL-6/STAT3 signalling could contribute to this.

6.3.2 EMT and migration *in vitro*

IL-6/STAT3 signalling has been shown to drive EMT in a wide range of cancers, including in breast cancer¹⁶⁷, promoting migration, metastasis and therapy resistance^{167,250} but has not previously been explored in ILC. Classically, EMT is characterised by down regulation of epithelial markers, in particular E-cadherin, and upregulation of mesenchymal markers, such as vimentin, and the transcription factors *SNAIL*, *SLUG*, *TWIST* and *ZEB1/2*. Although loss of E-cadherin is a key marker of EMT, most primary ILC tumours still display an epithelial phenotype, retaining expression of cytokeratins (CK8/18 and CK19) and show little upregulation of classical EMT markers such as vimentin and N-cadherin. In one study, only 3/148 ILC tumours displayed these mesenchymal-like features⁴⁶. However, ILC cell lines have been shown to have higher overall expression of EMT markers, such as *SNAIL*, *TWIST* and *VIM*, compared to IDC cell lines³⁵⁰ suggesting some form of intrinsic EMT-like program.

Expression of *SNAIL* is increased by IL-6 in CAF CM in the SUM44PE cells, although not significantly (Figure 6-3, logFC (CM/control) = 1.02, p-value=0.03, FDR=0.5), and has been shown to be a ILC-specific estrogen and tamoxifen induced gene^{252,350}. In ILC cell lines, *SNAIL* suppresses proliferation rather than driving an EMT phenotype, so may contribute to the IL-6 dependent suppression of proliferative genes rather than

EMT. The samples used for RNAseq were collected after 24 hours of IL-6 stimulation so more significant changes in classical EMT transcription factors may be detected at later time points. However, the enrichment for EMT gene sets and the elongated phenotype of the SUM44PE cells on IL-6 stimulation may be driven independently of the classical EMT transcription factors, which could account for the low level of classical mesenchymal markers seen in ILC tumours⁴⁶.

IL-6 stimulation of SUM44PE cells led to a more mesenchymal-like morphology on CAF-derived matrices (CDMs) and increased the speed of migration on collagen coated plates, demonstrating that the gene expression changes detected in Chapter 4 drive phenotypic changes that contribute to pro-tumorigenic features. A CRISPR screen targeting 26 genes from the IL-6 dependent gene signature (Figure 4-10 and Appendix 4) has since been carried out in the lab, using the 2D-random migration assay as the functional read out. This has identified *MUCL1*, a novel STAT3-target gene, and *BCL3*, a classical STAT3 target gene, as driving the IL-6 dependent increase in SUM44PE cell migration, and work is ongoing to validate these findings.

The ILC patient derived CAF-derived matrices (CDMs) used in Figure 6-4 and Figure 6-5 provide a biologically-relevant extracellular matrix for the SUM44PE cells to bind to, more closely recapitulating the breast microenvironment than plastic or collagen-coated plates^{200,375}. Unfortunately, due to issues with the Incucyte S3 during these experiments, cell migration could not be tracked on CDMs. Remodelling of the extracellular matrix has been shown to be key for ILC xenograft progression³⁵, although this has been shown to be driven by tumour cell intrinsic gene expression rather than induced by the microenvironment. As CAFs form such an abundant part of ILC tumours⁵⁵ and the extracellular matrix can influence many features of cancer, the CDMs provide a very useful *in vitro* model to study ILC.

The IL-6 driven elongation of the cells appears to be caused by transcriptional changes as an increase in AR was still seen for IL-6 pre-treated cells once IL-6 was removed. This IL-6 driven mesenchymal-like phenotype was dependent on STAT3 expression downstream of IL-6, providing further support to the central role of STAT3 in driving gene expression and functional changes in ILC cells. However, the elongation of unstimulated SUM44PE cells may also be dependent on STAT3 expression independent of IL-6, as an increase in mean AR was seen for WT and siCTRL

SUM44PE cells after 72 hours of adherence to CDM whilst siSTAT3 treated cells showed little change (Figure 6-5C). A wide range of signalling pathways can activate STAT3 independent of IL-6, and STAT3 activation has been shown to be induced by different extracellular matrices as well as contributing to remodelling these matrices. For example, melanoma cell adherence to stiffer extracellular matrices induced STAT3 activation and subsequent EMT gene expression³⁷⁶, and in a model of pancreatic ductal adenocarcinoma, STAT3 activation promoted extracellular matrix stiffening and remodelling, promoting progression *in vivo*³⁷⁷. Further work should be carried out to characterise the composition of the CDMs and any signalling changes induced in the cells on binding to CDMs.

Typically, ILC cells do not display a mesenchymal-like elongated phenotype in patient tumours^{24,46,378}, although spindle shaped ILC cells have been observed *in vitro* when cells are invading into collagen³⁶⁸. The mesenchymal-like phenotype observed in this study may be due to the 2D nature of the assays and the sparse seeding of cells – similar CAF driven EMT-like transcriptional programmes may occur in the breast but, potentially due to the constraints of the surrounding microenvironment and cells, may not result in an elongated phenotype. Organotypic assays or heterotypic spheroid assays may better model the behaviour of ILC cells in patients, although ILC cells are known to migrate as single cells rather than via collective migration, which is seen in IDC³⁶⁸, so our 2D migration assay may be a good representation of this behaviour.

6.3.3 Zebrafish xenograft model

6.3.3.1 SUM44PE cell dissemination in ZFEX model

The ZFEX model demonstrated that IL-6 significantly increases the likelihood of SUM44PE cell dissemination and increases the number of disseminated cells at 2 dpf. Despite the limitations of this model, not least the fact that it is not mammalian and embryos do not have breast tissue, this provides a useful proof-of-principle and is a fast, higher throughput, adaptable and cheaper alternative to the gold-standard intraductal mouse xenograft model of ILC.

Patient derived xenograft (PDX) models from colorectal⁸⁵, multiple-myeloma⁸⁹ and non-small cell lung (NSCLC) patients⁸⁷ implanted into zebrafish embryos have been shown to maintain features of the primary tumours as well as their responses to treatments, and recapitulate the same features seen when PDX models were

implanted into nude mice. Colorectal and multiple-myeloma PDXs implanted in the ZFEX model also show the same response and resistances to drugs as the patients they were derived from. Similarly, extravasation and invasion of breast cancer cell lines in the ZFEX model have been shown to increase with the increased 'aggressiveness' of these cell lines *in vitro* and *in vivo*³⁷⁹, although in this study cells were injected directly into the Duct of Cuvier so only models the later stages of metastasis. Co-injection with CAFs has been shown to increase dissemination of tumour cells in ZFEX models too^{86,90}. Importantly, dissemination of NSCLC PDX samples in the ZFEX assay closely correlated with lymph node involvement in the patients they were derived from⁸⁷. This implies that the increased dissemination of IL-6 treated SUM44PE cells may predict increased lymph node involvement in lobular patients with high IL-6 and downstream gene expression. High combined expression of the STAT3 target genes *MUC1* and *OSMR* has been shown to predict poor breast cancer specific survival for ER+ patients in general⁸³, but the impact of this pathway on clinical features has not been investigated in ILC specifically. Analysis of expression of *IL6*, *STAT3* and downstream genes in ILC tumour microarrays is ongoing by other members of our lab and will be correlated with clinical features.

Disseminated cells were seen in embryos injected with both untreated and IL-6 treated SUM44PEs, showing that intrinsic signalling in SUM44PE cells can promote dissemination, as has been shown in intraductal xenograft models³⁵, but is enhanced by IL-6 treatment. SUM44PE cells are injected into zebrafish embryos after a week of pre-treatment with IL-6. SUM44PE cells do not express IL-6, even at this time point (data not shown, RNAseq analysis ongoing), and although zebrafish do express IL-6 in embryos 2-4 dpf^{380,381}, there is a very low level of similarity between zebrafish IL-6 and human IL-6 (33.6%)³⁸¹. It is unlikely the human IL6R/gp130 complexes recognise the zebrafish IL-6 since the complexes are not able to recognise the much more similar mouse IL-6 protein (71.4% similar)¹⁶⁵. This suggests that the increase in migration and dissemination caused by IL-6 pre-treatment is driven by transcriptional programmes that are maintained after IL-6 withdrawal rather than continuous IL-6 stimulation in the embryo. This is consistent with the data in Figure 6-4, in which the IL-6 driven change in morphology was maintained and further promoted up to 48 hours after IL-6 withdrawal. An RNAseq experiment has since been carried out looking at gene expression changes driven by IL-6 in SUM44PE and MDA-MB-134VI ILC cancer cell

lines after 1 week of IL-6 stimulation, which will give us an insight into the transcriptional changes driving this pro-migratory mesenchymal-like phenotype. Knockdown of *MUCL1* and *BCL3* in SUM44PE cells – two genes in the IL-6 gene signature - reduced IL-6 dependent migration in the 2D *in vitro* assay in a CRISPR screen carried out by our group. The ZFEX assay will be used to determine whether knockdown of these genes also impacts the IL-6 dependent increase in dissemination of SUM44PE cells.

Proliferation of SUM44PE cells in both groups could be detected at 5dpf/3dpi in ZFEX model, although only a very low percentage of proliferating cells could be detected (Figure 6-12). Many factors may contribute to this low rate of proliferation, including the fact that even in optimal conditions *in vitro*, SUM44PE cells are slow growing with it taking up to 72 hours of incubation with EdU to label close to 100% of cells *in vitro* (Figure 6-1A). The embryos are maintained at 32°C during these experiments, which is a compromise between the optimal temperature for human cell growth (37°C) and for zebrafish embryo development (28-29°C) and may reduce SUM44PE cell proliferation. Also, although there is a high level of conservation between zebrafish and human genomes, human receptors may not recognise orthologous ligands secreted by zebrafish cells that promote proliferation such as IGF-1 and insulin which SUM44PE cells are dependent on for proliferation. In the human breast, the microenvironment would provide these pro-proliferative secreted factors. To further the relevance of the ZFEX model, SUM44PE cells could be co-injected with ILC CAFs to provide other factors, as well as IL-6, that promote their proliferation or zebrafish embryos engineered to express human IGF-1 and insulin could be used⁹¹. However, this low level of proliferation is consistent with what has been seen in mouse intraductal injection models of ILC³⁵ and in patients, with ILC tumours being very slow-growing and most recurrences being detected after 5-10 years³⁸². In future, we could use this assay to study the impact of IL-6/STAT3 pathway targeted drugs or endocrine targeted therapies on ILC cell line dissemination, as well as expand our studies to include some of the few ILC PDX models available⁷⁹.

6.3.3.2 SUM44PE driven angiogenesis in ZFEX model

CAF CM treatment of SUM44PE cells showed an IL-6 dependent significant enrichment for angiogenic gene sets in Chapter 4. Evidence of angiogenesis could be

seen in the ZFEX model injected with both untreated and IL-6 treated SUM44PE cells at 1 dpi, with abnormal blood vessels extending towards the primary site from the net-like SIV structure above the yolk. This demonstrates that the SUM44PE cells in this system are viable and capable of driving this pro-tumorigenic hallmark of cancer. Although no changes in expression of vascular endothelial growth factor genes was seen in the RNAseq data, unstimulated SUM44PE cells do express *VEGFA*, *B* and *C* (RNAseq data, not shown), which could account for the angiogenesis driven by untreated cells. Angiogenesis could be seen in embryos with and without disseminated cells and there was no correlation between number of disseminated cells and angiogenesis, suggesting that cell dissemination and angiogenesis happen via independent mechanisms. Usually assessments of angiogenesis in the ZFEX assay is carried out at 4 dpi, looking at vessel structures formed within primary tumours⁸⁵. However, our experiments were restricted to 2 dpi due to the requirement for a Home Office PPL for the use of zebrafish from 5 dpf. In future, experiments could be extended to 4dpi/6dpf to see if SUM44PE cells continue to survive in the primary site, how extensive angiogenesis is and if any disseminated cells begin to proliferate at distant sites.

Ideally, the rate of angiogenesis would have been scored at the time of collection, as was done for disseminated cells. From our small sample (n=28 out of 412 embryos used in total) analysed by confocal microscopy, there was no difference in the presence of angiogenesis between untreated and IL-6 treated cells. However, of the embryos with evidence of angiogenesis, there was a higher volume of vessels infiltrating into IL-6 pre-treated primary tumours. This suggests IL-6 induced gene expression leads to increased angiogenesis in this model. Interesting, the IL-6 induced gene *LRG1* (see Table xx) has been shown to drive angiogenesis in a number of systems^{275–277,346}. *LRG1* has been knocked down in SUM44PE cells as part of a CRISPR screen and did not impact IL-6 driven migration *in vitro* (data not shown) but may play a role in IL-6 driven angiogenesis, which will be assessed by using the *LRG1* knock down SUM44PE cells in the ZFEX model in future.

Chapter 6.4 Conclusions

The data in this chapter demonstrates that paracrine IL-6 signalling drives an EMT-like phenotype in ILC cells and promotes migration and dissemination *in vitro* and *in vivo* (Figure 6-15). I have also demonstrated the utility of CAF-derived matrices and the ZFEX assay to study ILC biology, a disease for which more models are necessary. Further work is needed to elucidate the downstream and potentially targetable mechanisms that IL-6 is activating to drive migration, to avoid inhibiting IL-6 anti-tumorigenic features and off-target effects in other tissues. This complements the work in Chapter 4 and 5, highlighting the functional consequences of the ILC-enriched IL-6/STAT3 signalling pathway.

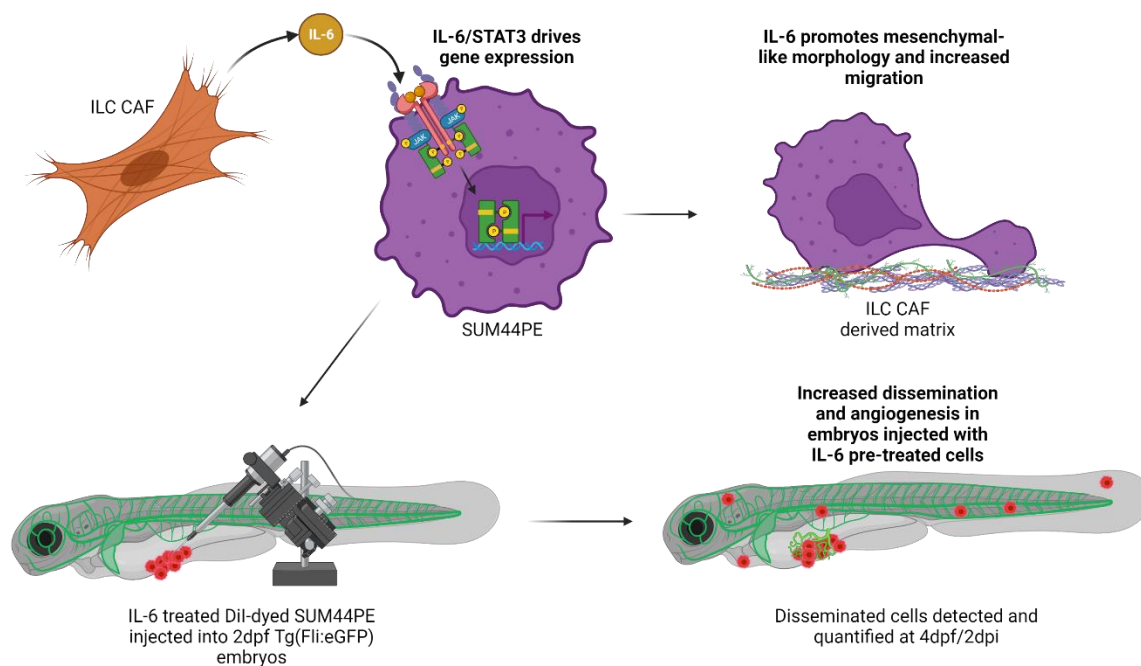


Figure 6-15. Schematic of the functional consequences of paracrine IL-6/STAT3 signalling.

Chapter 7 Discussion and Future Work

Although ILC and IDC have increasingly been recognised as distinct diseases, the vast majority of ILC research has focused on tumour cell autonomous phenotypes. This work forms part of one of the first in depth analysis of the ILC stroma and identified the paracrine IL-6/STAT3 pathway as enriched in ILC tumours and driving ILC tumour cell migration and dissemination, as well as suppressing estrogen receptor signalling.

Chapter 7.1 ILC Cancer Associated Fibroblasts

In Chapter 3, using primary patient derived CAFs, we identified a number genes specifically upregulated in the stroma of ILC tumours and not IDC tumours. This forms part of one of the first in depth transcriptomic studies of the ILC stromal TME. Although this work shows that ILC and IDC CAFs are not distinct based on their tumour of origin (Figure 3-1), 153 genes and a number of gene sets were differentially expressed between ILC and IDC CAFs. Interestingly, we found an ILC CAF enrichment for the mTOR signalling gene set (Figure 3-2) and high levels of IGF-1 (Figure 3-6) secreted by ILC CAFs. This suggests that the IGF-1/PI3K/mTOR pathway is enriched in both the tumour epithelium and stroma of ILC tumours, and that inhibition of this pathway may affect both compartments.

CECR1, a secreted adenosine deaminase, was found to be significantly more highly expressed in ILC CAFs, the stroma and bulk tumour analyses than in IDC (Table 3-3). This consistent enrichment in ILC suggests that it may play a role in ILC biology and should be investigated further. Generally, *CECR1* has been shown to be a good prognostic indicator in breast cancer²³⁷, but has not been investigated in ILC before. Of note, in survival analyses of bulk tumours, *CECR1* expression was associated with poorer outcome in ILC but not IDC tumours in the SCANB dataset¹³ but not in TCGA and METABRIC. It would be useful to interrogate the association of ILC stroma enriched genes with clinical features and outcomes in ILC tumours, looking specifically at stromal expression rather than in bulk analyses, by either IHC or spatial transcriptomics. The Edwards group at the University of Glasgow have recently completed a tumour microarray (TMA) containing 223 ILC tumours with 5 years follow up clinical and treatment data which could be used for this purpose.

This work was designed to investigate the differences between isolated ILC and IDC CAFs and to be used in conjunction with the laser capture microdissection dataset. Though we achieved the aims of this project, the use of bulk RNAseq from cultured CAFs limited our ability to investigate CAF intra-tumoural spatial and transcriptomic heterogeneity. The CAF lines did not fall into specific previously defined CAF subtype categories, rather showing mixed expression or high expression of some or all of the CAF subtype signatures investigated (Figure 3-3). This implies intra-cell line heterogeneity but could also be due to maintenance of these cell lines *in vitro*, whereas the subtype signatures were derived from freshly isolated CAFs^{101,254}. The work of Houthuijzen *et al* goes some of the way to answering the question of CAF heterogeneity in ILC, using single cell RNAseq (scRNAseq) to investigate normal fibroblasts and CAFs from the ILC *WapCre;E-cadherin^{fl/fl}/Pten^{fl/fl}* (WEPTn) mouse model. Although high stromal content as assessed by H&E and stroma-related gene expression has been associated with better outcomes for ILC patients^{3,141}, the underlying mechanisms of this association have not yet been elucidated. Further in depth investigation of CAFs in the TME of human ILC tumours using spatial transcriptomics and scRNAseq is currently ongoing by other research groups, and may begin to answer some of these unknowns.

Chapter 7.2 IL-6/STAT3 pathway in ILC

This work demonstrates that IL-6/STAT3 signalling is enriched in ILC and promotes a wide range of gene expression changes, including suppression of ER α signalling, and increased migration and dissemination in the ZFEX assay. Although the activation of this pathway promotes a wide range of pro-tumorigenic phenotypes, the influence of this signalling on therapy resistance and proliferation needs further interrogation and highlights that the IL-6/STAT3 pathway may not be a straightforward therapeutic target for ILC patients.

7.2.1 IL-6/STAT3 impact on breast cancer prognosis

Conflicting data exists for the correlation of IL-6 and STAT3 expression with survival outcomes and clinical features in breast cancer patients. Some studies have found that high pSTAT3 correlates with more aggressive disease¹⁵⁶, with TNBC having highest expression, whereas others have shown highest expression in lower risk

Luminal A breast cancers¹⁹⁴. High pSTAT3 or nuclear STAT3 has been found to be associated with improved breast cancer specific survival (BCSS) in ER+, ER- and HER2+ breast cancer^{178,193,194} and associated with positive prognostic indicators, such as ER expression, low Ki67 and Luminal A PAM50 subtype¹⁵⁷. Interestingly, in one study significantly more ILC tumours had high nuclear pSTAT3 (73%) compared to IDC tumours (44%). This may suggest that the association with 'good prognostic indicators' in this study may be caused by the enrichment for ILC tumours in the pSTAT3 high group. This highlights the need to investigate ILC tumours as a distinct group, although often studies are limited due to low numbers of ILC tumours reducing statistical power. However, combined high expression of *IL6* and pSTAT3¹⁶⁵ and high expression of STAT3 target genes, such as combined high *MUC1* and *OSMR*⁸³, has been shown to associate with decreased breast cancer specific survival (BCSS) in ER+ breast tumours, suggesting that combined IL-6/STAT3 or STAT3 target genes should be considered in survival analyses.

None of these studies specifically investigated STAT3 in ILC, with ILC tumours either excluded or combined with IDC tumours in an ER+ or Luminal group. Work is ongoing in our group to correlate ILC patient clinical features and outcomes with components of the IL-6/STAT3 pathway and some IL6GS genes, using a TMA with both ILC and IDC tumours. So far, this has corroborated higher stromal IL-6 in ILC compared to IDC patients and higher tumour cell expression of STAT3 in ILC, which correlates with poorer survival in ER+ ILC but not ER+ IDC (unpublished work by Zeanap Mabruk). However, this TMA only contained 62 ER+ ILC tumours compared to 476 ER+ IDC tumours, limiting the statistical power of this analysis. This work is currently being expanded into the newly completed Edwards' group TMA with 223 ILC tumours and expression of STAT3, pSTAT3, *IL6* and the IL6GS genes *MUCL1* and *BCL3* will be assessed.

7.2.2 IL-6/STAT3 pathway targeted therapy in breast cancer

Although IL-6/STAT3 signalling has been identified as a promising therapeutic target in breast cancer, clinical trials targeting this pathway have had limited success. Trials have recently been completed investigating the combination of ruxolitinib, a JAK inhibitor, with trastuzumab or exemestane in metastatic patients resistant to HER2-targeted therapy¹⁹⁵ or aromatase inhibitors¹⁹⁶, respectively. Though these

combinations were well tolerated, ruxolitinib treatment showed no clinical benefit in either trial. However, the STEM clinical trial (NCT02970682) using SFX-01, a stabilised sulforaphane derivative that directly targets STAT3⁸³, has shown efficacy in metastatic ER+ breast cancer patients progressing on endocrine therapy¹⁹⁷, though full results have not been published. This trial used 'RECIST' measurable disease as an inclusion criteria³⁸³, which often leads to exclusion of ILC patients due to the diffuse growth pattern of ILC primary and secondary tumours³⁸, so may not give any information on the efficacy of SFX-01 for ILC patients. However, SFX-01 has been shown to reduce stem-like features of metastatic ILC patient derived cell lines *in vitro* in the presence of endocrine therapies⁸³. This suggests that ILC patients could benefit from this treatment and highlights the need for ILC specific trials.

The CAF CM stimulated SUM44PE cells showed a number of gene expression changes associated with suppressed ER α -FOXA1 signalling and upregulation of genes related to endocrine-targeted therapy resistance (Chapter 5). STAT3 activation has been implicated in resistance to endocrine-targeted therapies^{83,131,163,169}, HER2-targeted therapies^{181,183–185,384}, chemotherapy^{167,312,385,386}, immunotherapy³⁴⁸ and CDK4/6 inhibitors (CDK4/6i)¹⁷⁵ in breast cancer. IL-6/STAT3 gene signatures are also upregulated in long-term estrogen deprived derivatives of both SUM44PE and MDA-MB-134VI ILC cell lines, even though these cell lines mediate resistance to aromatase inhibitors via different mechanisms⁶¹. This highlights STAT3 as a central node in breast cancer therapy resistance signalling. STAT3 has not been investigated in ILC specifically before and further work is needed to determine how paracrine activation of STAT3 and any of the downstream genes influence ILC cell line and patient therapy response and resistance.

7.2.3 STAT3 activity in CAFs

The work in this thesis has focused on STAT3 activity in ILC tumour cells. However, pSTAT3 was also detected in the nucleus of our primary CAFs in basal conditions (data not shown). STAT3 activity in breast CAFs has been shown to drive secretion of pro-migratory factors¹²⁸, including IL-6, as well promoting expression of microRNAs that promote endocrine targeted therapy resistance^{135,163} and increased metastasis *in vivo*¹⁸⁰. Stromal expression of STAT3 and expression of a CAF STAT3-dependent gene signature have been associated with poorer outcomes, particularly in basal-like

and TNBC^{121,178}, and tumours with high stromal STAT3 show reduced infiltration of cytotoxic T-cells. Although STAT3 activity in the TME has generally been shown to be pro-tumorigenic, investigation of the impact of STAT3 inhibition on the cells of the TME in ILC should be investigated. Currently, in collaboration with the Poli group at the University of Turin, we are carrying out analysis of conditioned media from a selection of our primary ILC and IDC CAFs after siSTAT3 treatment, to determine how STAT3 knockdown influences the human CAF secretome. Of note is that in preliminary analyses, STAT3 knockdown does not impact the expression of IL-6, contrasting with findings in mouse breast CAFs¹²⁸.

7.2.4 IL-6 driven migration

IL-6/STAT3 has been shown to promote migration and metastasis in many models of IDC and TNBC^{154,160,163,165} but this is the first work to demonstrate that IL-6 promotes migration and dissemination of ILC cells. Since writing this thesis, a CRISPR screen has been carried out, targeting 26 of the IL6GS genes, to determine if any of these downstream genes are driving the IL-6 dependent increase in 2D migration of SUM44PE cells. Knockdown of a number of genes suppressed the IL-6 induced increase in migration, with validation ongoing. Knockdown of *MUCL1*, a breast specific mucin and novel IL-6/STAT3 target gene, and *BCL3*, a transcription factor and well known IL-6/STAT3 target gene^{165,387}, look to be the most promising candidates. *MUCL1* in particular represents an interesting potential therapeutic target or biomarker, due to its tissue-restricted expression in healthy tissues, as it is exclusively expressed in mammary and salivary glands, and significantly increased expression in breast cancer compared to normal breast^{299,321,325}. *MUCL1* is particularly highly expressed in ILC compared to IDC (Figure 4-11) and has been implicated in driving EMT and migration³²³ in breast cancer cells downstream of HER2 amplification²⁹⁹ – however, this is the first work identifying *MUCL1* as an IL-6/STAT3 target. The ZFEX assay will be used to assess whether *MUCL1* and *BCL3* knockdown SUM44PE cells show reduced IL-6 induced dissemination.

Chapter 7.3 The IGF-1/PI3K pathway

The IGF-1 and downstream PI3K pathways have been shown to be hyper-activated in ILC due to E-cadherin loss driving autocrine signalling^{52,222} and activating mutations in *PIK3CA* and inactivating *PTEN* mutations³. Work by our group, using the laser capture microdissection data¹³, the FPPA (Chapter 3) and CAF CM stimulation of SUM44PE cells (Chapter 4 and 6), has also highlighted an enrichment for IGF-1 expression in the stroma of ILC and paracrine activation of the downstream PI3K pathway in ILC tumour cells, promoting proliferation. This provides further support to targeting this pathway in ILC.

Clinically, IGF-1 targeted therapies have not been very successful, possibly due to the multiple redundancies in the pathway and activating mutations in the PI3K pathway in ILC as well as adverse side effects⁵⁶. A number of other CAF-secreted factors, including IL-6, can also activate the PI3K pathway. However, OSI-906 treatment completely suppressed CAF CM induced SUM44PE proliferation (Fig 6-1), suggesting paracrine activation of IGF1R/InsR signalling alone promotes CAF CM induced proliferation *in vitro*. Targeting downstream components of the PI3K pathway may have more benefit in ILC, inactivating both cell autonomous and paracrine induced hyper activation of the pathway. However, a recent retrospective analysis of trials using PI3K α inhibitor alpelisib or the mTOR inhibitor everolimus in combination with ETs demonstrated no difference in progression free survival between ILC and IDC patients, despite the increased activity of the PI3K/mTOR pathway in ILC³⁶⁰. A recent trial using alpelisib as a monotherapy in advanced patients with PI3K mutant tumours did show clinical benefit in ER+ disease⁵⁷. As PI3K is so frequently mutated in ILC, this suggests alpelisib monotherapy may benefit advanced ILC patients. Interestingly, this trial included four ILC patients, and three out of the top four most responsive patients had *CDH1* mutations, although the trial did not specifically report on the response in ILC patients.

STAT3 activation downstream of PI3K β in a model of PTEN inactivation suppressed cytotoxic T-cell infiltration¹⁵² and, in the *K14Cre;E-cadherin^{fl/fl};Tp53^{fl/fl}* (KEP) ILC model, resistance to an mTOR inhibitor was caused by immunosuppression⁵⁴. Potentially, combined targeting of both the PI3K/mTOR pathway and the IL-6/STAT3 pathway could promote formation of an immune-responsive TME in ILC, improving response to

immune checkpoint inhibitors, which have had limited success in ILC⁷⁷. Both STAT3¹⁷⁵ and PI3K³⁸⁸ pathway activation have also been shown to promote CDK4/6i resistance, with the combination of fulvestrant, an AKT inhibitor and a CDK4/6i reducing growth of fulvestrant-CDK4/6i resistant PDX models. The interactions between these two ILC enriched pathways should be further investigated.

Chapter 7.4 Future *in vivo* work

This work *in vitro*, using the ZFEX assay and interrogating patient datasets has allowed us to extensively explore CAFs and paracrine signalling in ILC. However, our work could be further expanded and validated using more ILC cell line models, both mouse and human, and ILC PDXs *in vitro* and in the ZFEX assay. The ZFEX assay in particular is a valuable model for ILC as it is high throughput and uses smaller cell numbers than mouse models do. In future, we could use the ZFEX assay to investigate IL-6/STAT3 targeted therapies as well as carrying out CAF co-injection experiments.

7.4.1 Intraductal co-injection studies

Ideally, in future this work will be followed up with mouse intraductal xenograft models, which have been shown to closely recapitulate the progression and dissemination of ILC tumours³⁵, with intraductally injected SUM44PE cells metastasising to the gastrointestinal tract, peritoneum and ovaries, common ILC metastatic sites. However, a limitation of these experiments in studying IL-6 signalling is that ILC cell lines do not express IL-6 (Figure 4-6)¹⁴ and the human gp130/IL6R receptor complex does not recognise mouse IL-6 that would be secreted by the host microenvironment¹⁶⁵. To account for this, lobular cell lines could be engineered to constitutively express human IL-6, as was done by Siersbæk *et al* with T47D cells¹⁶⁵, or mice could be engineered to express human IL-6 under the control of a fibroblast or CAF specific promoter.

Previous unpublished work by our lab has shown that intraductal co-injection of both human and mouse ILC cell lines with primary human or mouse CAFs, respectively, enhances primary tumour growth. However, metastases were not assessed and transplanted CAFs have been shown to be lost 4 weeks post implantation in xenografts³⁸⁹. This suggests that co-injection of CAFs supports the early stages of tumorigenesis in these co-injection studies. In the ZFEX model, co-injection of CAFs with tumour cells has also been shown to enhance tumour cell dissemination

dependent on paracrine signalling⁹⁰ and direct contact⁸⁶. Co-injection of ILC CAFs and tumour cells, either using IL-6 knockdown CAFs or SUM44PE cells with IL6GS genes knocked out, into immunocompromised mice or the ZFEX assay could be used in future.

7.4.2 ILC GEMMs

An issue with co-injection xenograft studies is that they do not precisely recapitulate tumorigenesis *in vivo* – both our primary CAF lines and ILC tumour cell lines are derived from established tumours rather than early stage ILC. Two well characterised genetically engineered mouse models (GEMMs) of ILC have been developed - the KEP (*K14-Cre;Cdh1^{fl/fl};Tp53^{fl/fl}*) model³⁶ and WEPTn (*Wap-Cre;Cdh1^{fl/fl};Pten^{fl/fl}*)³⁷ both model ILC *in vivo*. The KEP model is thought to more closely recapitulate a model of E-cadherin loss rather than ILC specifically, as p53 is only mutated in 8% of ILC tumours compared to 44% of IDC tumours³. The WEPTn model more closely recapitulates ILC tumorigenesis and has more relevant mutations, with *PTEN* inactivation leading to hyper-activation of the PI3K pathway as seen in human tumours. Interestingly, *in vitro* studies in our group using tumour cells and CAFs derived from both these models show that KEP CAFs do not express IL-6 whereas WEPTn CAFs do. Stimulation of ‘normal’ fibroblasts (NFs) derived from the WEPTn model *in vitro* with tumour cell conditioned media has also been shown to induce NFs to express IL-6 (Julia Houthuijzen, personal communication). Members of our group are currently investigating the paracrine IL-6/STAT3 pathway in the WEPTn model *in vitro* and may potentially carry out future *in vivo* experiments via intraductal co-injection studies or use of the WEPTn GEMM.

Chapter 7.5 The immune microenvironment of ILC

The aims of my PhD research focused on understanding and characterising the stromal tumour microenvironment, specifically the role of CAFs, and the interactions between CAFs and tumour cells in ILC. Whilst CAFs are the predominant non-transformed cells in the TME of ILC⁵, cells of the immune system and their crosstalk with tumour and stromal cells play an essential role in tumour biology.

A number of CAF subtypes, including the CAF-S1 and iCAF subtypes defined by Costa *et al* and Houthuijzen *et al*, have been shown to be immunosuppressive^{101,254}. Interestingly, further classification of the CAF-S1 subtype by scRNAseq identified some clusters that promoted expression of Treg markers and immune checkpoint receptors in T cells, and were enriched in immunotherapy resistant melanoma patients, whereas other clusters were correlated with higher T-cell cytolytic activity¹¹⁸. STAT3 activity in breast cancer cell lines has also been shown to promote an immunosuppressive microenvironment. Interestingly, PI3K β induced activation of STAT3 promotes immune evasion in a model of *PTEN* loss in breast cancer¹⁵², a common mutation in ILC. However, breast cancer secreted IL-6 has been shown to decrease the number of suppressive Treg cells in the TME of a Kindlin-1 null model of breast cancer³⁹⁰. This highlights the complex impact of CAFs and CAF-induced signalling in tumour cells on the immune microenvironment.

Both the ZFEX and SUM44PE intraductal mouse xenografts take place in immune-compromised animals, as the adaptive immune system is not fully active in zebrafish until 3 weeks post fertilisation^{380,391}. This means the interactions between CAFs, tumour cells and the adaptive immune system, as well as the influence of IL-6 signalling on these interactions, cannot be studied in these models. The use of GEMMs as well as *in vitro* assays and examination of patient samples would allow investigation of these interactions.

However, both zebrafish embryos and nude mice have components of the innate immune system, so these models can be used to understand interactions between ILC cell lines, CAFs and innate immune cells such as macrophages. Neutrophils and macrophages are recruited to colorectal tumour cell masses in the ZFEX assay as early as 1dpi but only impact tumour engraftment at 4dpi³⁹², two days later than our experiments were extended to. Colorectal cell lines that persist in the ZFEX assay to 4dpi polarised zebrafish macrophages towards a pro-tumorigenic M2-like phenotype whereas cell lines with a lower engraftment rate are cleared by neutrophils and M1-like macrophages. Interestingly, in a single pilot experiment, 3/3 embryos injected with IL-6 pre-treated SUM44PE cells still had tumour cell masses at 4dpi whereas only 1/3 embryos injected with untreated cells still had tumour masses (data not shown). Although we do not know the reasons, this could reflect increased immunosuppressive

signalling and M2-macrophage polarisation by IL-6 treated SUM44PE cells. This is preliminary data and so follow up experiments - potentially using transgenic zebrafish embryos with fluorescently labelled macrophages³⁸⁰ - need to be carried out before any conclusions can be drawn.

A tumour-associated macrophage signature has been shown to correlate with poorer survival in ILC patients³⁹³ and an increased in the ratio of pro-tumorigenic M2 macrophages to anti-tumorigenic M1 macrophages has been found in the TME of ILC compared to IDC¹⁴. Whilst ILC tumour cell conditioned media was found to contribute to this polarisation, partially via IL-33, the authors did not investigate signalling from other components of the TME, such as CAFs, and how paracrine signalling from the TME to tumour cells could influence this polarisation. The adenosine deaminase *CECR1*, identified as an ILC tumour, stroma and CAF enriched gene (Chapter 3), has been shown to promote M2 polarisation²³³ and correlate with M2-related gene signatures in glioma and breast cancer²³⁵. IL-6 has also been shown to directly promote M2 polarisation^{177,394}. *IL6* was also significantly more highly expressed in the 'immune-related' subtype of ILC tumours, which have high expression of macrophage-related genes³, relative to the proliferative subtype in the TCGA dataset (Figure 4-9). Interestingly, CAF CM treated SUM44PE cells were also enriched for the 'immune-related' ILC subtype signature, and the IL6GS genes *CHI3L2*, *S100A9* and *IL4R* have been shown to influence macrophage infiltration and phenotypes^{328,332,395}.

This suggests that reciprocal signalling between different compartments of the TME could potentially contribute to the enrichment of M2-macrophages in ILC and emphasises the need to investigate the interactions between tumour cells and the stromal and immune compartments in concert. Limited success has been seen for T-cell targeted immunotherapy in ILC, with only 6/23 patients responding to PD-1 blockade in the recent GELATo trial⁷⁷, with 4/6 of the responsive tumours being triple negative ILC. Focus on macrophage-targeted therapies in ILC could lead to better outcomes for some ILC patients but requires further research.

Chapter 7.6 Conclusions

Overall, this work highlights differences between ILC and IDC CAFs (Chapter 3), provides further evidence for enhanced IGF-1/PI3K signalling in both the stroma and epithelium of ILC (Chapters 3, 4 and 6), and identified the ILC-enriched paracrine IL-6/STAT3 pathway (Chapter 4) suppressing ER α signalling (Chapter 5) and driving increased migration, angiogenesis and dissemination (Chapter 6), summarised in Figure 7-1.

Our primary ILC and IDC CAFs provide a useful model for studying tumour-stroma interactions *in vitro* and the ZFEX assay is a novel model for the investigation of ILC tumour cell dissemination. The CAF RNAseq dataset in particular also provides a valuable resource for the further study of the ILC TME. Further work is needed to fully untangle the complex consequences of the IL-6/STAT3 pathway in ILC, particularly focusing on the role of STAT3 in therapy resistance and potentially its influence on the immune microenvironment of ILC.

This research highlights the unique features and distinct biology of ILC and emphasises the need to investigate ILC as a distinct disease from the more common and more frequently researched IDC subtype of breast cancer. Future work should focus on validation of these findings in larger ILC patient datasets as well as relevant *in vivo* models of ILC and expanding studies to investigate how tumour cells, CAFs and immune cells interact and influence the biology of ILC tumours.

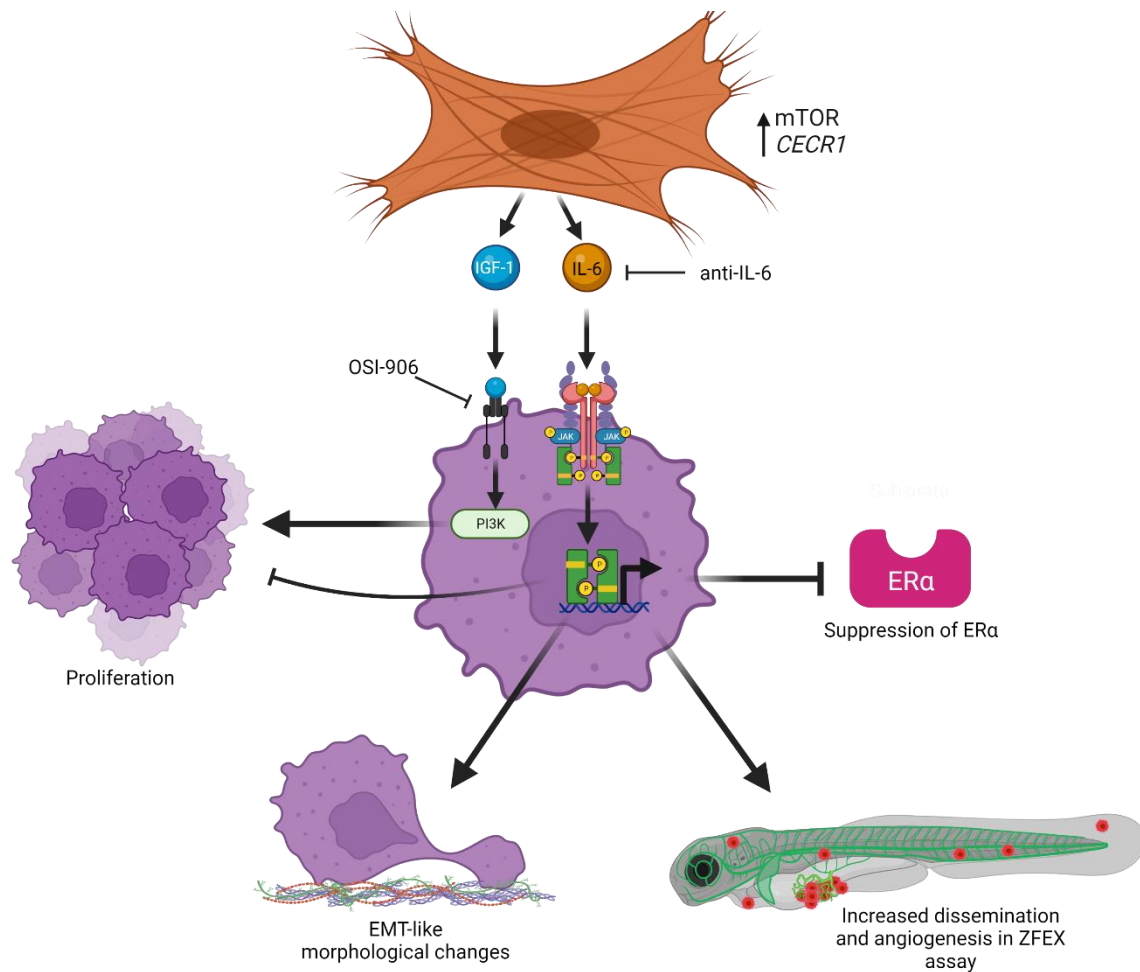


Figure 7-1. Summary of the findings of this thesis. ILC CAFs express high levels of the adenosine deaminase *CECR1* and *mTOR* gene sets relative to IDC CAFs. ILC CAFs secrete IGF-1 and IL-6. Paracrine IGF-1 signalling promotes ILC cell line proliferation. Paracrine IL-6 signalling promotes STAT3 activation and subsequent gene expression. IL-6/STAT3 signalling suppresses ERα expression and pro-proliferative genes and promotes EMT and tumour cell dissemination and angiogenesis in the ZFEX model. The IL-6/STAT3 pathway is enriched in ILC tumours and may represent a potential therapeutic target for ILC.

References

1. Cancer Research UK. Breast cancer statistics | Cancer Research UK. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero> (2014).
2. Office for National Statistics. Number of people diagnosed with Lobular Breast Cancer, England, 2006 to 2016 - Office for National Statistics. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/adhocs/009819numberofpeoplediagnosedwithlobularbreastcancerengland2006to2016>.
3. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163**, 506–519 (2015).
4. Desmedt, C. *et al.* Genomic Characterization of Primary Invasive Lobular Breast Cancer. *J. Clin. Oncol.* **34**, 1872–1881 (2016).
5. Nakagawa, S. *et al.* Tumor microenvironment in invasive lobular carcinoma: possible therapeutic targets. *Breast Cancer Res. Treat.* **155**, 65–75 (2015).
6. Millar, E. K. A. *et al.* Tumour stroma ratio assessment using digital image analysis predicts survival in triple negative and luminal breast cancer. *Cancers (Basel)*. **12**, 1–14 (2020).
7. Arpino, G., Bardou, V. J., Clark, G. M. & Elledge, R. M. Infiltrating lobular carcinoma of the breast: Tumor characteristics and clinical outcome. *Breast Cancer Res.* **6**, 7–11 (2004).
8. Rakha, E. A. *et al.* Invasive lobular carcinoma of the breast: Response to hormonal therapy and outcomes. *Eur. J. Cancer* **44**, 73–83 (2008).
9. Chen, Z. *et al.* Invasive lobular carcinoma of the breast: A special histological type compared with invasive ductal carcinoma. *PLoS One* **12**, 1–17 (2017).
10. Adachi, Y. *et al.* Comparison of clinical outcomes between luminal invasive ductal carcinoma and luminal invasive lobular carcinoma. *BMC Cancer* **16**, 1–9 (2016).
11. Korhonen, T. *et al.* The impact of lobular and ductal breast cancer histology on the metastatic behavior and long term survival of breast cancer patients. *Breast* **22**, 1119–1124 (2013).
12. Mathew, A. *et al.* Distinct Pattern of Metastases in Patients with Invasive Lobular Carcinoma of the Breast. *Geburtshilfe Frauenheilkd.* **77**, 660–666 (2017).
13. Gómez-Cuadrado, L. *et al.* Characterisation of the Stromal Microenvironment in Lobular Breast Cancer. *Cancers (Basel)*. **14**, 904 (2022).
14. Onkar, S. *et al.* Immune landscape in invasive ductal and lobular breast cancer reveals a divergent macrophage-driven microenvironment. *Nat. Cancer* **4**, 516–534 (2023).
15. Desmedt, C. *et al.* Immune infiltration in invasive lobular breast cancer. *J. Natl. Cancer Inst.* **110**, 768–776 (2018).

16. Tubiana-Hulin, M. *et al.* Response to neoadjuvant chemotherapy in lobular and ductal breast carcinomas: a retrospective study on 860 patients from one institution. *Ann. Oncol.* **17**, 1228–1233 (2006).
17. Truin, W. *et al.* Differences in Response and Surgical Management with Neoadjuvant Chemotherapy in Invasive Lobular Versus Ductal Breast Cancer. *Ann. Surg. Oncol.* **23**, 51–57 (2015).
18. Riba, L. A., Russell, T., Alapati, A., Davis, R. B. & James, T. A. Characterizing response to neoadjuvant chemotherapy in invasive lobular breast carcinoma. *J. Surg. Res.* **233**, 436–443 (2019).
19. Timbres, J. *et al.* Survival outcomes in invasive lobular carcinoma compared to oestrogen receptor-positive invasive ductal carcinoma. *Cancers (Basel)*. **13**, 1–19 (2021).
20. Dossus, L. & Benusiglio, P. R. Lobular breast cancer: Incidence and genetic and non-genetic risk factors. *Breast Cancer Res.* **17**, 1–8 (2015).
21. Kotsopoulos, J. *et al.* Risk factors for ductal and lobular breast cancer: Results from the nurses' health study. *Breast Cancer Res.* **12**, (2010).
22. Van Baelen, K. *et al.* Current and future diagnostic and treatment strategies for patients with invasive lobular breast cancer. *Annals of Oncology* vol. 33 769–785 at <https://doi.org/10.1016/j.annonc.2022.05.006> (2022).
23. Iorfida, M. *et al.* Invasive lobular breast cancer: subtypes and outcome. *Breast Cancer Res. Treat.* **133**, 713–723 (2012).
24. Christgen, M. *et al.* Lobular breast cancer: Histomorphology and different concepts of a special spectrum of tumors. *Cancers (Basel)*. **13**, 1–34 (2021).
25. Filho, O. M. *et al.* Relative effectiveness of letrozole compared with tamoxifen for patients with lobular carcinoma in the BIG 1-98 Trial. *J. Clin. Oncol.* **33**, 2772–2778 (2015).
26. Rakha, E. A. *et al.* Invasive lobular carcinoma of the breast: Response to hormonal therapy and outcomes. *Eur. J. Cancer* **44**, 73–83 (2008).
27. Okines, A. *et al.* Clinical outcomes in patients with triple negative or HER2 positive lobular breast cancer: a single institution experience. *Breast Cancer Res. Treat.* **192**, 563–571 (2022).
28. Joshi, U. *et al.* Clinical outcomes and prognostic factors in triple-negative invasive lobular carcinoma of the breast. *Breast Cancer Res. Treat.* **200**, 217–224 (2023).
29. Dalenc, F. *et al.* Impact of lobular versus ductal histology on overall survival in metastatic breast cancer: a French retrospective multicentre cohort study. *Eur. J. Cancer* **164**, 70–79 (2022).
30. Gluz, O. *et al.* Efficacy of deescalated chemotherapy according to PAM50 subtypes, immune and proliferation genes in triple-negative early breast cancer: Primary translational analysis of the WSG-ADAPT-TN trial. *Int. J. Cancer* **146**, 262–271 (2020).

31. Conforti, F. *et al.* Biological and clinical features of triple negative Invasive Lobular Carcinomas of the breast. Clinical outcome and actionable molecular alterations. *Breast* **59**, 94–101 (2021).
32. Porter, A. J., Evans, E. B., Foxcroft, L. M., Simpson, P. T. & Lakhani, S. R. Mammographic and ultrasound features of invasive lobular carcinoma of the breast. *J. Med. Imaging Radiat. Oncol.* **58**, 1–10 (2014).
33. Wilson, N., Ironside, A., Diana, A. & Oikonomidou, O. Lobular Breast Cancer: A Review. *Front. Oncol.* **10**, 1–13 (2021).
34. Richard, F. *et al.* Characterization of stromal tumor-infiltrating lymphocytes and genomic alterations in metastatic lobular breast cancer. *Clin. Cancer Res.* **26**, 6254–6265 (2020).
35. Sflomos, G. *et al.* Intraductal xenografts show lobular carcinoma cells rely on their own extracellular matrix and LOXL1. *EMBO Mol. Med.* **13**, 1–19 (2021).
36. Derksen, P. W. B. *et al.* Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell* **10**, 437–449 (2006).
37. Boelens, M. C. *et al.* PTEN Loss in E-Cadherin-Deficient Mouse Mammary Epithelial Cells Rescues Apoptosis and Results in Development of Classical Invasive Lobular Carcinoma. *Cell Rep.* **16**, 2087–2101 (2016).
38. Abel, M. K. *et al.* Decreased enrollment in breast cancer trials by histologic subtype: does invasive lobular carcinoma resist RECIST? *npj Breast Cancer* **7**, 18–20 (2021).
39. Du, T. *et al.* Invasive lobular and ductal breast carcinoma differ in immune response, protein translation efficiency and metabolism. *Sci. Rep.* **8**, 1–11 (2018).
40. Hogan, M. P. *et al.* Comparison of 18F-FDG PET/CT for systemic staging of newly diagnosed invasive lobular carcinoma versus invasive ductal carcinoma. *J. Nucl. Med.* **56**, 1674–1680 (2015).
41. Covington, M. F. *et al.* Prospective Pilot Study of 18 F-Fluoroestradiol PET/CT in Patients With Invasive Lobular Carcinomas . *Am. J. Roentgenol.* (2023) doi:10.2214/ajr.22.28809.
42. Ulaner, G. A. *et al.* Head-to-Head Evaluation of 18F-FES and 18F-FDG PET/CT in Metastatic Invasive Lobular Breast Cancer. *J. Nucl. Med.* **62**, 326–331 (2021).
43. Eshet, Y. *et al.* The Role of 68Ga-FAPI PET/CT in Detection of Metastatic Lobular Breast Cancer. *Clin. Nucl. Med.* **48**, 228–232 (2023).
44. Pestalozzi, B. C. *et al.* Distinct clinical and prognostic features of infiltrating lobular carcinoma of the breast: Combined results of 15 International Breast Cancer Study Group clinical trials. *J. Clin. Oncol.* **26**, 3006–3014 (2008).
45. Thill, M. *et al.* Treatment and outcome in metastatic lobular breast cancer in the prospective German research platform OPAL. *Breast Cancer Res. Treat.* **198**, 545–553 (2023).

46. McCart Reed, A. E. *et al.* An epithelial to mesenchymal transition programme does not usually drive the phenotype of invasive lobular carcinomas. *J. Pathol.* **238**, 489–494 (2016).
47. Tasdemir, N. *et al.* Comprehensive phenotypic characterization human invasive lobular carcinoma cell lines in 2D and 3D cultures. *Cancer Res.* **78**, 6209–6222 (2018).
48. Tasdemir, N. *et al.* Proteomic and transcriptomic profiling identifies mediators of anchorage-independent growth and roles of inhibitor of differentiation proteins in invasive lobular carcinoma. *Sci. Rep.* **10**, 1–15 (2020).
49. Schackmann, R. C. J. *et al.* Cytosolic p120-catenin regulates growth of metastatic lobular carcinoma through Rock1-mediated anoikis resistance. *J. Clin. Invest.* **121**, 3176–3188 (2011).
50. Van De Ven, R. A. H. *et al.* Nuclear p120-catenin regulates the anoikis resistance of mouse lobular breast cancer cells through Kaiso-dependent Wnt11 expression. *DMM Dis. Model. Mech.* **8**, 373–384 (2015).
51. Rätze, M. A. K. *et al.* Loss of E-cadherin leads to Id2-dependent inhibition of cell cycle progression in metastatic lobular breast cancer. *Oncogene* (2022) doi:10.1038/s41388-022-02314-w.
52. Nagle, A. M. *et al.* Loss of E-cadherin enhances IGF1-IGF1R pathway activation and sensitizes breast cancers to anti-IGF1R/InsR inhibitors. *Clin. Cancer Res.* **24**, 5165–5177 (2018).
53. Teo, K. *et al.* E-cadherin loss induces targetable autocrine activation of growth factor signalling in lobular breast cancer. *Sci. Rep.* **8**, (2018).
54. Klarenbeek, S. *et al.* Response of metastatic mouse invasive lobular carcinoma to mTOR inhibition is partly mediated by the adaptive immune system. *Oncoimmunology* **9**, (2020).
55. Nakagawa, S. *et al.* Tumor microenvironment in invasive lobular carcinoma: possible therapeutic targets. *Breast Cancer Res. Treat.* **155**, 65–75 (2016).
56. Lee, J. S., Tocheny, C. E. & Shaw, L. M. The Insulin-like Growth Factor Signaling Pathway in Breast Cancer: An Elusive Therapeutic Target. *Life* vol. 12 at <https://doi.org/10.3390/life12121992> (2022).
57. Savas, P. *et al.* Alpelisib Monotherapy for PI3K-Altered, Pretreated Advanced Breast Cancer: A Phase II Study. *Cancer Discov.* **12**, 2058–2073 (2022).
58. Fuso, P. *et al.* PI3K Inhibitors in Advanced Breast Cancer: The Past, The Present, New Challenges and Future Perspectives. *Cancers (Basel)*. **14**, 1–21 (2022).
59. Nardone, A. *et al.* A Distinct Chromatin State Drives Therapeutic Resistance in Invasive Lobular Breast Cancer. *Cancer Res.* 3673–3686 (2022) doi:10.1158/0008-5472.can-21-3186.
60. Sottnik, J. L. *et al.* Mediator of DNA Damage Checkpoint 1 (MDC1) is a Novel Estrogen Receptor Coregulator in invasive Lobular Carcinoma of the Breast.

- Mol. Cancer Res.* **19**, 1270–1282 (2021).
61. Du, T. *et al.* Key regulators of lipid metabolism drive endocrine resistance in invasive lobular breast cancer. *Breast Cancer Res.* **20**, 1–15 (2018).
 62. Sikora, M. J. *et al.* WNT4 mediates estrogen receptor signaling and endocrine resistance in invasive lobular carcinoma cell lines. *Breast Cancer Res.* **18**, 1–16 (2016).
 63. Zhou, J. *et al.* Stem Cells and Cellular Origins of Breast Cancer : Updates in the Rationale , Controversies , and Therapeutic Implications. **9**, 1–12 (2019).
 64. Zhang, M., Lee, A. V & Rosen, J. M. The Cellular Origin and Evolution of Breast Cancer. (2023).
 65. Gross, K., Wronski, A., Skibinski, A., Phillips, S. & Kuperwasser, C. Cell Fate Decisions During Breast Cancer Development. doi:10.3390/jdb4010004.
 66. Rorive, S. *et al.* Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity ´. (2015) doi:10.1038/nature14665.
 67. Derksen, P. W. B. *et al.* Mammary-specific inactivation of E-cadherin and p53 impairs functional gland development and leads to pleomorphic invasive lobular carcinoma in mice. *DMM Dis. Model. Mech.* **4**, 347–358 (2011).
 68. Arthur, L. M. *et al.* Molecular changes in lobular breast cancers in response to endocrine therapy. *Cancer Res.* **74**, 5371–5376 (2014).
 69. Selli, C. *et al.* Molecular changes during extended neoadjuvant letrozole treatment of breast cancer: distinguishing acquired resistance from dormant tumours. *Breast Cancer Res.* **21**, 1–15 (2019).
 70. Gao, J. J. *et al.* CDK4/6 inhibitor treatment for patients with hormone receptor-positive, HER2-negative, advanced or metastatic breast cancer: a US Food and Drug Administration pooled analysis. *Lancet Oncol.* **21**, 250–260 (2020).
 71. NCT02764541. Palbociclib and Endocrine Therapy for LOBular Breast Cancer Preoperative Study (PELOPS). <https://clinicaltrials.gov/show/NCT02764541> (2016).
 72. Truin, W. *et al.* Differences in Response and Surgical Management with Neoadjuvant Chemotherapy in Invasive Lobular Versus Ductal Breast Cancer. *Ann. Surg. Oncol.* **23**, 51–57 (2016).
 73. Tubiana-Hulin, M. *et al.* Response to neoadjuvant chemotherapy in lobular and ductal breast carcinomas: A retrospective study on 860 patients from one institution. *Ann. Oncol.* **17**, 1228–1233 (2006).
 74. Bajrami, I. *et al.* E-Cadherin/ROS1 inhibitor synthetic lethality in breast cancer. *Cancer Discov.* **8**, 498–515 (2018).
 75. Agostinetto, E. *et al.* ROSALINE: a phase II, neoadjuvant study targeting ROS1 in combination with endocrine therapy in invasive lobular carcinoma of the breast. *Futur. Oncol.* **18**, 2383–2392 (2022).
 76. Crizotinib in Lobular Breast, Diffuse Gastric and Triple Negative Lobular Breast

- Cancer or CDH1-mutated Solid Tumours - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03620643?term=rolo&cond=Lobular+Breast+Carcinoma&draw=2&rank=1>.
77. Voorwerk, L. *et al.* PD-L1 blockade in combination with carboplatin as immune induction in metastatic lobular breast cancer: the GELATO trial. *Nat. Cancer* **4**, (2023).
 78. Christgen, M. & Derksen, P. W. B. Lobular breast cancer: Molecular basis, mouse and cellular models. *Breast Cancer Res.* **17**, 1–9 (2015).
 79. Sflomos, G. *et al.* Atlas of Lobular Breast Cancer Models: Challenges and Strategic Directions. *Cancers (Basel)*. **13**, 5396 (2021).
 80. Derksen, P. W. B. *et al.* Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell* **10**, 437–449 (2006).
 81. Desmedt, C. *et al.* Genomic Characterization of Primary Invasive Lobular Breast Cancer. in *Journal of Clinical Oncology* vol. 34 1872–1880 (2016).
 82. Guillen, K. P. *et al.* A human breast cancer-derived xenograft and organoid platform for drug discovery and precision oncology. *Nat. Cancer* **3**, 232–250 (2022).
 83. Simões, B. M. *et al.* Targeting STAT3 signaling using stabilised sulforaphane (SFX-01) inhibits endocrine resistant stem-like cells in ER-positive breast cancer. *Oncogene* **39**, 4896–4908 (2020).
 84. Sflomos, G. *et al.* A Preclinical Model for ER α -Positive Breast Cancer Points to the Epithelial Microenvironment as Determinant of Luminal Phenotype and Hormone Response. *Cancer Cell* **29**, 407–422 (2016).
 85. Fior, R. *et al.* Single-cell functional and chemosensitive profiling of combinatorial colorectal therapy in zebrafish xenografts. *Proc. Natl. Acad. Sci.* **114**, E8234–E8243 (2017).
 86. Hurtado, P. *et al.* Modelling metastasis in zebrafish unveils regulatory interactions of cancer-associated fibroblasts with circulating tumour cells. *Front. Cell Dev. Biol.* **11**, 1–16 (2023).
 87. Ali, Z. *et al.* Zebrafish patient-derived xenograft models predict lymph node involvement and treatment outcome in non-small cell lung cancer. *J. Exp. Clin. Cancer Res.* **41**, 1–18 (2022).
 88. Zhao, C. *et al.* A novel xenograft model in zebrafish for high-resolution investigating dynamics of neovascularization in tumors. *PLoS One* **6**, 1–9 (2011).
 89. Lin, J. *et al.* A clinically relevant in vivo zebrafish model of human multiple myeloma to study preclinical therapeutic efficacy. *Blood* **128**, 249–252 (2016).
 90. Liu, C. *et al.* A zebrafish model discovers a novel mechanism of stromal fibroblast-mediated cancer metastasis. *Clin. Cancer Res.* **23**, 4769–4779 (2017).

91. Wang, L. *et al.* Patient-derived Heterogeneous Xenograft Model of Pancreatic Cancer Using Zebrafish Larvae as Hosts for Comparative Drug Assessment. *J. Vis. Exp.* **2019**, 2–7 (2019).
92. Costa, B., Estrada, M. F., Mendes, R. V. & Fior, R. Zebrafish avatars towards personalized medicine— a comparative review between avatar models. *Cells* **9**, (2020).
93. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
94. Cerami, E. *et al.* The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
95. Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, (2013).
96. Onuchic, V. *et al.* Epigenomic Deconvolution of Breast Tumors Reveals Metabolic Coupling between Constituent Cell Types. *Cell Rep.* **17**, 2075–2086 (2016).
97. Chhabra, Y. & Weeraratna, A. T. Fibroblasts in cancer: Unity in heterogeneity. *Cell* **186**, 1580–1609 (2023).
98. Sahai, E. *et al.* A framework for advancing our understanding of cancer-associated fibroblasts. *Nat. Rev. Cancer* (2020) doi:10.1038/s41568-019-0238-1.
99. Houthuijzen, J. M. & Jonkers, J. Cancer-associated fibroblasts as key regulators of the breast cancer tumor microenvironment. *Cancer Metastasis Rev.* **37**, 577–597 (2018).
100. Özdemir, B. *et al.* Depletion of Carcinoma-Associated Fibroblasts and Fibrosis Induces Immunosuppression and Accelerates Pancreas Cancer with Reduced Survival. *Cancer Cell* **25**, 719–734 (2014).
101. Houthuijzen, J. M. *et al.* CD26-negative and CD26-positive tissue-resident fibroblasts contribute to functionally distinct CAF subpopulations in breast cancer. *Nat. Commun.* **14**, 183 (2023).
102. Raz, Y. *et al.* Bone marrow-derived fibroblasts are a functionally distinct stromal cell population in breast cancer. *J. Exp. Med.* **215**, 3075–3093 (2018).
103. Andreucci, E. *et al.* Extracellular Lactic Acidosis of the Tumor Microenvironment Drives Adipocyte-to-Myofibroblast Transition Fueling the Generation of Cancer-Associated Fibroblasts. *Cells* **12**, (2023).
104. Zhu, Q. *et al.* Adipocyte mesenchymal transition contributes to mammary tumor progression. *Cell Rep.* **40**, (2022).
105. Bartoschek, M. *et al.* Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. *Nat. Commun.* **9**, 5150 (2018).
106. Kidd, S. *et al.* Origins of the tumor microenvironment: Quantitative assessment

- of adipose-derived and bone marrow-derived stroma. *PLoS One* **7**, (2012).
107. Sharon, Y. *et al.* Tumor-derived osteopontin reprograms normal mammary fibroblasts to promote inflammation and tumor growth in breast cancer. *Cancer Res.* **75**, 963–973 (2015).
 108. Mishra, P. J. *et al.* Carcinoma-associated fibroblast-like differentiation of human mesenchymal stem cells. *Cancer Res.* **68**, 4331–4339 (2008).
 109. Zhu, Y. *et al.* Breast Cancer Exosome-Derived miR-425-5p Induces Cancer-Associated Fibroblast-Like Properties in Human Mammary Fibroblasts by TGF β 1/ROS Signaling Pathway. *Oxid. Med. Cell. Longev.* **2022**, (2022).
 110. Butti, R. *et al.* Tumor-derived osteopontin drives the resident fibroblast to myofibroblast differentiation through Twist1 to promote breast cancer progression. *Oncogene* **40**, 2002–2017 (2021).
 111. Weber, C. E. *et al.* Osteopontin mediates an MZF1-TGF- β 1-dependent transformation of mesenchymal stem cells into cancer-associated fibroblasts in breast cancer. *Oncogene* **34**, 4821–4833 (2015).
 112. Jotzu, C. *et al.* Adipose tissue derived stem cells differentiate into carcinoma-associated fibroblast-like cells under the influence of tumor derived factors. *Cell. Oncol.* **34**, 55–67 (2011).
 113. Zeng, W. *et al.* CCL18 signaling from tumor-associated macrophages activates fibroblasts to adopt a chemoresistance-inducing phenotype. *Oncogene* **42**, 224–237 (2023).
 114. Shangguan, L. *et al.* Inhibition of TGF- β /Smad signaling by BAMBI blocks differentiation of human mesenchymal stem cells to carcinoma-associated fibroblasts and abolishes their protumor effects. *Stem Cells* **30**, 2810–2819 (2012).
 115. Toullec, A. *et al.* Oxidative stress promotes myofibroblast differentiation and tumour spreading. *EMBO Mol. Med.* **2**, 211–230 (2010).
 116. Calvo, F. *et al.* Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and maintenance of cancer-associated fibroblasts. *Nat. Cell Biol.* **15**, 637–646 (2013).
 117. Costa, A. *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **33**, 463-479.e10 (2018).
 118. Kieffer, Y. *et al.* Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discov.* **10**, 1330–1351 (2020).
 119. Pelon, F. *et al.* Cancer-associated fibroblast heterogeneity in axillary lymph nodes drives metastases in breast cancer through complementary mechanisms. *Nat. Commun.* **11**, 404 (2020).
 120. Bonneau, C. *et al.* A subset of activated fibroblasts is associated with distant relapse in early luminal breast cancer. *Breast Cancer Res.* **22**, 1–22 (2020).
 121. Avalle, L. *et al.* STAT3 induces breast cancer growth via ANGPTL4, MMP13 and STC1 secretion by cancer associated fibroblasts. *Oncogene* 1–12 (2022)

doi:10.1038/s41388-021-02172-y.

122. Matsumura, Y. *et al.* Stromal fibroblasts induce metastatic tumor cell clusters via epithelial–mesenchymal plasticity. *Life Sci. Alliance* **2**, 1–24 (2019).
123. Ershaid, N. *et al.* NLRP3 inflammasome in fibroblasts links tissue damage with inflammation in breast cancer progression and metastasis. *Nat. Commun.* **10**, 4375 (2019).
124. Guo, Z. *et al.* Cancer-associated fibroblasts induce growth and radioresistance of breast cancer cells through paracrine IL-6. *Cell Death Discov.* **9**, 6 (2023).
125. O’Connell, J. T. *et al.* VEGF-A and Tenascin-C produced by S100A4 + stromal cells are important for metastatic colonization. *Proc. Natl. Acad. Sci.* **108**, 16002–16007 (2011).
126. Orimo, A. *et al.* Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* **121**, 335–348 (2005).
127. Al-Jomah, N., Al-Mohanna, F. H. & Aboussekhra, A. Tocilizumab suppresses the pro-carcinogenic effects of breast cancer-associated fibroblasts through inhibition of the STAT3/AUF1 pathway. *Carcinogenesis* **42**, 1439–1448 (2021).
128. Avalle, L. *et al.* STAT3 induces breast cancer growth via ANGPTL4, MMP13 and STC1 secretion by cancer associated fibroblasts. *Oncogene* **41**, 1456–1467 (2022).
129. Brechbuhl, H. M. *et al.* Fibroblast subtypes define a metastatic matrisome in breast cancer. *{JCI} Insight* **5**, (2020).
130. Bayer, S. V. *et al.* DDR2 controls breast tumor stiffness and metastasis by regulating integrin mediated mechanotransduction in cfs. *Elife* **8**, 1–25 (2019).
131. Dhimolea, E. *et al.* Pleiotropic Mechanisms Drive Endocrine Resistance in the Three-Dimensional Bone Microenvironment. *Cancer Res.* **81**, 371–383 (2021).
132. Gui, Y. *et al.* Metastatic Breast Carcinoma–Associated Fibroblasts Have Enhanced Protumorigenic Properties Related to Increased IGF2 Expression. *Clin. Cancer Res.* (2019) doi:10.1158/1078-0432.ccr-19-1268.
133. Leyh, B., Dittmer, A., Lange, T., Martens, J. W. M. & Dittmer, J. Stromal cells promote anti-estrogen resistance of breast cancer cells through an insulin-like growth factor binding protein 5 (IGFBP5)/B-cell leukemia/lymphoma 3 (Bcl-3) axis. *Oncotarget* **6**, 39307–39328 (2015).
134. Dittmer, A. & Dittmer, J. A CAF-Fueled TIMP-1/CD63/ITGB1/STAT3 Feedback Loop Promotes Migration and Growth of Breast Cancer Cells. *Cancers (Basel)*. **14**, (2022).
135. Gao, Y. *et al.* CD63+ Cancer-Associated Fibroblasts Confer Tamoxifen Resistance to Breast Cancer Cells through Exosomal miR-22. *Adv. Sci.* **7**, 1–17 (2020).
136. Sansone, P. *et al.* Evolution of cancer stem-like cells in endocrine-resistant metastatic breast cancers is mediated by stromal microvesicles. *Cancer Res.*

- 77**, 1927–1941 (2017).
137. Wang, B. *et al.* Cancer-Associated Fibroblasts Promote Radioresistance of Breast Cancer Cells via the HGF/c-Met Signaling Pathway. *Int. J. Radiat. Oncol.* **116**, 640–654 (2022).
 138. Gok Yavuz, B. *et al.* Cancer associated fibroblasts sculpt tumour microenvironment by recruiting monocytes and inducing immunosuppressive PD-1 + TAMs. *Sci. Rep.* **9**, 1–15 (2019).
 139. Jungwirth, U. *et al.* Impairment of a distinct cancer-associated fibroblast population limits tumour growth and metastasis. *Nat. Commun.* **12**, (2021).
 140. Park, C. K., Jung, W. H. & Koo, J. S. Expression of cancer-associated fibroblast-related proteins differs between invasive lobular carcinoma and invasive ductal carcinoma. *Breast Cancer Res. Treat.* **159**, 55–69 (2016).
 141. Forsare, C. *et al.* The Prognostic Role of Intratumoral Stromal Content in Lobular Breast Cancer. *Cancers (Basel)*. **14**, (2022).
 142. Ebrahimsade, S., Westhoff, C. C. & Barth, P. J. CD34+ fibrocytes are preserved in most invasive lobular carcinomas of the breast. *Pathol. Res. Pract.* **203**, 695–698 (2007).
 143. Westhoff, C. C. *et al.* Prognostic relevance of the loss of stromal CD34 positive fibroblasts in invasive lobular carcinoma of the breast. *Virchows Arch.* (2020) doi:10.1007/s00428-020-02835-3.
 144. Manore, S. G., Doheny, D. L., Wong, G. L. & Lo, H. W. IL-6/JAK/STAT3 Signaling in Breast Cancer Metastasis: Biology and Treatment. *Front. Oncol.* **12**, 1–17 (2022).
 145. Ma, J. H., Qin, L. & Li, X. Role of STAT3 signaling pathway in breast cancer. *Cell Commun. Signal.* **18**, 1–13 (2020).
 146. Chung, C. D. *et al.* Specific inhibition of Stat3 signal transduction by PIAS3. *Science (80-.)*. **278**, 1803–1805 (1997).
 147. Avale, L., Camporeale, A., Camperi, A. & Poli, V. STAT3 in cancer: A double edged sword. *Cytokine* **98**, 42–50 (2017).
 148. Haan, S., Keller, J. F., Behrmann, I., Heinrich, P. C. & Haan, C. Multiple reasons for an inefficient STAT1 response upon IL-6-type cytokine stimulation. *Cell. Signal.* **17**, 1542–1550 (2005).
 149. Xu, L. *et al.* IGF1/IGF1R/STAT3 signaling-inducible IFITM2 promotes gastric cancer growth and metastasis. *Cancer Lett.* **393**, 76–85 (2017).
 150. Dong, Z. *et al.* FGF2-induced STAT3 activation regulates pathologic neovascularization. *Exp. Eye Res.* **187**, 107775 (2019).
 151. Balanis, N. *et al.* Epithelial to mesenchymal transition promotes breast cancer progression via a fibronectin-dependent STAT3 signaling pathway. *J. Biol. Chem.* **288**, 17954–17967 (2013).
 152. Bergholz, J. S. *et al.* PI3K β controls immune evasion in PTEN-deficient breast

- tumours. *Nature* (2023) doi:10.1038/s41586-023-05940-w.
153. Tang, L. Y. *et al.* Transforming growth factor- β (TGF- β) directly activates the JAK1-STAT3 axis to induce hepatic fibrosis in coordination with the SMAD pathway. *J. Biol. Chem.* **292**, 4302–4312 (2017).
 154. Bui, Q. T. *et al.* Essential role of Notch4/STAT3 signaling in epithelial–mesenchymal transition of tamoxifen-resistant human breast cancer. *Cancer Lett.* **390**, 115–125 (2017).
 155. Behera, R., Kumar, V., Lohite, K., Karnik, S. & Kundu, G. C. Activation of JAK2/STAT3 signaling by osteopontin promotes tumor growth in human breast cancer cells. *Carcinogenesis* **31**, 192–200 (2010).
 156. Wang, Y. *et al.* p-STAT3 expression in breast cancer correlates negatively with tumor size and HER2 status. *Medicine (Baltimore)*. **100**, e25124 (2021).
 157. Aleskandarany, M. A. *et al.* The prognostic significance of STAT3 in invasive breast cancer: analysis of protein and mRNA expressions in large cohorts. *Breast Cancer Res. Treat.* **156**, 9–20 (2016).
 158. Casneuf, T. *et al.* Interleukin-6 is a potential therapeutic target in interleukin-6 dependent, estrogen receptor- α - Positive breast cancer. *Breast Cancer Targets Ther.* **8**, 13–27 (2016).
 159. Ray, P., Ghosh, S. K., Zhang, D. H. & Ray, A. Repression of interleukin-6 gene expression by 17 β -estradiol: Inhibition of the DNA-binding activity of the transcription factors NF-IL6 and NF- κ B by the estrogen receptor. *FEBS Lett.* **409**, 79–85 (1997).
 160. Chang, Q. *et al.* The IL-6/JAK/Stat3 feed-forward loop drives tumorigenesis and metastasis. *Neoplasia (United States)* **15**, 848–862 (2013).
 161. Jiang, M. *et al.* Interleukin-6 trans-signaling pathway promotes immunosuppressive myeloid-derived suppressor cells via suppression of suppressor of cytokine signaling 3 in breast cancer. *Front. Immunol.* **8**, (2017).
 162. Ortiz-Montero, P., Londoño-Vallejo, A. & Vernot, J. P. Senescence-associated IL-6 and IL-8 cytokines induce a self- and cross-reinforced senescence/inflammatory milieu strengthening tumorigenic capabilities in the MCF-7 breast cancer cell line. *Cell Commun. Signal.* **15**, 1–18 (2017).
 163. Sansone, P. *et al.* Self-renewal of CD133hi cells by IL6/Notch3 signalling regulates endocrine resistance in metastatic breast cancer. *Nat. Commun.* **7**, 1–10 (2016).
 164. Chen, Y. *et al.* STAT3, a poor survival predictor, is associated with lymph node metastasis from breast cancer. *J. Breast Cancer* **16**, 40–49 (2013).
 165. Siersbæk, R. *et al.* IL6/STAT3 Signaling Hijacks Estrogen Receptor α Enhancers to Drive Breast Cancer Metastasis. *Cancer Cell* **38**, 412–423.e9 (2020).
 166. Al-Harbi, B. & Aboussekhra, A. Cucurbitacin I (JSI-124)-dependent inhibition of STAT3 permanently suppresses the pro-carcinogenic effects of active breast

- cancer-associated fibroblasts. *Mol. Carcinog.* **60**, 242–251 (2021).
167. Stevens, L. E. *et al.* JAK-STAT Signaling in Inflammatory Breast Cancer Enables Chemotherapy-Resistant Cell States. *Cancer Res.* **83**, 264–284 (2023).
 168. Jones, L. M. *et al.* STAT3 establishes an immunosuppressive microenvironment during the early stages of breast carcinogenesis to promote tumor growth and metastasis. *Cancer Res.* **76**, 1416–1428 (2016).
 169. Castellaro, A. M., Rodriguez-Baili, M. C., Di Tada, C. E. & Gil, G. A. Tumor-associated macrophages induce endocrine therapy resistance in ER+ breast cancer cells. *Cancers (Basel)*. **11**, 1–29 (2019).
 170. Kutys, M. L. *et al.* Uncovering mutation-specific morphogenic phenotypes and paracrine-mediated vessel dysfunction in a biomimetic vascularized mammary duct platform. *Nat. Commun.* **11**, (2020).
 171. Barbieri, I. *et al.* Stat3 is required for anchorage-independent growth and metastasis but not for mammary tumor development downstream of the ErbB-2 oncogene. *Mol. Carcinog.* **49**, 114–120 (2010).
 172. Zhang, F. *et al.* Elevated STAT3 signaling-mediated upregulation of MMP-2/9 confers enhanced invasion ability in multidrug-resistant breast cancer cells. *Int. J. Mol. Sci.* **16**, 24772–24790 (2015).
 173. Sullivan, N. J. *et al.* Interleukin-6 induces an epithelial-mesenchymal transition phenotype in human breast cancer cells. *Oncogene* **28**, 2940–2947 (2009).
 174. Gao, X. *et al.* PIM1 is responsible for IL-6-induced breast cancer cell EMT and stemness via c-myc activation. *Breast Cancer* **26**, 663–671 (2019).
 175. Kettner, N. M. *et al.* Combined inhibition of STAT3 and DNA repair in palbociclib-resistant er-positive breast cancer. *Clin. Cancer Res.* **25**, 3996–4013 (2019).
 176. Aouad, P. *et al.* Epithelial-mesenchymal plasticity determines estrogen receptor positive breast cancer dormancy and epithelial reconversion drives recurrence. *Nat. Commun.* **13**, (2022).
 177. Zhao, C. *et al.* CAA-derived IL-6 induced M2 macrophage polarization by activating STAT3. *BMC Cancer* **23**, 392 (2023).
 178. Morrow, E. *et al.* High expression of STAT3 within the tumour-associated stroma predicts poor outcome in breast cancer patients. *Cancer Med.* 1–16 (2023) doi:10.1002/cam4.6014.
 179. Hendrayani, S.-F., Al-Khalaf, H. H. & Aboussekhra, A. The Cytokine IL-6 Reactivates Breast Stromal Fibroblasts through Transcription Factor STAT3-dependent Up-regulation of the RNA-binding Protein AUF1. *J. Biol. Chem.* **289**, 30962–30976 (2014).
 180. Orso, F. *et al.* Stroma-derived miR-214 coordinates tumor dissemination. *J. Exp. Clin. Cancer Res.* **42**, 20 (2023).
 181. Sonnenblick, A. *et al.* Constitutive phosphorylated STAT3-associated gene signature is predictive for trastuzumab resistance in primary HER2-positive breast cancer. *BMC Med.* **13**, 1–10 (2015).

182. Xu, H. *et al.* Tumor-derived mesenchymal-stem-cell-secreted IL-6 enhances resistance to cisplatin via the STAT3 pathway in breast cancer. *Oncol. Lett.* **15**, 9142–9150 (2018).
183. Huang, W. C. *et al.* Interleukin-6 expression contributes to lapatinib resistance through maintenance of stemness property in HER2-positive breast cancer cells. *Oncotarget* **7**, 62352–62363 (2016).
184. Marusyk, A. *et al.* Spatial proximity to fibroblasts impacts molecular features and therapeutic sensitivity of breast cancer cells influencing clinical outcomes. *Cancer Res.* **76**, 6495–6506 (2016).
185. Mao, Y. *et al.* Cancer-associated fibroblasts induce trastuzumab resistance in HER2 positive breast cancer cells. *Mol. Biosyst.* **11**, 1029–1040 (2015).
186. Chung, A. W. *et al.* Tocilizumab overcomes chemotherapy resistance in mesenchymal stem-like breast cancer by negating autocrine IL-1A induction of IL-6. *npj Breast Cancer* **8**, (2022).
187. Li, H. *et al.* Drug design targeting protein-protein interactions (PPIs) using multiple ligand simultaneous docking (MLSD) and drug repositioning: Discovery of raloxifene and bazedoxifene as novel inhibitors of IL-6/GP130 interface. *J. Med. Chem.* **57**, 632–641 (2014).
188. Fu, S., Chen, X., Lo, H. W. & Lin, J. Combined bazedoxifene and paclitaxel treatments inhibit cell viability, cell migration, colony formation, and tumor growth and induce apoptosis in breast cancer. *Cancer Lett.* **448**, 11–19 (2019).
189. Schust, J., Sperl, B., Hollis, A., Mayer, T. U. & Berg, T. Stattic: A Small-Molecule Inhibitor of STAT3 Activation and Dimerization. *Chem. Biol.* **13**, 1235–1242 (2006).
190. Poria, D. K., Sheshadri, N., Balamurugan, K., Sharan, S. & Sterneck, E. The STAT3 inhibitor Stattic acts independently of STAT3 to decrease histone acetylation and modulate gene expression. *J. Biol. Chem.* **296**, 100220 (2021).
191. Bharadwaj, U. *et al.* Correction: Small-molecule inhibition of STAT3 in radioresistant head and neck squamous cell carcinoma. *Oncotarget* **10**, 1603 (2019).
192. Rašková, M. *et al.* The Role of IL-6 in Cancer Cell Invasiveness and Metastasis—Overview and Therapeutic Opportunities. *Cells* **11**, (2022).
193. Sonnenblick, A. *et al.* STAT3 activation in HER2-positive breast cancers: Analysis of data from a large prospective trial. *Int. J. Cancer* **148**, 1529–1535 (2021).
194. Sonnenblick, A. *et al.* P-STAT3 in luminal breast cancer: Integrated RNA-protein pooled analysis and results from the BIG 2-98 phase III trial. *Int. J. Oncol.* **52**, 424–432 (2018).
195. Kearney, M. *et al.* Phase I/II trial of ruxolitinib in combination with trastuzumab in metastatic HER2 positive breast cancer. *Breast Cancer Res. Treat.* **189**, 177–185 (2021).

196. Makhlin, I. *et al.* Ruxolitinib and exemestane for estrogen receptor positive, aromatase inhibitor resistant advanced breast cancer. *npj Breast Cancer* **8**, (2022).
197. Howell, S. J. *et al.* Final results of the STEM trial: SFX-01 in the treatment and evaluation of ER+ Her2- metastatic breast cancer (mBC). *Ann. Oncol.* **30**, v122 (2019).
198. Speirs, V. *et al.* Short-term primary culture of epithelial cells derived from human breast tumours. *Br. J. Cancer* **78**, 1421–1429 (1998).
199. Avalle, L. *et al.* Liver-Specific siRNA-Mediated Stat3 or C3 Knockdown Improves the Outcome of Experimental Autoimmune Myocarditis. *Mol. Ther. - Methods Clin. Dev.* **18**, 62–72 (2020).
200. Kaukonen, R., Jacquemet, G., Hamidi, H. & Ivaska, J. Cell-derived matrices for studying cell proliferation and directional migration in a complex 3D microenvironment. *Nat. Protoc.* **12**, 2376–2390 (2017).
201. Ershov, D. *et al.* TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. **19**, (2022).
202. MacLeod, K. *et al.* Predicting Response to Radiotherapy in Cancer-Induced Bone Pain: Cytokines as a Potential Biomarker? *Clin. Oncol.* **32**, e203–e208 (2020).
203. Macleod, K. G., Serrels, B. & Carragher, N. O. Reverse Phase Protein Arrays and Drug Discovery. in *Proteomics for Drug Discovery: Methods and Protocols* (eds. Lazar, I. M., Kontoyianni, M. & Lazar, A. C.) 153–169 (Springer New York, 2017). doi:10.1007/978-1-4939-7201-2_10.
204. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
205. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
206. Binns, D. *et al.* QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics* **25**, 3045–3046 (2009).
207. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. www.pnas.org/cgi/doi/10.1073/pnas.0506580102 (2005).
208. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* **14**, (2013).
209. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
210. Paul Shannon, 1 *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13**, 2498–2504 (2003).
211. Warde-Farley, D. *et al.* The GeneMANIA prediction server: Biological network

- integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **38**, 214–220 (2010).
212. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: A network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, (2010).
213. Hoon, M. J. L. De, Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. **20**, 1453–1454 (2004).
214. Saldanha, A. J. Java Treeview - Extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
215. Aran, D., Sirota, M. & Butte, A. J. Systematic pan-cancer analysis of tumour purity. *Nat. Commun.* **6**, 1–12 (2015).
216. Oh, E. Y. *et al.* Extensive rewiring of epithelial-stromal co-expression networks in breast cancer. *Genome Biol.* **16**, (2015).
217. Adachi, Y. *et al.* Comparison of clinical outcomes between luminal invasive ductal carcinoma and luminal invasive lobular carcinoma. *{BMC} Cancer* **16**, (2016).
218. Brechbuhl, H. M. *et al.* Fibroblast subtypes define a metastatic matrisome in breast cancer. *JCI Insight* **5**, (2020).
219. Kieffer, Y. *et al.* Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discov.* **10**, 1330–1351 (2020).
220. Chen, Y. *et al.* edgeR: differential analysis of sequence read count data User's Guide. *R Packag.* 1–121 (2020).
221. Sherman, B. T. *et al.* DAVID : a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). **50**, 216–221 (2022).
222. Teo, K. *et al.* E-cadherin loss induces targetable autocrine activation of growth factor signalling in lobular breast cancer. *Sci. Rep.* **8**, 1–14 (2018).
223. De Francesco, E. M. *et al.* GPER mediates the angiocrine actions induced by IGF1 through the HIF-1 α /VEGF pathway in the breast tumor microenvironment. *Breast Cancer Res.* **19**, 1–14 (2017).
224. De Francesco, E. M. *et al.* GPER mediates activation of HIF1 α /VEGF signaling by estrogens. *Cancer Res.* **74**, 4053–4064 (2014).
225. De Francesco, E. M. *et al.* HIF-1 α /GPER signaling mediates the expression of VEGF induced by hypoxia in breast cancer associated fibroblasts (CAFs). *Breast Cancer Res.* **15**, 1–18 (2013).
226. Ma, X. J., Dahiya, S., Richardson, E., Erlander, M. & Sgroi, D. C. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res.* **11**, 1–18 (2009).
227. Da, J., Lu, M. & Wang, Z. Estrogen Receptor Alpha (ER α)-Associated Fibroblasts Promote Cell Growth in Prostate Cancer. *Cell Biochem. Biophys.* **73**, 793–798 (2015).

228. Yeh, C. R. *et al.* Estrogen receptor α in cancer associated fibroblasts suppresses prostate cancer invasion via reducing CCL5, IL6 and macrophage infiltration in the tumor microenvironment. *Mol. Cancer* **15**, 1–14 (2016).
229. Kumar, M. M. *et al.* Role of estrogen receptor alpha in human cervical cancer-associated fibroblasts: a transcriptomic study. *Tumor Biol.* **37**, 4409–4420 (2016).
230. Finak, G. *et al.* Stromal gene expression predicts clinical outcome in breast cancer. *Nat. Med.* **14**, 518–527 (2008).
231. Muchlińska, A. *et al.* Alpha-smooth muscle actin-positive cancer-associated fibroblasts secreting osteopontin promote growth of luminal breast cancer. *Cell. Mol. Biol. Lett.* **27**, (2022).
232. Punj, V., Matta, H. & Chaudhary, P. M. X-linked ectodermal dysplasia receptor is downregulated in breast cancer via promoter methylation. *Clin. Cancer Res.* **16**, 1140–1148 (2010).
233. Zhu, C. *et al.* Activation of CECR1 in M2-like TAMs promotes paracrine stimulation-mediated glial tumor progression. doi:10.1093/neuonc/nw251.
234. Zhu, C. *et al.* CECR1-mediated cross talk between macrophages and vascular mural cells promotes neovascularization in malignant glioma. *Nat. Publ. Gr.* **36**, 5356–5368 (2017).
235. Kutryb-Zajac, B. *et al.* Adenosine deaminase inhibition suppresses progression of 4T1 murine breast cancer by adenosine receptor-dependent mechanisms. *J. Cell. Mol. Med.* **22**, 5939–5954 (2018).
236. Kutryb-Zajac, B. *et al.* Macrophage-derived adenosine deaminase 2 correlates with m2 macrophage phenotype in triple negative breast cancer. *Int. J. Mol. Sci.* **22**, 1–19 (2021).
237. Gao, Z. W. *et al.* Distinct Roles of Adenosine Deaminase Isoenzymes ADA1 and ADA2: A Pan-Cancer Analysis. *Front. Immunol.* **13**, 1–12 (2022).
238. Takahashi, H. *et al.* Akt3 is a novel regulator of cancer-associated fibroblasts in head and neck squamous cell carcinoma. *Cancers (Basel)*. **13**, 1–17 (2021).
239. Heits, N. *et al.* Influence of mTOR-inhibitors and mycophenolic acid on human cholangiocellular carcinoma and cancer associated fibroblasts. *BMC Cancer* **16**, 1–12 (2016).
240. Zang, C. *et al.* Targeting multiple tyrosine kinase receptors with dovitinib blocks invasion and the interaction between tumor cells and cancer-associated fibroblasts in breast cancer. *Cell Cycle* **14**, 1291–1299 (2015).
241. Bae, W. J., Kim, S., Ahn, J. M., Han, J. H. & Lee, D. Estrogen-responsive cancer-associated fibroblasts promote invasive property of gastric cancer in a paracrine manner via CD147 production. *FASEB J.* **36**, 1–12 (2022).
242. Narbe, U. *et al.* The estrogen receptor coactivator AIB1 is a new putative prognostic biomarker in ER-positive/HER2-negative invasive lobular carcinoma of the breast. *Breast Cancer Res. Treat.* **175**, 305–316 (2019).

243. Bramwell, V. H. C. *et al.* Serial plasma osteopontin levels have prognostic value in metastatic breast cancer. *Clin. Cancer Res.* **12**, 3337–3343 (2006).
244. McBryan, J., Howlin, J., Napoletano, S. & Martin, F. Amphiregulin: Role in mammary gland development and breast cancer. *J. Mammary Gland Biol. Neoplasia* **13**, 159–169 (2008).
245. Wang, S. F. *et al.* Mitochondrial stress adaptation promotes resistance to aromatase inhibitor in human breast cancer cells via ROS/calcium up-regulated amphiregulin–estrogen receptor loop signaling. *Cancer Lett.* **523**, 82–99 (2021).
246. Wang, X., Masri, S., Phung, S. & Chen, S. The role of amphiregulin in exemestane-resistant breast cancer cells: Evidence of an autocrine loop. *Cancer Res.* **68**, 2259–2265 (2008).
247. Ishimoto, T. *et al.* Activation of Transforming Growth Factor Beta 1 Signaling in Gastric Cancer-associated Fibroblasts Increases Their Motility, via Expression of Rhomboid 5 Homolog 2, and Ability to Induce Invasiveness of Gastric Cancer Cells. *Gastroenterology* **153**, 191-204.e16 (2017).
248. Biffi, G. *et al.* IL1-induced Jak/STAT signaling is antagonized by TGF β to shape CAF heterogeneity in pancreatic ductal adenocarcinoma. *Cancer Discov.* **9**, 282–301 (2019).
249. Arpino, V., Brock, M. & Gill, S. E. The role of TIMPs in regulation of extracellular matrix proteolysis. *Matrix Biol.* **44–46**, 247–254 (2015).
250. Goulet, C. R. *et al.* Cancer-associated fibroblasts induce epithelial-mesenchymal transition of bladder cancer cells through paracrine IL-6 signalling. *BMC Cancer* **19**, 1–13 (2019).
251. Chen, F. *et al.* Single-Cell Transcriptomic Heterogeneity in Invasive Ductal and Lobular Breast Cancer Cells. *Cancer Res.* **81**, 268–281 (2021).
252. Sikora, M. J. *et al.* Invasive lobular carcinoma cell lines are characterized by unique estrogen-mediated gene expression patterns and altered tamoxifen response. *Cancer Res.* **74**, 1463–1474 (2014).
253. Tasdemir, N. *et al.* Comprehensive Phenotypic Characterization of Human Invasive Lobular Carcinoma Cell Lines in 2D and 3D Cultures. *Cancer Res.* **78**, 6209–6222 (2018).
254. Costa, A. *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **33**, 463--479.e10 (2018).
255. Li, A., Chen, P., Leng, Y. & Kang, J. Histone deacetylase 6 regulates the immunosuppressive properties of cancer-associated fibroblasts in breast cancer through the STAT3–COX2-dependent pathway. *Oncogene* **37**, 5952–5966 (2018).
256. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
257. Regis, G., Pensa, S., Boselli, D., Novelli, F. & Poli, V. Ups and downs: The STAT1:STAT3 seesaw of Interferon and gp130 receptor signalling. *Semin. Cell*

- Dev. Biol.* **19**, 351–359 (2008).
258. Lee, S., Rauch, J. & Kolch, W. Targeting MAPK signaling in cancer: Mechanisms of drug resistance and sensitivity. *Int. J. Mol. Sci.* **21**, 1–29 (2020).
 259. Yu, P. N., Cheung, Z. H. & Ip, N. Y. STAT3 as a downstream mediator of Trk signaling and functions. *J. Biol. Chem.* **281**, 15636–15644 (2006).
 260. Tang, L.-Y. *et al.* Transforming Growth Factor- β (TGF- β) Directly Activates the JAK1-STAT3 Axis to Induce Hepatic Fibrosis in Coordination with the SMAD Pathway. *J. Biol. Chem.* **292**, 4302–4312 (2017).
 261. Rawlings, J. S., Rosler, K. M. & Harrison, D. A. The JAK/STAT signaling pathway. *J. Cell Sci.* **117**, 1281–1283 (2004).
 262. Pickert, G. *et al.* STAT3 links IL-22 signaling in intestinal epithelial cells to mucosal wound healing. *J. Exp. Med.* **206**, 1465–1472 (2009).
 263. Dudka, A. A., Sweet, S. M. M. & Heath, J. K. Signal Transducers and Activators of Transcription-3 Binding to the Fibroblast Growth Factor Receptor Is Activated by Receptor Amplification. *Cancer Res.* **70**, 3391–3401 (2010).
 264. Johnson, D. E., O’Keefe, R. A. & Grandis, J. R. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat. Rev. Clin. Oncol.* **15**, 234–248 (2018).
 265. Calon, A. *et al.* Dependency of Colorectal Cancer on a TGF- β -Driven Program in Stromal Cells for Metastasis Initiation. *Cancer Cell* **22**, 571–584 (2012).
 266. Turner, N. *et al.* FGFR1 Amplification Drives Endocrine Therapy Resistance and Is a Therapeutic Target in Breast Cancer. *Cancer Res.* **70**, 2085–2094 (2010).
 267. Aran, D., Hu, Z. & Butte, A. J. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **18**, 1–14 (2017).
 268. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545 LP – 15550 (2005).
 269. Kucera, M., Isserlin, R., Arkhangorodsky, A. & Bader, G. D. AutoAnnotate: A Cytoscape app for summarizing networks with semantic annotations [version 1; referees: 2 approved]. *F1000Research* **5**, 1–12 (2016).
 270. Junttila, I. S. Tuning the cytokine responses: An update on interleukin (IL)-4 and IL-13 receptor complexes. *Front. Immunol.* **9**, (2018).
 271. Rose-John, S., Winthrop, K. & Calabrese, L. The role of IL-6 in host defence against infections: Immunobiology and clinical implications. *Nat. Rev. Rheumatol.* **13**, 399–409 (2017).
 272. Hirano, T. IL-6 in inflammation, autoimmunity and cancer. *Int. Immunol.* **33**, 127–148 (2021).
 273. Tanaka, T., Narazaki, M. & Kishimoto, T. IL-6 in Inflammation, Immunity, and Disease. **6**, 1–16 (2014).
 274. Dritsoula, A. *et al.* Angiopathic activity of LRG1 is induced by the IL-6/STAT3 pathway. *Sci. Rep.* **12**, 1–14 (2022).

275. He, L. *et al.* LRG1 mediated by ATF3 promotes growth and angiogenesis of gastric cancer by regulating the SRC/STAT3/VEGFA pathway. *Gastric Cancer* **25**, 527–541 (2022).
276. Cai, D. *et al.* LRG1 in pancreatic cancer cells promotes inflammatory factor synthesis and the angiogenesis of HUVECs by activating VEGFR signaling. *J. Gastrointest. Oncol.* **13**, 400–412 (2022).
277. Wang, X. *et al.* LRG1 promotes angiogenesis by modulating endothelial TGF- β signalling. *Nature* **499**, 306–311 (2013).
278. Jemmerson, R., Staskus, K., Higgins, L., Conklin, K. & Kelekar, A. Intracellular leucine - rich alpha - 2 - glycoprotein - 1 competes with Apaf - 1 for binding cytochrome c in protecting MCF - 7 breast cancer cells from apoptosis. *Apoptosis* **26**, 71–82 (2021).
279. Zou, Z., Ohta, T., Miura, F. & Oki, S. ChIP-Atlas 2021 update: a data-mining suite for exploring epigenomic landscapes by fully integrating ChIP-seq, ATAC-seq and Bisulfite-seq data. *Nucleic Acids Res.* **50**, 175–182 (2022).
280. Parri, E., Kuusanmäki, H., van Adrichem, A. J., Kaustio, M. & Wennerberg, K. Identification of novel regulators of STAT3 activity. *PLoS One* **15**, 1–18 (2020).
281. Camilli, C., Hoeh, A. E., De Rossi, G., Moss, S. E. & Greenwood, J. LRG1: an emerging player in disease pathogenesis. *J. Biomed. Sci.* **29**, 1–29 (2022).
282. Zhong, B. *et al.* Colorectal cancer-associated fibroblasts promote metastasis by up-regulating LRG1 through stromal IL-6/STAT3 signaling. *Cell Death Dis.* **13**, (2022).
283. Yu, B. *et al.* LRG1 facilitates corneal fibrotic response by inducing neutrophil chemotaxis via Stat3 signaling in alkali-burned mouse corneas. *Am. J. Physiol. - Cell Physiol.* **321**, C415–C428 (2021).
284. Zhu, H. *et al.* AKR1C1 activates STAT3 to promote the metastasis of non-small cell lung cancer. *Theranostics* **8**, 676–692 (2018).
285. Muoio, M. G. *et al.* Activation of the s100a7/rage pathway by igf-1 contributes to angiogenesis in breast cancer. *Cancers (Basel)*. **13**, 1–19 (2021).
286. Vecchi, L. *et al.* Interleukin-6 Signaling in Triple Negative Breast Cancer Cells Elicits the Annexin A1/Formyl Peptide Receptor 1 Axis and Affects the Tumor Microenvironment. *Cells* **11**, (2022).
287. Guo, C. *et al.* Liposomal nanoparticles carrying anti-IL6R antibody to the tumour microenvironment inhibit metastasis in two molecular subtypes of breast cancer mouse models. *Theranostics* **7**, 775–788 (2017).
288. Rodriguez-Barrueco, R. *et al.* Inhibition of the autocrine IL-6–JAK2–STAT3–calprotectin axis as targeted therapy for HR–/HER2+ breast cancers. *Genes Dev.* **29**, 1631–1648 (2015).
289. Walia, V. *et al.* Loss of breast epithelial marker hCLCA2 promotes epithelial-to-mesenchymal transition and indicates higher risk of metastasis. *Oncogene* **31**, 2237–2246 (2012).

290. Walia, V. *et al.* hCLCA2 is a p53-inducible inhibitor of breast cancer cell proliferation. *Cancer Res.* **69**, 6624–6632 (2009).
291. Ramena, G., Yin, Y., Yu, Y., Walia, V. & Elble, R. C. CLCA2 interactor EVA1 is required for mammary epithelial cell differentiation. *PLoS One* **11**, 1–21 (2016).
292. Goulet, B. *et al.* Nuclear localization of maspin is essential for its inhibition of tumor growth and metastasis. *Lab. Investig.* **91**, 1181–1187 (2011).
293. Machowska, M., Wachowicz, K., Sopel, M. & Rzepecki, R. Nuclear location of tumor suppressor protein maspin inhibits proliferation of breast cancer cells without affecting proliferation of normal epithelial cells. *BMC Cancer* **14**, (2014).
294. Sakabe, T., Wakahara, M., Shiota, G. & Umekita, Y. Role of cytoplasmic localization of maspin in promoting cell invasion in breast cancer with aggressive phenotype. *Sci. Rep.* **11**, 1–12 (2021).
295. Longhi, M. T. *et al.* PI3K-AKT, JAK2-STAT3 pathways and cell–cell contact regulate maspin subcellular localization. *Cell Commun. Signal.* **19**, 1–16 (2021).
296. Qu, J. *et al.* AKR1B10 promotes breast cancer cell proliferation and migration via the PI3K/AKT/NF- κ B signaling pathway. *Cell Biosci.* **11**, 1–13 (2021).
297. Marino, N. *et al.* FAM83A is a potential biomarker for breast cancer initiation. *Biomark. Res.* **10**, 1–14 (2022).
298. Valque, H., Gouyer, V., Gottrand, F. & Desseyn, J. L. MUC5B Leads to Aggressive Behavior of Breast Cancer MCF7 Cells. *PLoS One* **7**, (2012).
299. Conley, S. J. *et al.* HER2 drives Mucin-like 1 to control proliferation in breast cancer cells. *Oncogene* **35**, 4225–4234 (2016).
300. Mendoza-Villanueva, D. *et al.* The C/EBP δ protein is stabilized by estrogen receptor α activity, inhibits SNAI2 expression and associates with good prognosis in breast cancer. *Oncogene* **35**, 6166–6176 (2016).
301. Ou Yang, T.-H., Cheng, W., Zheng, T., Maurer, M. A. & Anastassiou, D. Breast Cancer Prognostic Biomarker Using Attractor Metagenes and the FGD3–SUSD3 Metagene. *Cancer Epidemiol. Biomarkers Prev.* **23**, 2850–2856 (2014).
302. Willis, S. *et al.* High Expression of FGD3, a Putative Regulator of Cell Morphology and Motility, Is Prognostic of Favorable Outcome in Multiple Cancers. *JCO Precis. Oncol.* 1–13 (2017) doi:10.1200/po.17.00009.
303. Susini, T. *et al.* Immunohistochemical evaluation of fgd3 expression: A new strong prognostic factor in invasive breast cancer. *Cancers (Basel)*. **13**, (2021).
304. Moy, I. *et al.* Estrogen-dependent sushi domain containing 3 regulates cytoskeleton organization and migration in breast cancer cells. *Oncogene* **34**, 362–372 (2015).
305. Balamurugan, K. *et al.* The tumour suppressor C/EBP θ inhibits FBXW7 expression and promotes mammary tumour metastasis. *EMBO J.* **29**, 4106–4117 (2010).
306. Balamurugan, K. *et al.* C/EBP δ links IL-6 and HIF-1 signaling to promote breast

- cancer stem cell-associated phenotypes. *Oncogene* **38**, 3765–3780 (2019).
307. Du, T. *et al.* Invasive lobular and ductal breast carcinoma differ in immune response, protein translation efficiency and metabolism. *Sci. Rep.* **8**, (2018).
308. Demaria, M. *et al.* A STAT3-mediated metabolic switch is involved in tumour transformation and STAT3 addiction. *Aging (Albany, NY)*. **2**, 823–842 (2010).
309. Wang, T. *et al.* JAK/STAT3-Regulated Fatty Acid β -Oxidation Is Critical for Breast Cancer Stem Cell Self-Renewal and Chemoresistance. *Cell Metab.* **27**, 136-150.e5 (2018).
310. Oshi, M. *et al.* The E2F Pathway Score as a Predictive Biomarker of Response to Neoadjuvant Therapy in ER+/HER2- Breast Cancer. *Cells* **9**, (2020).
311. Luo, K. *et al.* Co-delivery of paclitaxel and STAT3 siRNA by a multifunctional nanocomplex for targeted treatment of metastatic breast cancer. *Acta Biomater.* **134**, 649–663 (2021).
312. Wang, S., Yao, Y., Yao, M., Fu, P. & Wang, W. Interleukin-22 promotes triple negative breast cancer cells migration and paclitaxel resistance through JAK-STAT3/MAPKs/AKT signaling pathways. *Biochem. Biophys. Res. Commun.* **503**, 1605–1609 (2018).
313. Wang, L. *et al.* CAFs enhance paclitaxel resistance by inducing EMT through the IL-6/JAK2/STAT3 pathway. *Oncol. Rep.* **39**, 2081–2090 (2018).
314. Rodriguez, M. J. *et al.* Targeting mTOR to overcome resistance to hormone and CDK4/6 inhibitors in ER-positive breast cancer models. *Sci. Rep.* **13**, 2710 (2023).
315. De Angelis, C. *et al.* Activation of the IFN signaling pathway is associated with resistance to CDK4/6 inhibitors and immune checkpoint activation in ER-positive breast cancer. *Clin. Cancer Res.* **27**, 4870–4882 (2021).
316. Xu, D., Zhang, Y. & Jin, F. The role of AKR1 family in tamoxifen resistant invasive lobular breast cancer based on data mining. *BMC Cancer* **21**, 1–14 (2021).
317. Czapiewski, P. *et al.* BCL3 expression is strongly associated with the occurrence of breast cancer relapse under tamoxifen treatment in a retrospective cohort study. *Virchows Arch.* (2022) doi:10.1007/s00428-021-03238-8.
318. Wang, L. *et al.* Cancer-associated fibroblasts enhance metastatic potential of lung cancer cells through IL-6/STAT3 signaling pathway. *Oncotarget* **8**, 76116–76128 (2017).
319. Kwan, Y. P. *et al.* Lrg1 promotes metastatic dissemination of melanoma through regulating egfr/stat3 signalling. *Cancers (Basel)*. **13**, 1–16 (2021).
320. van Weverwijk, A. *et al.* Metabolic adaptability in metastatic breast cancer by AKR1B10-dependent balancing of glycolysis and fatty acid oxidation. *Nat. Commun.* **10**, (2019).
321. O'Brien, N. *et al.* Use of a panel of novel genes for differentiating breast cancer from non-breast tissues. *Tumor Biol.* **28**, 312–317 (2007).

322. García, E. P. *et al.* MUC5B silencing reduces chemo-resistance of MCF-7 breast tumor cells and impairs maturation of dendritic cells. *Int. J. Oncol.* **48**, 2113–2123 (2016).
323. Li, Q. H. *et al.* Small breast epithelial mucin promotes the invasion and metastasis of breast cancer cells via promoting epithelial-to-mesenchymal transition. *Oncol. Rep.* **44**, 509–518 (2020).
324. Abdulla, M. *et al.* Targeting MUCL1 protein inhibits cell proliferation and EMT by deregulating β -catenin and increases irinotecan sensitivity in colorectal cancer. *Int. J. Oncol.* **60**, 1–12 (2022).
325. Zhang, Y., Lun, X. & Guo, W. Expression of TRPC1 and SBEM protein in breast cancer tissue and its relationship with clinicopathological features and prognosis of patients. *Oncol. Lett.* **20**, 1–1 (2020).
326. Liu, L. *et al.* Small breast epithelial mucin tumor tissue expression is associated with increased risk of recurrence and death in triple-negative breast cancer patients. *Diagn. Pathol.* **8**, 1–10 (2013).
327. Jia, L. *et al.* A 10-Gene Signature for Predicting the Response to Neoadjuvant Trastuzumab Therapy in HER2-Positive Breast Cancer. *Clin. Breast Cancer* **21**, e654–e664 (2021).
328. Ma, R. Y. *et al.* Monocyte-derived macrophages promote breast cancer bone metastasis outgrowth. *J. Exp. Med.* **217**, (2020).
329. Venmar, K. T., Carter, K. J., Hwang, D. G., Dozier, E. A. & Fingleton, B. IL4 receptor ILR4 α regulates metastatic colonization by mammary tumors through multiple signaling pathways. *Cancer Res.* **74**, 4329–4340 (2014).
330. Parveen, S. *et al.* Therapeutic targeting with DABIL-4 depletes myeloid suppressor cells in 4T1 triple-negative breast cancer model. *Mol. Oncol.* **15**, 1330–1344 (2021).
331. Liu, T. *et al.* Tumor-associated macrophages in human breast cancer produce new monocyte attracting and pro-angiogenic factor YKL-39 indicative for increased metastasis after neoadjuvant chemotherapy. *Oncoimmunology* **7**, 1–17 (2018).
332. Liu, L. *et al.* CHI3L2 Is a Novel Prognostic Biomarker and Correlated With Immune Infiltrates in Gliomas. *Front. Oncol.* **11**, 1–13 (2021).
333. Yin, C. *et al.* RAGE-binding S100A8/A9 promotes the migration and invasion of human breast cancer cells through actin polymerization and epithelial-mesenchymal transition. *Breast Cancer Res. Treat.* **142**, 297–309 (2013).
334. Liu, H. *et al.* S100A7 enhances invasion of human breast cancer MDA-MB-468 cells through activation of nuclear factor- κ B signaling. *World J. Surg. Oncol.* **11**, 1–8 (2013).
335. Nasser, M. W. *et al.* RAGE mediates S100A7-induced breast cancer growth and metastasis by modulating the tumor microenvironment. *Cancer Res.* **75**, 974–985 (2015).

336. Rigracciolo, D. C. *et al.* Focal Adhesion Kinase (FAK)-Hippo/YAP transduction signaling mediates the stimulatory effects exerted by S100A8/A9-RAGE system in triple-negative breast cancer (TNBC). *J. Exp. Clin. Cancer Res.* **41**, 1–25 (2022).
337. Zhang, S. *et al.* Distinct prognostic values of S100 mRNA expression in breast cancer. *Nat. Publ. Gr.* 1–11 (2017) doi:10.1038/srep39786.
338. Miller, P. *et al.* Elevated S100A8 protein expression in breast cancer cells and breast tumor stroma is prognostic of poor disease outcome. *Breast Cancer Res. Treat.* **166**, 85–94 (2017).
339. Allaoui, R. *et al.* Cancer-associated fibroblast-secreted CXCL16 attracts monocytes to promote stroma activation in triple-negative breast cancers. *Nat. Commun.* **7**, (2016).
340. Xu, X. *et al.* S100A9 promotes human lung fibroblast cells activation through receptor for advanced glycation end-product-mediated extracellular-regulated kinase 1/2, mitogen-activated protein-kinase and nuclear factor- κ B-dependent pathways. *Clin. Exp. Immunol.* **173**, 523–535 (2013).
341. Wang, S. *et al.* S100A8/A9 in inflammation. *Front. Immunol.* **9**, (2018).
342. Allgöwer, C. *et al.* Friend or foe: S100 proteins in cancer. *Cancers* vol. 12 (2020).
343. Nasser, M. W. *et al.* S100A7 enhances mammary tumorigenesis through upregulation of inflammatory pathways. *Cancer Res.* **72**, 604–615 (2012).
344. Desmedt, C. *et al.* Immune Infiltration in Invasive Lobular Breast Cancer. *{JNCI} J. Natl. Cancer Inst.* **110**, 768–776 (2018).
345. Bertucci, F. *et al.* Lobular and ductal carcinomas of the breast have distinct genomic and expression profiles. *Oncogene* **27**, 5359–5372 (2008).
346. Li, Z. *et al.* Exosomal Leucine-Rich-Alpha2-Glycoprotein 1 Derived from Non-Small-Cell Lung Cancer Cells Promotes Angiogenesis via TGF- β Signal Pathway. *Mol. Ther. - Oncolytics* **14**, 313–322 (2019).
347. Sreekumar, S. *et al.* Differential Regulation and Targeting of Estrogen Receptor α Turnover in Invasive Lobular Breast Carcinoma. *Endocrinol. (United States)* **161**, 1–14 (2020).
348. Wang, H. *et al.* Tumor immunological phenotype signature-based high-throughput screening for the discovery of combination immunotherapy compounds. *Sci. Adv.* **7**, 1–15 (2021).
349. Shackelford, M. T. *et al.* Estrogen Regulation of mTOR Signaling and Mitochondrial Function in Invasive Lobular Carcinoma Cell Lines Requires WNT4. *Cancers (Basel)*. **12**, 2931 (2020).
350. Bossart, E. A. *et al.* SNAIL is induced by tamoxifen and leads to growth inhibition in invasive lobular breast carcinoma. *Breast Cancer Res. Treat.* **175**, 327–337 (2019).
351. Ginestier, C. *et al.* Article ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome. **1**, 555–567

- (2007).
352. Sarmiento-Castro, A. *et al.* Increased Expression of Interleukin-1 Receptor Characterizes Anti-estrogen-Resistant ALDH⁺ Breast Cancer Stem Cells. *Stem Cell Reports* **15**, 307–316 (2020).
 353. Lin, Q. *et al.* CYP1B1-catalyzed 4-OHE2 promotes the castration resistance of prostate cancer stem cells by estrogen receptor α -mediated IL6 activation. *Cell Commun. Signal.* **20**, 1–16 (2022).
 354. Hanker, A. B., Sudhan, D. R. & Arteaga, C. L. Overcoming Endocrine Resistance in Breast Cancer. *Cancer Cell* **37**, 496–513 (2020).
 355. Tsoi, H., Man, E. P. S., Chau, K. M. & Khoo, U. S. Targeting the IL-6/STAT3 signalling cascade to reverse tamoxifen resistance in estrogen receptor positive breast cancer. *Cancers (Basel)*. **13**, (2021).
 356. Litzenburger, U. M. *et al.* Constitutive IDO expression in human cancer is sustained by an autocrine signaling loop involving IL-6, STAT3 and the AHR. *Oncotarget* **5**, 1038–1051 (2014).
 357. Belous, A. R., Hachey, D. L., Dawling, S., Roodi, N. & Parl, F. F. Cytochrome P450 1B1 – Mediated Estrogen Metabolism Results in Estrogen-Deoxyribonucleoside Adduct Formation. 812–817 (2007) doi:10.1158/0008-5472.CAN-06-2133.
 358. Balamurugan, K. & Sterneck, E. The Many Faces of C / EBP δ and their Relevance for Inflammation and Cancer. **9**, (2013).
 359. Palmieri, C. *et al.* Site-specific CpG methylation in the CCAAT / enhancer binding protein delta (CEBP d) CpG island in breast cancer is associated with metastatic relapse. 732–738 (2012) doi:10.1038/bjc.2012.308.
 360. Mouabbi, J. A. *et al.* Histology-based survival outcomes in hormone receptor-positive metastatic breast cancer treated with targeted therapies. *npj Breast Cancer* **8**, 1–7 (2022).
 361. Gao, J. J. *et al.* Overall survival in patients with hormone receptor-positive, HER2-negative, advanced or metastatic breast cancer treated with a cyclin-dependent kinase 4/6 inhibitor plus fulvestrant: a US Food and Drug Administration pooled analysis. *Lancet Oncol.* **22**, 1573–1581 (2021).
 362. Liu, C., Jiang, Y. & Han, B. Mir-613 suppresses chemoresistance and stemness in triple-negative breast cancer by targeting fam83a. *Cancer Manag. Res.* **12**, 12623–12633 (2020).
 363. Nisticò, C. *et al.* Lipid droplet biosynthesis impairment through dgat2 inhibition sensitizes mcf7 breast cancer cells to radiation. *Int. J. Mol. Sci.* **22**, 1–18 (2021).
 364. Simões, B. M. *et al.* Anti-estrogen Resistance in Human Breast Tumors Is Driven by JAG1-NOTCH4-Dependent Cancer Stem Cell Activity. *Cell Rep.* **12**, 1968–1977 (2015).
 365. Dinakar, Y. H. *et al.* Role of STAT3 in the initiation, progression, proliferation and metastasis of breast cancer and strategies to deliver JAK and STAT3

- inhibitors. *Life Sci.* **309**, 120996 (2022).
366. Turnham, D. J. *et al.* Bcl-3 promotes multi-modal tumour cell migration via NF- κ B1 mediated regulation of Cdc42. *Carcinogenesis* **41**, 1432–1443 (2020).
 367. Wakefield, A. *et al.* Bcl3 selectively promotes metastasis of ERBB2-driven mammary tumors. *Cancer Res.* **73**, 745–755 (2013).
 368. Koorman, T. *et al.* Spatial collagen stiffening promotes collective breast cancer cell invasion by reinforcing extracellular matrix alignment. *Oncogene* **41**, 2458–2469 (2022).
 369. Emad, A. *et al.* Superior breast cancer metastasis risk stratification using an epithelial-mesenchymal-amoeoid transition gene signature. *Breast Cancer Res.* **22**, 1–13 (2020).
 370. Taube, J. H. *et al.* Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15449–15454 (2010).
 371. Christgen, M. & Derksen, P. W. B. Lobular breast cancer: molecular basis, mouse and cellular models. *Breast Cancer Res.* **17**, (2015).
 372. Davis, S. L. *et al.* A Phase I Dose-Escalation Study of Linsitinib (OSI-906), a Small-Molecule Dual Insulin-Like Growth Factor-1 Receptor/Insulin Receptor Kinase Inhibitor, in Combination with Irinotecan in Patients with Advanced Cancer. *Oncologist* **23**, 1409-e140 (2018).
 373. Di Cosimo, S. *et al.* Combination of the mTOR inhibitor ridaforolimus and the anti-IGF1R monoclonal antibody dalotuzumab: Preclinical characterization and phase I clinical trial. *Clin. Cancer Res.* **21**, 49–59 (2015).
 374. Padmanaban, V. *et al.* E-cadherin is required for metastasis in multiple models of breast cancer. *Nature* **324**, 297–314 (2019).
 375. Quiros, R. M. *et al.* Ovarian normal and tumor-associated fibroblasts retain in vivo stromal characteristics in a 3-D matrix-dependent manner. *Gynecol. Oncol.* **110**, 99–109 (2008).
 376. Chen, Y. *et al.* SNF5, a core subunit of SWI/SNF complex, regulates melanoma cancer cell growth, metastasis, and immune escape in response to matrix stiffness. *Transl. Oncol.* **17**, 101335 (2022).
 377. Laklai, H. *et al.* Genotype tunes pancreatic ductal adenocarcinoma tissue tension to induce matricellular fibrosis and tumor progression. *Nat. Med.* **22**, 497–505 (2016).
 378. McCart Reed, A. E., Kalinowski, L., Simpson, P. T. & Lakhani, S. R. Invasive lobular carcinoma of the breast: the increasing importance of this special subtype. *Breast Cancer Research* vol. 23 1–16 at <https://doi.org/10.1186/s13058-020-01384-6> (2021).
 379. Drabsch, Y., He, S., Zhang, L., Snaar-Jagalska, B. E. & ten Dijke, P. Transforming growth factor- β signalling controls human breast cancer metastasis in a zebrafish xenograft model. *Breast Cancer Res.* **15**, (2013).

380. Nguyen-Chi, M. *et al.* Identification of polarized macrophage subsets in zebrafish. *Elife* **4**, 1–14 (2015).
381. Varela, M., Dios, S., Novoa, B. & Figueras, A. Characterisation, expression and ontogeny of interleukin-6 and its receptors in zebrafish (*Danio rerio*). *Dev. Comp. Immunol.* **37**, 97–106 (2012).
382. Pestalozzi, B. C. *et al.* Distinct Clinical and Prognostic Features of Infiltrating Lobular Carcinoma of the Breast: Combined Results of 15 International Breast Cancer Study Group Clinical Trials. *J. Clin. Oncol.* **26**, 3006–3014 (2008).
383. Evgen Pharma. SFX-01 in the Treatment and Evaluation of Metastatic Breast Cancer - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT02970682?term=sfx&cond=Breast+Cancer&draw=2&rank=1> (2019).
384. Ahmed, S. *et al.* IL-8 secreted by tumor associated macrophages contribute to lapatinib resistance in HER2-positive locally advanced breast cancer via activation of Src/STAT3/ERK1/2-mediated EGFR signaling. *Biochim. Biophys. Acta - Mol. Cell Res.* **1868**, 118995 (2021).
385. Vasiyani, H. *et al.* DNA damage induces STING mediated IL-6-STAT3 survival pathway in triple-negative breast cancer cells and decreased survival of breast cancer patients. *Apoptosis* **27**, 961–978 (2022).
386. Yang, C. *et al.* Increased drug resistance in breast cancer by tumor-associated macrophages through IL-10/STAT3/bcl-2 signaling pathway. *Med. Oncol.* **32**, (2015).
387. Brocke-Heidrich, K. *et al.* BCL3 is induced by IL-6 via Stat3 binding to intronic enhancer HS4 and represses its own transcription. *Oncogene* **25**, 7297–7304 (2006).
388. Alves, C. L. *et al.* Co-targeting CDK4/6 and AKT with endocrine therapy prevents progression in CDK4/6 inhibitor and endocrine therapy-resistant breast cancer. *Nat. Commun.* **12**, 1–15 (2021).
389. Duda, D. G. *et al.* Differential transplantability of tumor-associated stromal cells. *Cancer Res.* **64**, 5920–5924 (2004).
390. Webb, E. R. *et al.* Kindlin-1 regulates IL-6 secretion and modulates the immune environment in breast cancer models. *Elife* **12**, 1–21 (2023).
391. Lam, S. H., Chua, H. L., Gong, Z., Lam, T. J. & Sin, Y. M. Development and maturation of the immune system in zebrafish, *Danio rerio*: A gene expression profiling, in situ hybridization and immunological study. *Dev. Comp. Immunol.* **28**, 9–28 (2004).
392. Póvoa, V. *et al.* Innate immune evasion revealed in a colorectal zebrafish xenograft model. *Nat. Commun.* **12**, 1–15 (2021).
393. Tuit, S. *et al.* Transcriptional Signature Derived from Murine Tumor-Associated Macrophages Correlates with Poor Outcome in Breast Cancer Patients. *Cell Rep.* **29**, 1221-1235.e5 (2019).

394. Yang, L., Guo, P., Wang, P., Wang, W. & Liu, J. IL-6/ERK signaling pathway participates in type I IFN-programmed, unconventional M2-like macrophage polarization. *Sci. Rep.* **13**, 1–14 (2023).
395. Lim, S. Y., Yuzhalin, A. E., Gordon-Weeks, A. N. & Muschel, R. J. Tumor-infiltrating monocytes/macrophages promote tumor invasion and migration by upregulating S100A8 and S100A9 expression in cancer cells. *Oncogene* **35**, 5735–5745 (2016).

Appendix 1 Image analysis macros

```

///first duplicate channel and slice of interest and split channels
///make mask of nuclei
// make sure clicked on nuclei image before running
// make sure changed redirect in line 22 to appropriate image name
run("Smooth");
run("Unsharp Mask...", "radius=50 mask=0.90");
run("Smooth");
setAutoThreshold("Otsu dark");
setThreshold(300, 65535, "raw");
run("Make Binary");
run("Open");
run("Outline");
run("Smooth");
run("Make Binary");
run("Fill Holes");
run("Watershed");
run("Erode");
run("Gaussian Blur...", "sigma=2");
run("Convert to Mask");

/// redirect to other image --> edit for each image
run("Set Measurements...", "area mean shape integrated area_fraction limit display redirect=C2-SUM44PE-IL6CCM-IL61-STAT3-3-1.oir decimal=3");

run("Analyze Particles...", "size=25-Infinity display summarize");

```

Figure Ap 1-1. Nuclear intensity of STAT3 and ER α macro. For each image, the channels were first split then the DAPI (nuclear) channel image was selected and the name of the image in line 22 was changed to the STAT3/ER α channel image. The macro smooths the background of the nuclear channel, then thresholds the nuclei, fills in any holes and generates a mask of the nuclei. This mask is then applied to the STAT3/ER α channel and the mean gray value of STAT3/ER α in the nucleus of each cell is measured.

```

1
2 run("Duplicate...", " ")
3 run("Smooth");
4 run("Unsharp Mask...", "radius=50 mask=0.90");
5 run("Smooth");
6 setAutoThreshold("Mean dark");
7 //run("Threshold...");
8 setAutoThreshold("Mean dark");
9 run("Make Binary");
10 run("Open");
11 run("Outline");
12 run("Smooth");
13 run("Make Binary");
14 run("Fill Holes");
15 run("Watershed");
16 run("Erode");
17 run("Gaussian Blur...", "sigma=2");
18 run("Convert to Mask");
19 // exclude particles smaller than 25 um^2
20 run("Analyze Particles...", "size=25-Infinity show=Outlines display summarize add stack");|

```

Figure Ap 1-2. EdU+ cell quantification macro. This macro was applied to the DAPI (nuclei) and EdU-TAMRA channels to count the total number of nuclei and number of EdU+ nuclei. Written by Dr Amy Davies.

```

1  makeRectangle(2, 3, 1052, 689);
2  run("Crop");
3  run("Gaussian Blur...", "sigma=2");
4  setAutoThreshold("Otsu dark");
5  //run("Threshold...");
6  //setThreshold(115, 255);
7  setOption("BlackBackground", true);
8  run("Convert to Mask");
9
10 run("Convert to Mask");
11 run("Analyze Particles...", "size=50.00-Infinity show=Outlines display exclude include summarize overlay");
12
13

```

Figure Ap 1-3. Cell shape analysis macro. This macro was applied to brightfield images of SUM44PE cells adhered to CAF-derived matrices taken on the Incucyte S3 at 10X. Images were first cropped to remove borders then the background was blurred and images were thresholded to leave the outline of cells. These outlines were then converted to masks and the aspect ratio for each individual cell in each image and the mean aspect ratio of all cells larger than $50 \mu\text{m}^2$ in an image was measured.

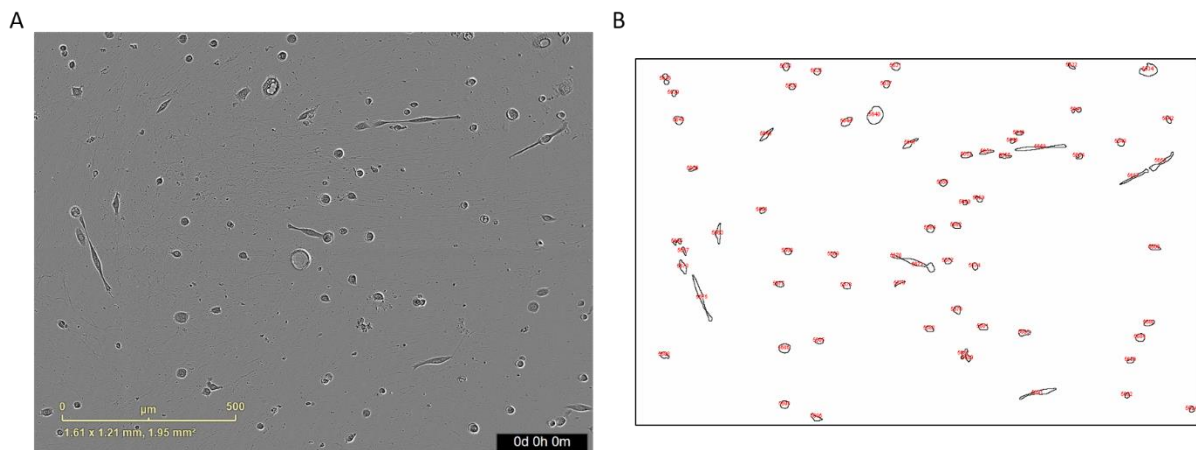


Figure Ap 1-4. Representative images of SUM44PE cell shape analysis on CDMs. (A) Brightfield image of SUM44PE cells stimulated with 10 ng/ml IL-6 adhered to ED26 CDMs for 72 hours and (B) cell outlines generated and measured by the macro in Figure Ap 1-3.

Appendix 2 FPPA and RPPA antibody lists

Appendix 2.1 FPPA antibodies

Target	Catalogue No.
Amphiregulin	MAB262-100
CCL1/I-309	MAB272-100
CCL11/Eotaxin	MAB320-100
CCL17/TARC	MAB364-100
CCL2/JE/MCP-1	MAB679-100
CCL21/6Ckine	MAB3661
CCL22/MDC	MAB336-100
CCL27/CTACK	MAB3761
CCL28	MAB7171
CCL3/MIP-1 alpha	AF-270-NA
CCL5/RANTES	MAB678-100
CX3CL1/Fractalkine	MAB3652-100
CXCL10/IP-10/CRG-2	MAB266
CXCL11/I-TAC	MAB672-100
CXCL12/SDF-1	MAB350-100
CXCL13/BLC/BCA-1	MAB801-100
CXCL16	AF976
CXCL4/PF4	MAB7591-100
CxCL7 / NAP-2	MAB393
CXCL9/MIG	MAB392-100
DPPIV/CD26	MAB1180-100
EGF	MAB636-100
Endoglin/CD105	MAB10972-100
Gas1	MAB2636
HGF	MAB694-100
GM-CSF	MAB615-100
IFN-gamma	MAB28252-100
IGF-I	MAB291-100
IGFBP-2	MAB6741-100
IGFBP-3	MAB305-100
IGFBP-5	MAB8751
IGFBP-6	MAB8761
IL-1 alpha / IL-1F1	MAB200-100

IL-1 beta IL-1F2	MAB601-100
IL-10	MAB2172-100
IL-11	MAB618-100
IL-12 p70	MAB611-100
IL-13	MAB213-100
IL-17 / IL-17A	MAB317-100
IL-1ra/ IL-1F3	MAB280
IL-2	MAB602-100
IL-23 (duoet)	842896
IL-3	MAB603
IL-4	MAB604
IL-5	MAB405-100
IL-6	MAB206-100
IL-7	MAB207
Leptin	MAB398-100
M-CSF	MAB616-100
MMP-2 duoet	843027
MMP-3	AF513
MMP-7	MAB9072-100
MMP-9	MAB936-100
NOV/CCN3/IGFBP-9	AF1640
Osteopontin	MAB14332-100
Pentraxin-3/ TSG-14	MAB1826
Thrombospondin-2	MAB16351
TIMP-1	MAB970-100
TNF-alpha	MAB610-100
TREM-1	AF1278
VEGF	MAB293-100
granzyme B (duoet)	844296
CXCL1/GRO alpha Protein	MAB275-100

Table Ap 2-1. Antibodies used for the FPPA experiment in Chapter 3. All antibodies from R&D systems.

Appendix 2.2 RPPA antibodies

Antibody	Supplier	Catalogue Number
FAK1	Cell Signaling Technologies	3285
FAK1 P Y397	Cell Signaling Technologies	3283
IkB-alpha	Cell Signaling Technologies	4812
IKK alpha/beta P Ser176/Ser177	Cell Signaling Technologies	2078
JAK1	Cell Signaling Technologies	3332
MEK1/2	Cell Signaling Technologies	9122
MEK1/2 P Ser217/221	Cell Signaling Technologies	9154
MNK1 (MKNK) P Thr197,Thr202	Cell Signaling Technologies	2111
Met P Tyr1234	Signalway	11227-1
PI3 Kinase p110-alpha	Cell Signaling Technologies	4249
PKA RII P Ser96	Epitomics	1151-1
PKC (pan) P Ser660 (beta-2)	Cell Signaling Technologies	9371
PKC substrate P (R/K)X(S*)(Hyd)(R/k)	Cell Signaling Technologies	2261
PKC-zeta	Cell Signaling Technologies	9372
PKC-zeta/lambda P Thr410/403	Cell Signaling Technologies	9378
IGF-1R beta P Tyr1162,Tyr1163	Invitrogen (Biosource)	44-804G
ErbB-1/EGFR	Cell Signaling Technologies	2232
ErbB-2/Her2/EGFR P Tyr1248/Tyr1173	Cell Signaling Technologies	2244
ErbB-3/Her3/EGFR	Cell Signaling Technologies	4754
ErbB-3/Her3/EGFR P Tyr1289	Cell Signaling Technologies	4791
Stat5	Invitrogen (Biosource)	44-368G
Stat5 P Tyr694	Cell Signaling Technologies	9351
Akt P Thr308	Cell Signaling Technologies	2965
Met	Cell Signaling Technologies	4560
Stat3 P Tyr705	Cell Signaling Technologies	9138
IkB-alpha P Ser32	Cell Signaling Technologies	2859
S6 Ribosomal protein P Ser235,Ser236	Cell Signaling Technologies	2211
p44/42 MAPK (ERK1/2)	Cell Signaling Technologies	9102
p44/42 MAPK (ERK1/2) P Thr202/Thr185,Tyr204/Tyr187	Cell Signaling Technologies	4370
Src	Cell Signaling Technologies	2109
Akt	Cell Signaling Technologies	9272
Akt P Ser473	Cell Signaling Technologies	4060
NFkB p65 Ser536	Cell Signaling Technologies	3033
c-Jun P Ser73	Cell Signaling Technologies	9164
c-Myc	Cell Signaling Technologies	5605
E-Cadherin	Cell Signaling Technologies	3195
Rb	Abcam/Epitomics	ab113074

4E-BP1 P Ser65	Cell Signaling Technologies	9451
4E-BP1 P Thr37,Thr46	Cell Signaling Technologies	2855
beta-Catenin	Cell Signaling Technologies	9562
beta-Catenin P Ser33,Ser37,Thr41	Cell Signaling Technologies	9561
beta-Catenin P Thr41,Ser45	Cell Signaling Technologies	9565
VEGFR P Tyr1175	Cell Signaling Technologies	2478
PTEN	Cell Signaling Technologies	9552
Tsc-2 (Tuberin) P Thr1462	Cell Signaling Technologies	3617
Tsc-2 (Tuberin)	Cell Signaling Technologies	3612
p70 S6 Kinase P Thr389	Cell Signaling Technologies	9205S
PTEN P Ser380,Thr382,Thr383	Cell Signaling Technologies	9554
SAPK/JNK (JNK2)	Cell Signaling Technologies	9258
p70 S6 Kinase P Thr421,Ser424	Cell Signaling Technologies	9204
GSK-3-alpha/beta P Ser21/Ser9	Cell Signaling Technologies	9331
Stat6 P Tyr641	Cell Signaling Technologies	9361
p38 MAPK PThr180,Tyr182	Cell Signaling Technologies	9211
p38 MAPK	Cell Signaling Technologies	9212
mTOR P Ser2448	Cell Signaling Technologies	2971
mTOR	Cell Signaling Technologies	2972
Stat1 P Ser727	Invitrogen (Biosource)	44-382G
Raf P Ser259	Cell Signaling Technologies	9421
PLC-gamma1 P Tyr783	Cell Signaling Technologies	2821
PDK-1 P Ser241	Cell Signaling Technologies	3061
p90 S6 kinase (Rsk1-3) P Thr359,Ser363	Cell Signaling Technologies	9344
PDK-1	Cell Signaling Technologies	3062
p70 S6 Kinase	Cell Signaling Technologies	9202
JAK1 P Tyr1022,Thr1023	Invitrogen (Biosource)	44-422G
c-Myc P Thr58,Ser62	Epitomics	1203-1
SAPK/JNK P Thr183,Tyr185	Cell Signaling Technologies	4668
Rb P Ser807,Ser811	Cell Signaling Technologies	9308
Raf P Ser338	Cell Signaling Technologies	9427
Stat6	Cell Signaling Technologies	9362
Src (family) P Tyr416	Cell Signaling Technologies	2101
Smad2/3 P Ser465/Ser423,Ser467/Ser425	Cell Signaling Technologies	8828
Smad1/5 P Ser463/Ser465	Cell Signaling Technologies	9516
AMPK alpha	Cell Signaling Technologies	2532
AMPK alpha P Thr172	Cell Signaling Technologies	2535
c-Jun N-term	Epitomics	1254-1
EGFR P Tyr1086		369700
IRS-1	Cell Signaling Technologies	2382
GSK-3-beta	Cell Signaling Technologies	9315

PKC-alpha	Beckton Dickinson	610108
Ras	Beckton Dickinson	610001
Stat1	Cell Signaling Technologies	9176
IRS-1 P S636/639	Cell Signaling Technologies	2388
PLC-gamma1	Cell Signaling Technologies	2822
SHP2 P Tyr542	Cell Signaling Technologies	3751
Stat3	Cell Signaling Technologies	12640
VEGFR P Tyr1059	Cell Signaling Technologies	3817
NFkB p105/p50	GeneTex	GTX110585
Calmodulin	Calbiochem	NB12
IGF-1R beta	Cell Signaling Technologies	3027
Grb2	Cell Signaling Technologies	3972
4E-BP1	Cell Signaling Technologies	9452
EGFR P Y992	Cell Signaling Technologies	2235
EGFR P Y1068	Cell Signaling Technologies	2234
MEK 1/2 P Ser217/221	Cell Signaling Technologies	9121
Grb2 P Y237	Abcam	ab32037
B-Raf [EP152Y]	Abcam	ab33899
YAP1 [EP1674Y]	Abcam	ab52771
mTOR P S2448 [EP2426(2)Y]	abcam	ab109268
STAT1 alpha P Y701 [EPR3147]	abcam	ab109457
Eph1A [EPR1786]	abcam	ab133598
EphB3 [EPR8280]	abcam	ab133742
EphB2 [EPR10072(B)]	abcam	ab150411
FOXA1	Millipore (Upstate)	AB4124
Stat3 P Y705	CST	9131
TCF1 (C63O9)	Cell Signaling Technologies	2203
LEF1 (C12A5)	Cell Signaling Technologies	2230
FOXO3a (75D8)	Cell Signaling Technologies	2497
TCF4 (C9B9)	Cell Signaling Technologies	2565
TCF3 (D15G11)	Cell Signaling Technologies	2883
Smad2 (C86F7)	Cell Signaling Technologies	3122
YAP P Ser127	Cell Signaling Technologies	4911
TIMP2 (D18B7)	Cell Signaling Technologies	5738
GATA3 (D13C9)	Cell Signaling Technologies	5852
Stat3 P Ser727	Cell Signaling Technologies	9134
TGF beta (56E4)	Cell Signaling Technologies	3709
Akt substrate P (RXXS/T) (110B7E)	Cell Signaling Technologies	9614
PKA substrate P (RRXS/T) (100G7E)	Cell Signaling Technologies	9624
PKA	Abcam	ab26322
CamKII alpha (22B1) P Thr286	Abcam	ab2724

PKC-alpha P Thr638	Abcam	ab32502
GAPDH	Abcam	ab9484
Raf1 (C-12)	Santa Cruz	sc-133
p90 S6 kinase (Rsk1-3)	Santa Cruz	sc-231
PKC-gamma P Thr514	GeneTex	GTX25778
mTOR P Ser2481	Millipore (Upstate)	09-343SP

Table Ap 2-2. RPPA antibodies.

Appendix 3 Gene sets used for Gene Set Enrichment Analysis

Gene Set ID	Description	Figure used in	Genes	Reference
Houthuijzen iCAFs	Identified from scRNAseq of WEPTn tumours - inflammatory CAFs (iCAFs) enriched for immune modulatory signalling	Figure 3-3	C4B,CAV1,CFB,CXCL12,CYP7B1,CHL1,FGF7,HMCN2,LY6C1,C3,EMB,PRG4,SFRP4,STEAP4	Houthuijzen <i>et al</i> , 2023 ¹
Houthuijzen _myCAFs	Identified from scRNAseq of WEPTn tumours - myofibroblastic CAFs (myCAFs) enriched for ECM modulatory signalling	Figure 3-3	ACTA2,AK1,ASPN,ASS1,BPGM,C1QTNF3,CCN2,CCN4,CD59A,CDH11,CILP,CNN2,COL12A1,COL8A1,CPXM2,CRABP1,CRLF1,CTSC,DDAH1,DKK3,ECRG4,EDNRA,F2R,FZD1,GASK1B,GNG11,GPC1,GPR153,IGFBP3,IL34,ITGB5,KIF26B,LOXL3,LRRC15,MFAP4,MICAL2,NCAM1,NKD2,OCIAD2,OLFML3,PLAAT3,PMEPA1,PTK7,RFLNB,RUNX1,SDC1,SERPINE2,SLC20A1,SMIM3,SPON1,TGFB1,TGFB3,TNC	Houthuijzen <i>et al</i> , 2023 ¹
CAF_S1	One of four subtypes identified by FACs of pre-treatment breast tumours - immune modulatory and defined by CD29 ^{Med} /FAP ^{Hi} /FSP1 ^{Low-Hi} /αSMA ^{Hi} /PDGFRβ ^{Med-Hi} /CAV1 ^{Low}	Figure 3-3	LAMA2,SFRP4,PDGFRA,LRRC15,GREM1,SFRP2,KIAA1462,DCN,CPXM2,PTGFR,WNT2,ADAMTSL1,FBLN5,DPP4,P4HA3,CCDC80,PRICKLE1,GALNT12,ABI3BP,GALNTL2,RSPO3,LOC100288077,PODN,CHRD1,CILP,PRDM6,CCL11,CAPN6,PDGFRL,PTGS2,SEMA3D,DPT,NOX4,ZFH4,EPYC,DCHS1,DNM3,COL6A6,CPZ,FLRT2,NEGR1,BMPER,LOC255480,CXCL12,HTRA3,DIO2,TNNT3,COL6A3,KCNK2,KERA,COL11A1,STMN2,SYNDIG1,ISLR,COL10A1,ADH1B,OMD,GAS1,PTX3,FLNC,SPON1,CST1,CNTN3,COMP,CD177,MFAP4,LSP1,F13A1,PLA2G2A,GDF10,UST,FHAD1,OGN,LIPG,SORCS2,FGF7,ABCA9,WISP2,BEND6,ADH1C,GRIA3,FAT4,FNDC1,SCARA5,IGFL2,KIAA1199,PI16,LOC728640,C3orf55,AOX1,CLCN4,HS3ST3A1,TMEM176B,CST2,FBN1,GXYLT2,FGL2,GABRB2,COL3A1,SLC1A7	Costa <i>et al</i> , 2018 ²

CAF_S4	One of four subtypes identified by FACs of pre-treatment breast tumours - ECM modulatory and defined by CD29 ^{Hi} /FAP ^{Neg} /FSP1 ^{Low-Med} /αSMA ^{Hi} /PDGFRβ ^{Low-Med} /CAV1 ^{Neg-Low}	Figure 3-3	CASQ2,ADAP2,HIGD1B,NRARP,KCNA5,ESAM,CDH6,TMEM74B,AVPR1A,COX4I2,PPP1R14A,ARHGAP15,OR51E2,OR51E1,GJA4,AKAP6,FHL5,GPR20,WFDC1,CACNA1H,ATP1A2,DPY19L2,SLC2A4,RBM20,CD93,PI15,TPSG1,P2RY14,PDGFB,SLCO1C1,SH2D3C,KCNK17,KIAA1274,RASGRP2,LINGO1,NRIP2,ANGPT2,C4orf32,LRR10B,SEMA5B,38231,AOC3,CCDC141,SGCA,CD4,C21orf7,SPTB,MYOM1,DGKB,GPR39,KLHL30,NFASC,C20orf202,TRPC4,SLC38A11,SCN3A,RGS6,RERGL,PDE5A,NEURL1B,PLCE1,SYNPO2,KCNMB1,FMO3,PCDH1,AGAP11,SSTR2,GPR116,EPSS8L1,ARHGAP44,DGKG,HEYL,ACTC1,NPY1R,TRPC6,ACTG2,CAPN12,RBPMS2,SCN4B,RYR2,LMOD1,OSTBETA,CDH8,ENPEP,JPH2,GUCY1A2,GUCY1B3,FAM13C,CSRNP3,CRHR2,ACAN,RASL12,GPR4,HRC,CCDC102B,KCNJ8,PHYHIP1L,RASD2,VSIG2,SLC8A1	Costa <i>et al</i> , 2018 ²
Reactive-like ILC subtype	Genes significantly enriched in the reactive-like subtype of ILC tumours, enriched for epithelial and stromal signalling	Figure 4-9	1227 genes - details can be found at https://www.sciencedirect.com/science/article/pii/S0092867415011952?via%3Dihub Supplementary table S8	Ciriello <i>et al</i> , 2015 ³
Immune Related ILC subtype	Genes significantly enriched in the immune related subtype of ILC tumours, enriched for immune signalling	Figure 4-9	272 genes - details can be found at https://www.sciencedirect.com/science/article/pii/S0092867415011952?via%3Dihub Supplementary table S8	Ciriello <i>et al</i> , 2015 ³

Nardone 2022 FOXA1 ILC 120	ILC-specific FOXA1 target genes	Figure 5-3	MAML3,CERS6,IRF2BP2,NAP1L4,ITGA6,TSKU,RHOB,FOXA1,ARRB1,FLNB,WHSC1L1,C1GALT1,PLEKHF2,MOB2,GAB2,BRF2,SPRY4,RNF141,CASZ1,USP47,KIAA0040,SBF1,ALG8,CDC14A,PROSC,EZR,RAB11FIP1,ASB13,ACER3,MAP3K1,VAV3,GLUL,TSSC4,KCNJ11,HK1,LETM2,EXOC2,NAV2,RAPGEF5,GATA4,SLC22A4,H2AFJ,FAM46C,LSM1,NDUFC2,SLAIN2,C15orf59,AFF1,SNAPC2,DAAM1,SPRED2,HIST1H2BH,PPARD,ZNRF2,SECISBP2L,SNX24,MBOAT2,IPO7,PPID,ABCD3,WDR35,ZNF703,MINOS1,BCAR3,SLC29A1,MUT,AFTPH,CHPT1,FIG4,MCF2L,TACC1,DDX60L,PPP6R3,PER3,NUCB2,ARHGAP29,USP35,SOD2,LGALS3BP,ABCB8,CHN2,WEE1,CTSD,ABLIM3,LGR4,PPP1CA,KIAA0232,SLC1A4,LGALS8,FADD,TSPAN5,FZD7,RPS3,DDHD2,KIF16B,F3,MLLT4,STIM1,IGSF9,BMP5,PPFIBP2,IMMP2L,MGP,PPP1R3D,NUAK2,RRM1,COMMD3,CMBL,THRSP,SLC35E2B,FAIM3,SCOC,BAG4,NBL1,FSIP1,TMEM106B,ARHGEF3,HUS1B,KIF13B,RERE	Nardone <i>et al</i> , 2022 ⁴
Sikora 2014 ILC E2 UP	ILC-specific E2 upregulated genes	Figure 5-3	UXS1,TTC39A,CDKN1C,KIF16B,LEPR,NT5DC3,TOB1,SGCB,EML1,ZIC3,FASN,PELI1,CAMTA1,MICALL1,DSCR3,RBL1,ARMCX2,MAP3K8,PTPN1,ZFP36,WDR76,MSN,SNAI1,ZNF814,KCND3,ARHGEF26,TFCP2L1,CCDC88A,WNT4,DOCK10	Sikora <i>et al</i> , 2014 ⁵
Sikora 2014 ILC E2 DOWN	ILC-specific E2 downregulated genes	Figure 5-3	RAPGEF4,PDE4B,STK32B,NAV3,HLF,NLRP1,SNAP25,GNAI1,RORA,CXXC4,SCGB1D2,CHN2,DACH1,SYNE1,SPARC,SMAGP,GALNT14,RNF128,F3,RNASE4,OVGP1,HGD,LPAR6,ENOX1,QDPR,SIX2,ZNF277,ATP1A1,PLXNC1,EA2,DPYD,HEY1,NR3C2,F7,CCT6B,GLI3,FAH,ANG,TCEAL1,ABLIM1,BACH1,ALS2CR8,CRTC3,NPAS2,RAB15,CBX7,SLC44A4	Sikora <i>et al</i> , 2014 ⁵

Xu 2021 SUM44PE TAMR	Genes upregulated in both tamoxifen resistant and treated SUM44PE cells compared to parental and control cells	Figure 5-3	APOD,QPRT,DNAJC15,MET,SCNN1A,AZGP1,AZGP1,KMO,DDC,IL17RB,PDLIM3,SLC44A4,UGT1A3,MUC1,TARP,MXI1,AKR1C1,TACC1,UGT1A1,CRIP2,RAB31,MMP16,TMPRSS11E,BASP1,UGDH,PHTF2,NT5DC2,LGALS3BP,IFI27,SCGB1D2,PLEKHA5,LAPTM4B,AKR1C2,ATP6V0A4,RGS2,SERHL2,PTPN11,STAT1,GPX2,VTCN1,PSCA,S100A8,ADD3,SLC1A1,DHRS2,GSTM4,CTSO,TIMP1,GULP1,ESR1,FOXI1,C14orf159,SERPINA5,CYP1B1,PDLIM5,RAPGEF2,DSG2,SEL1L3,CALD1,SLC26A2,PIK3R1,KYNU,CREBZF,ADCY3,GJA1,CEBPD,SCARB2,OAS2,STEAP3,RUNX1,UBA6,UBXN4,IFITM2,SCCPDH,RARRES3,PEG10,BTBD3,TRAK2,SERPINA3,CLUAP1,IGF1R,PTHLH,GREB1,STC1,CPD,SYT1,PALLD,MAPT,IFITM3,HSPA2,PKIG,CDH3,DLG1,HMGCS1,PYGL,MYOF,APBB2,EIF2AK2,FN1,KLF4,FRYL,EEF1A2,CD47,SYBU,IGFBP5,ZFP36L1,RPS6KA2,PSD3,AKR1C3,TESMIN,C9orf3,PDS5A,IL6ST,SNRK,PLEKHB1,CTNND2,FAM134B,RAB38,NCK2,SCGB2A2,PLSCR1,UBE2K,LY75-CD302/CD302/LY75,SEMA3C,ABAT,GT2E1,NMI,CLDN3,NEBL,HSPA4L,EMP1,TFF1,MFAP3L,SLC35A3,PTPN12,FAT1,TNFRSF10B,NAT1,NPY1R,SLC39A6,MAP3K7,EIF1AX,TMEM33,MRPS30,RCN1,RBPJ,ARFGEF1,GABARAPL3///GABARAPL1,SMARCA5,RAI14,NMU,COL3A1,ISG15,MSX1,WAC,MYO1B,SECISBP2L,EFR3A,NRCAM,PAPOLA,MAPKAPK2,MED28	Xu <i>et al</i> , 2021 ⁶
Selli 2019 Letrozole Resistant UP	Genes upregulated in letrozole resistant tumours vs dormant tumours after treatment	Figure 5-3	FOSB,LEP,PIP,ACKR1,HSPB6,CRYAB,MYH11,MAOA,MFAP4,OSR1,FABP4,PLIN4,SAA1,CFD,ACTG2,CLDN5,VWF,MYOM1,FOS,THRSF,IL33,PTX3,KRT14,CAVIN2,DPT,CD36,RE RGL,ANKRD30BP3,ID4,PPP1R1A,CLCA2,GPX3,ACACB,CCL14,HBB,CCL15/CCL14,PIGR,HMGCS2,PLIN1,MIR675,H19,FLNC,C7,DNAASE1L3,GPAM,RPRML,FHL1,NR2F1,PROM1,TNNT3,CD34,MEOX1,TUBB2B,SPRY2,DCD,ADH1C,ESAM,IGF1,PTN,C10orf10,GRIA2,PDK4,SPARCL1,CPB1,ANXA1,LDB2,PTGDS,RBP7,ADH1A,COX7A1,EHD2,KRT5,HBA2,HBA1,LIFR,COL14A1,GJC3,ARAP3,KRT15,SVEP1,SFRP1,ADAMTS1,ADGRF5,EMCN,TMEM119,AKR1C2,GNG11,IGFBP6,CLDN11,RASIP1,KRT17,NR4A1,ZFP36,GHR,SOX10,NFIB,YBX3,STC2,GEM,TACC1,CD93,ALDH1A1,JAM2,C1QTNF1,PTGS2,CLDN8,DPYSL2,CYR61,MEOX2,CRISPLD2,VTCN1,SYNM,NOVA1,AKR1C3,ADM,STEAP4,PREX2,GYG2,EGR3,CAV1,DUSP1,TMEM47,PCOLCE2,EGR1,SLC26A3,HMCN1,RSPO3,CNKSRR3,RIBC1,CDR1,NTRK2,RGS5,SELP,LPL,CCDC80,GAS1,PPP1R3C,TGFB2,RCAN2,STK32B,CILP,EFEMP1,GOS2,ARHGEF10,AQP1,TSPAN7,ANGPTL2,RDH5,TAT,S1PR1,ELN,ADCY4,NPY1R,PROS2P,RHOU,ADGRL4,FBLN7,APOD,RAPGEF3,ECSCR,ACTA2,MTURN,CYYR1,MYLK,HSPB8,TXNIP,CD248,JCHAIN,CXCL2,LIMS2	Selli <i>et al</i> , 2019 ⁷

Selli 2019 Letrozole Resistant Down	Genes downregulated in letrozole resistant tumours vs dormant tumours after treatment	Figure 5-3	S100P, TFF1, HIST1H3C, ANKRD30B, HIST1H2APS5, ANKRD30BP1, AGR2, CXADR1, TOP2A, HIST1H2AJ, CEP55, SPP1, MMP11, CYP2A6, HIST1H2AI, SLC9A3R1, SCGB2A1, ISG15, AURKA, HIST2H3C, POTE2, POTE3, HIST2H3A, ANKRD30A, POTE, FPR3, SCGB2A2, HMGB3, CXADR, CP, SCGB1D2, HIST1H2BF, PKIB, TFF3, KRT8, MIR559, CLEC3A, COL11A1, POTE, TERC, MCM2, DLEU7, AGR3, HIST1H2AM, HIST1H3J, NAT1, HIST1H2AE, CKS2, IGHG2, SNO, RA79, C19orf33, MAL2, EPCAM, SMIM22, PRC1, ASCL2, CYP4Z2P, MMP9, CENPF, UCP2, IFI30, CKAP2L, IGKV3-20, COL10A1, CD24P4, KRT19, IGKV1D-17, IGKV3D-11, IGKV1-9, CD24, IGKV3OR2-268, HIST1H3F, SERPINA3, IGKV3-11, NCAPG, IGKV3D-20, HIST1H3D, UBE2C, HIST1H2AH, RIPPLY3, POTE, MIR3143, C1orf64, PAFAH1B3, POTEH, SLC16A3, ANLN, HIST1H1B, CALML5, BMP1B, CRIP1, HIST1H2BG, CAPN13, PTTG3P, CXC, L9, PRR11, APOC1, GBP5, BAMBI, BCAS1, SLC4A8, CFB, LINC00993, HOXC9, SPINT2, UBD, CNB1, IGKV1D-33, CAMP, IGKV1-33, NUSAP1, CYP4Z1, HDAC11, HIST3H2A, KRT8P33, TPX2, PLEKHS1, CDH1, RNU4ATAC, KRT7, EPSTI1, CDCA5, HIST1H2BD, UGT2B28, GP2, PLK4, POTE, MIR3621, CLGN, HIST1H1C, TMEM210, LRRC26, SYCP2, RSPH1, PVALB, TFAP2B, ARHGAP39, OR4K15, NEK2, GINS2, IGHV2-70, PRODH, LRRC6, SLITRK6, MIR4728, CDKN3, ENPP1, SCARNA5, C1QC, MUC1, UGT2B10, HLA-DQA2, IGLL1, ECT2, EXO1, SEMA4B, IFIT1, ESR1, ARHGAP11B, CXCL10, CKMT1B, ANKRD30BP3, LYZ, COMP, CTSZ, PTTG1, APOE, SPDEF, HIST1H3I, FOXA1, PRLR, MMP13, SORD, HIST2H2AB, TSPAN13, CD9, KRT16, H2AFZ, SLC16A6, HIST1H2AD, STMND1, FAM234B, HIST1H2BH, RAB17, SLC39A6, S100A8, C4B, RRES3, C4B_2, HIST1H4L, RAB25, S100A14, HLA-DQA1, GPRC5A, HIST1H4H, PPDPF, HIST1H2BE, UGT2B7, MYB, KIF11, CCR1, BCAS4, ITGB6, ORC6, ENPP5, CNTNAP2, M, IFI6, LTBB, BUB1, BPNT1, PEG10, MSR1, EFHD1, HIST1H4B, HSPA1A, KCNJ3, MATN3, FBP1, MCM4, C1QB, UQCRCQ, EPYC, CKMT1A, TREM1, LAMP5, CA12, SYT13, FCER1G, UBE2T, SFN, SLC12A2, LRP2, CAPS, HIST1H4I, TNFRSF11B, OVOL2, SNORA80E, C8orf74, POLR3K, TRIB3, SHROOM1, UGT2A3, PTTG2, HIST1H3B, HELZ2, SIX4, FAM49B, RSAD2	Selli <i>et al</i> , 2019 ⁷
Simões 2020 STAT3 CSC	STAT3 related genes upregulated in ALDH+ ET treated cells compared to ALDH- cells from metastatic breast cancer patients	Figure 5-3 and 5-6	OSMR, CTTN, MUC1, PBX1, ERFFI1, KITLG, FERMT2, LBP, CDH1, ERBB3, PRLR, FLRT3, CCND1, EPAS1, MAGED2, CEACAM1, CD151, SPTBN1, IL6ST, MARVELD2, LAMB2, TNFSF10, AGT, ERBB2, PLTP, FHL2, ITGB6	Simões <i>et al</i> , 2020 ⁸

Emad 2020 EMT	Genes commonly regulated by TGF β treatment, upregulation of EMT transcription factors and down regulation of E-cadherin in breast cancer cell lines	Figure 6-3	C20ORF19, SNX10, TP73L, KCNK1, BDKRB2, ANXA8, ANXA8L1, LOC728113, RHBDF2, LOC653562, SLC6A10P, SLC6A8, KRT18, CDS1, THBD, NEFM, RPS6KA1, SMPDL3B, ABCA12, RHOD, KRT14, PRKCH, ZBED2, C10ORF10, LRRC1, STAP2, JUP, IL4R, PERP, FGFBP1, MYO1D, FAT2, WWC1, FZD3, ZNF165, SNCA, KIAA1815, PRSS8, SH2D3A, GNAL, BIK, CDH3, KIAA0888, KRT5, GJB3, KIAA0040, CELSR2, F11R, NUP62CL, SERPINB1, SPINT2, POLR3G, ELMO3, IL1RN, TSPAN1, IFI30, PLS1, LOC729884, TMPRSS11E, C1ORF116, ALOX15B, COL17A1, RTE1, TNFRSF6B, LAD1, PTPN3, MST1R, EPHA1, E2F5, KRT6B, STAC, ITGB4, C6ORF105, GLS2, ANXA3, DST, ARHGAP25, DSC2, SLC6A8, LSR, CLDN1, FLJ20366, CYP4F11, CCND2, FGFR2, ABLIM1, XDH, CAMK2B, DSG3, OCLN, NAIP, KRT17, SAA2, SAA1, PRRG4, ANK3, C10ORF116, TMPRSS4, CST6, NDRG1, S100A8, CORO1A, KLK5, EXPH5, IRX4, IRF6, HOOK1, ARTN, FLJ12684, SLC2A9, KLK8, TMEM40, TRIM29, HBEGF, ALDH1A3, RBM35B, MYO5C, CYP27B1, IL1B, NMU, KRT16, CDH1, JAG2, VSNL1, RLN2, CTSL2, SYK, SAA1, EPB41L4B, RNF128, ST14, LEPREL1, PI3, AP1M2, CKMT1A, CKMT1B, GRHL2, ARHGAP8, LOC553158, IGFBP2, IL18, CA9, S100A14, CA2, KRT15, EVA1, TMEM30B, S100A7, KLK7, LGALS7, FST, CXADR, SLPI, RBM35A, RAB25, UCHL1, LK10, TACSTD1, SERPINB2, SPRR1A, FGFR3, SPRR1B, FBLN5, GREM1, COL3A1, COL1A2, DCN, CDH2, ENPP2, POSTN, RGS4, C5ORF13, PRRX1, FBN1, SRGN, SPOCK1, PRR16, DLC1, BIN1, RGL1, IGFBP4, PVRL3, CDH11, OLFML3, MMP2, MYL9, COL5A2, CTGF, PLEKHC1, ZEB1, ROR1, PTGER2, CHN1, PMP22, TRAM2, TAGLN, TNFAIP6, CREB3L1, UGDH, HAS2, DNAJB4, CDKN2C, CCDC92, WNT5A, IGFBP3, PPM1D, FILIP1L, PDGFC, TBX3, XYLT1, FAP, DPT, STC1, KRT81, MMP1, HS3ST2, LMCD1, N-PAC, 38961, NR2F1, SCCPDH, MLPH, LTBP2, TPM1, ANKRD25, DDR2, SEMA5A, TGFB1I1, PCOLCE, STARD13, NID1, SYNC1, ENOX1, FSTL1, VIM, MME, C10ORF56, LTBP1, NRP1, THY1, NEBL, TNS3, ECM1, FBLN1, TUBA1A, COPZ2, CYBRD1, PPAP2B, PTX3, FADS2, BGN, TSHZ1, ZBTB38,	Taube <i>et al</i> , 2010 ⁹ and Emad <i>et al</i> , 2020 ¹⁰
------------------	--	------------	---	---

Table Ap 3-1. Details of gene sets used for gene set enrichment analysis.

Appendix 4 IL-6 dependent gene signature literature review

A literature review of genes significantly upregulated by CAF CM treatment of SUM44PE cells, dependent on IL-6 (see Chapter 4). Carried out in March 2023, using Pubmed and the search terms “Gene” + “breast cancer” or “IL-6” or “STAT3”. Genes highlighted in orange are part of a previously published IL-6 response signature across ER+ cell lines¹¹ and those highlighted in yellow are not reported to have STAT3 bound to the promoters (chip-atlas.org). OE, over-expression, KD, knockdown, ER, estrogen receptor, BC, breast cancer, NA, not available, DFS, disease free survival, OS, overall survival, TNBC, triple negative breast cancer, EMT, epithelial to mesenchymal transition, AI, aromatase inhibitors

Gene	Description	Significantly higher in ER+ ILC v IDC in TCGA?	Role in Breast Cancer	Role with IL-6/STAT3
ACOX2	acyl-CoA oxidase 2(ACOX2)	YES	Involved in fatty acid metabolism, higher expression in luminal A breast cancer. ¹² Intronic start variant highly expressed in ER+ BC, ERa regulated and associated with good outcome ¹³	STAT3 knockdown decreases Acox2 expression in mouse respiratory epithelial cells ¹⁴
AKR1B10	aldo-keto reductase family 1 member B10(AKR1B10)	NO	Drives metastasis ^{15–17} Activates PI3K, MAPK pathways and EMT ¹⁷ . High expression = worse DFS, ¹⁷ in ER- and HER2+, not ER+ ¹⁶ High serum levels in BC patients ¹⁸ High expression promotes fatty acid oxidation ¹⁶ and lipid messenger signalling ¹⁹	<i>Stat3</i> and <i>Akr1b10</i> upregulated in non-obese type-2 diabetes rat model ²⁰ OE of thyroid enzyme DIO downregulates <i>STAT3</i> and <i>AKR1B10</i> in renal cell carcinoma ²¹ KD reduces <i>IL6</i> expression in colorectal cancer cells ²²

ASPH	aspartate beta-hydroxylase(ASPH)	NO	Tamoxifen resistance in luminal B ²³ and doxorubicin resistance in TNBC ²⁴ Associated with poor survival ^{25,26} Promotes metastasis <i>in vitro</i> ^{26,27}	NA
BBOX1	gamma-butyrobetaine hydroxylase 1(BBOX1)	YES	Promotes TNBC growth via stabilising IP3R3 mediated calcium signalling ²⁸ High expression in HER2+ recurrent tumours ²⁹	NA
BCL3	b-cell lymphoma 3 protein(BCL3)	YES	High expression correlates with very poor relapse free survival in ILC ³⁰ IGFBP5 downregulation promotes fulvestrant resistance in MCF7 cells via BCL3 activation ³¹ Higher BCL3 TF activity in isolated ILC tumour cells from PDXs than non-ILC cells ³² Promotes BC cell migration ^{33,34}	Well known IL-6/STAT3 target ^{35,36}
C1orf167	chromosome 1 open reading frame 167(C1orf167)	NA	NA	NA
CDC7	cell division cycle 7(CDC7)	NO	CDC7 inhibition synergistic with EGFR inhibition ³⁷ and DNA damaging chemotherapies ³⁸ Low expression in TNBC = poor prognosis ³⁹	STAT3 and CDC7 levels high in male BC patient serum ⁴⁰ CDC7 inhibition reduces STAT3 activation ⁴¹
CEBPD	CCAAT enhancer binding protein delta(CEBPD)	YES	Promoter methylation associated with metastasis ⁴² <i>In vitro</i> suppresses MCF7 EMT and migration ⁴³	Common ER+ breast cell line IL-6 target gene ¹¹

			Promotes CSC formation in TNBC and in ER+ BC in the presence of IL-6 or in hypoxic conditions ⁴⁴	Drives IL6R expression and increased <i>CEBPD</i> expression promotes STAT3+ CSC formation ⁴⁴
<i>CHI3L2</i>	chitinase 3 like 2(<i>CHI3L2</i>) AKA: YKL39	YES	Expressed by macrophages, promotes angiogenesis ⁴⁵ High expression after chemotherapy correlates with high rate of metastasis Chemo-resistant patients have high M2 <i>CHI3L2</i> + macrophages ⁴⁶	NA
<i>CLCA2</i>	chloride channel accessory 2(<i>CLCA2</i>)	YES	Inhibits proliferation ⁴⁷ and EMT ^{48,49}	NDRG1 KD in cervical and ovarian cancer cells upregulates <i>IL6</i> and downregulates <i>CLCA2</i> ⁵⁰ STAT3 inhibits <i>Clca5</i> (mouse orthologue) during mammary involution ⁵¹
<i>CYP1B1</i>	cytochrome P450 family 1 subfamily B member 1(<i>CYP1B1</i>)	YES	Mono-oxygenase involved in E2 metabolism, promotes tumorigenesis Up in tamoxifen resistant ILC cell lines ⁶ Driven by HIF1 α and ER α in ER+ BC ⁵² Promotes proliferation, migration and invasion via Wnt signalling ⁵³ Promotes EMT and chemo-resistance in TNBC cell lines ⁵⁴	Catalyses estrogen, promotes 4-OHE2 and ER α induction of IL-6 in prostate cancer cells ⁵⁵ Activated by IL-6 via downregulation of miR27b in colorectal cancer cells ⁵⁶ <i>IL6</i> and <i>CYP1B1</i> both induced by AHR ⁵⁷ and induced by IL-6 signalling
<i>DEPP1</i>	DEPP1 autophagy regulator(<i>DEPP1</i>)	YES	Decreased expression in breast cancer v normal tissue, low expression = poor survival ⁵⁸	NA
<i>DGAT2</i>	diacylglycerol O-acyltransferase 2(<i>DGAT2</i>)	YES	Inhibition leads to reduced EMT, migration and CSC gene expression, sensitised to radiotherapy ⁵⁹	Fatty acid metabolism, IL6 and DGAT2 involved in liver damage and type-2 diabetes

			Post-transcriptionally downregulated by IRE α in TNBC, high expression sensitises to starvation ⁶⁰	miR221-3p induces <i>IL6</i> and suppresses <i>DGAT2</i> expression in adipocytes ⁶¹
FAM83A	family with sequence similarity 83 member A(FAM83A)	NO	High expression in cancer vs normal breast, high expression = poor survival ^{62,63} Promotes chemoresistance and CSC phenotype in TNBC ⁶⁴ Drives resistance to trastuzumab ^{65,66} and EGFR targeted therapies ⁶⁷	NA
FGD3	FYVE, RhoGEF and PH domain containing 3(FGD3)	YES	High expression is associated with better prognosis ^{68–70} Inhibits cancer cell migration ⁷¹ Co-regulated with <i>SUSD3</i>	NA
GPCPD1	glycerophosphocholine phosphodiesterase 1(GPCPD1)	YES	Highest in ER- HER2+ BC, expression driven by HER2 ⁷² Induced by hypoxia, increases tumour growth via mitophagy in TNBC ⁷³ Promotes MCF7 migration ⁷⁴	STAT3 binding site in promoter, inhibit STAT3 reduces expression ⁷²
GSDMC	gasdermin C(GSDMC)	NO	High expression correlates with poor survival ^{75–78} Correlates with immune checkpoints in BC ⁷⁶	Expression driven by pSTAT3 and nuclear PD-L1, switches from apoptosis to pyroptosis ⁷⁵
H2BC5	H2B clustered histone 5(H2BC5)	NO	NA	NA
HMOX1	heme oxygenase 1(HMOX1)	NO	Upregulated in chemotherapy-resistant TNBC cell lines ⁷⁹ , KD increases response to chemotherapies <i>in vitro</i> ⁸⁰ Suppresses TGF β induced EMT in MCF7 cells ⁸¹	Inhibition of JAK2 suppressed <i>Hmox1</i> expression in dendritic cells and prevents the formation of an immunosuppressive TME in a model of PDAC ⁸⁵ HMOX1 stabilised by 14-3-3 ζ and promotes growth in hepatocellular

			Associated with poorer survival for TNBC in univariate analyses ^{82,83} and poorer DFS across all BC ⁸⁴	carcinoma <i>in vitro</i> and <i>in vivo</i> dependent on HMOX1 promoting STAT3 activation ⁸⁶
HPX	hemopexin(HPX)	YES	Up in BC compared to normal breast ^{87,88}	HPX and IL-6 are both acute phase response proteins
IL13RA1	interleukin 13 receptor subunit alpha 1(IL13RA1)	NO	High expression = poor prognosis ^{89,90} Expression in MDSCs increases lung metastases in 4T1 model which was suppressed with anti-IL-13 ⁹¹ Part of a glycolysis related gene signature that predicts poor OS ⁹⁰	KD in macrophages suppresses ox-LDL induced STAT3 activation and IL-6 upregulation ⁹² Induced by trans-IL-6 signalling in airway smooth muscle cells ⁹³
IL4R	interleukin 4 receptor(IL4R)	YES	KD suppresses 4T1 lung metastasis ⁹⁴ IL4 stimulation enhances TNBC cell proliferation via increased glutamine and glucose metabolism ⁹⁵ Depletion of IL4R+ bone macrophages suppresses metastatic outgrowth ⁹⁶ Toxin fused to IL4 selectively depletes IL4R+ immunosuppressive cells and TNBC cells, suppresses tumour growth <i>in vivo</i> ⁹⁷	IL-6 treatment of MDSCs induces IL4R expression dependent on JAK/STAT signalling ⁹⁸ IL4 stimulation activates STAT3 and promotes increased IL4R expression in CD8 T-cells ⁹⁹ IL4R/STAT6/STAT3/PPAR γ pathway promotes M2 macrophage polarization in adipose tissue, reducing inflammation and improving insulin sensitivity ¹⁰⁰
LBP	lipopolysaccharide binding protein(LBP)	NO	NA	STAT3-related gene up in ALDH high endocrine resistant BC samples ⁸
LRG1	leucine rich alpha-2-glycoprotein 1(LRG1)	YES	High expression in metastatic lymph nodes ¹⁰¹ High expression in BC v normal breast ¹⁰²	Activated by IL-6/STAT3, drives metastasis in colorectal cancer ¹⁰⁶

			Protective against apoptosis in MCF7s ¹⁰³ Upregulated in PIK3CA mutant tumours, high expression = good response to AI therapy ¹⁰⁴ High expression = poor prognosis ¹⁰⁵	Promotes gastric cancer proliferation and angiogenesis via STAT3/VEGFA, high expression = poor survival ¹⁰² IL-6 treatment of endothelial cells drives LRG1 dependent angiogenesis ¹⁰⁷ Promotes neutrophil recruitment and fibrosis in damaged cornea via activation of IL6/STAT3 ¹⁰⁸
LRP2	LDL receptor related protein 2(LRP2)	YES	Low expression correlates with distant metastasis free survival ¹⁰⁹ Induced by estrogen and progesterone in mammary stem cells and luminal progenitors	Leptin and clusterin signal through LRP2 to activate STAT3 ^{110,111} STAT3 and LRP2 mutated in peripheral T-cell lymphoma ¹¹²
LRRC4C	leucine rich repeat containing 4C(LRRC4C)	YES	NA	NA
LTF	lactotransferrin(LTF)	YES	High concentration of recombinant LTF inhibits migration ^{113,114} and causes apoptosis in BC cell lines ¹¹⁵ Reduced expression in mets v primary ¹¹³ Overexpressed in PIK3CA mutant tumours ¹¹⁶ Lower in BC compared to normal breast	High expression = good prognosis in prostate cancer, OE LTF – reduced STAT3 and GM-CSF expression ¹¹⁷ Estrogen responsive gene, KD STAT3 in uterine tissue upregulates LTF ¹¹⁸ LTF and IL-6 both upregulated in lots of inflammatory diseases Bovine LTF inhibits migration and reverts EMT, downregulates IL-6/STAT3 in glioblastoma ¹¹⁹ and oral cell squamous carcinoma ¹²⁰

MUC5B	mucin 5B, oligomeric mucus/gel-forming(MUC5B)	NO	MUC5B-KD increases response to chemotherapies and reduces proliferation ¹²¹ OE promotes invasion, proliferation <i>in vitro</i> and <i>in vivo</i> ¹²²	Induced by constitutively active STAT3 ¹²³ Induced by <i>M. pneumonia</i> infection of lungs via IL-6/STAT3 ¹²⁴ IL-6 and MUC5B involved in lung diseases eg pulmonary fibrosis, asthma, cystic fibrosis
MUCL1	mucin like 1(MUCL1) (Small breast epithelial mucin SBEM)	YES	Up in HER2+ BC, ¹²⁵ associated with trastuzumab resistance and poor prognosis ¹²⁶ and induced downstream of HER2/PI3K ¹²⁷ Promotes EMT and invasion ¹²⁸ Higher expression in BC vs normal breast ^{129,130} Correlates with metastasis and poor prognosis ^{128,131}	NA
NFAM1	NFAT activating protein with ITAM motif 1(NFAM1)	YES	NA	Reduced IL-6 expression in CD40L stimulated NFAM1 ^{-/-} monocytes ¹³²
PKNOX2	PBX/knotted 1 homeobox 2(PKNOX2)	YES	NA	NA
PPP1R1B	protein phosphatase 1 regulatory inhibitor subunit 1B(PPP1R1B)	YES	Two isoforms (DARPP-32 and tDARPP) t-DARPP promotes trastuzumab resistant HER2+ BC via IGF1R/PI3K ¹³³ , <i>PPP1R1B</i> up in trastuzumab resistant BC ¹³⁴ Inhibits PKA and PP1 activity	DARPP-32 promotes STAT3 activation ^{136,137} and STAT3 dependent angiogenesis in gastric cancer ¹³⁸

			Low expression of DARPP-32 isoform = poor prognosis in ER+ and HER2+ ¹³⁵	
PRSS22	serine protease 22(PRSS22)	YES	NA	NA
RNF213	ring finger protein 213(RNF213)	NO	E3 ubiquitin ligase negatively regulated by the phosphatase PTP1B – RNF213 upregulation causes cell death in HER2+ cells <i>in vivo</i> in hypoxic conditions ¹³⁹	NA
S100A9	S100 calcium binding protein A9(S100A9)	NO	Commonly induced in ER+ breast cancer cell lines by IL-6 stimulation¹¹ Negative correlation with <i>ESR1</i> and positive correlation with HER2¹⁴⁰ Tumour and stroma expression significantly associated with poorer OS¹⁴⁰, high expression associated with poorer OS across all BC and Luminal A BC¹⁴¹ High infiltration of MDSCs expressing S100A9 in TNBC, promote immunosuppression¹⁴² Inhibition of S100A9 drives M2 to M1 switch in macrophages, relieves immunosuppression and decreases lung metastases in 4T1 model <i>in vivo</i>¹⁴³	Well known IL-6/STAT3 responsive gene in cancer and inflammation^{11,144,145} S100A9 promotes <i>IL6</i> expression and STAT3 activation in osteocytes¹⁴⁶, breast cancer cells¹⁴⁰ and fibroblasts¹⁴⁷ STAT3 induces S100A9 expression in myeloid cells, drives immunosuppression by suppressing dendritic cell maturation and increasing MDSC activity in cancer and inflammatory diseases^{148,149} Stabilises STAT3/CEBPβ complex¹⁴⁹
SECTM1	secreted and transmembrane 1(SECTM1)	NO	NA	Induced by interferon stimulation of monocytes and has a STAT1/STAT3 binding site in its promoter ¹⁵⁰

SEPTIN3	septin 3(SEPTIN3)	NO	NA	NA
SLC28A3	solute carrier family 28 member 3(SLC28A3)	YES	NA	NA
SUSD3	sushi domain containing 3(SUSD3)	YES	Co-regulated with <i>FGD3</i> Estrogen responsive, marker of AI response, good prognosis ^{70,151} KD decreases cell cycle progression, reduces migration and stress fibre formation, co-localises with E-cadherin ¹⁵¹	NA
TC2N	tandem C2 domains, nuclear(TC2N)	NO	Low expression associated with poorer outcomes OE in MCF7 and MDA-MB-231 suppresses proliferation <i>in vitro</i> and <i>in vivo</i> by preventing Akt activity ¹⁵²	NA
THRSP	thyroid hormone responsive(THRSP)	YES	High expression associated with increased recurrence ¹⁵³ Increased expression in breast cancer and KD reduces cell viability, increases apoptosis ¹⁵⁴	NA
TMC5	transmembrane channel like 5(TMC5)	YES	Commonly induced by IL-6 in ER+ breast cancer cell lines ¹¹ Significantly more highly expressed in <i>PIK3CA</i> mutant ER+ BC ¹¹⁶ Lower expression predictive of increased risk of relapse ¹⁵⁵	Closely correlates with STAT3 expression and predicted to be regulated by STAT3 in pancreatic adenocarcinoma ¹⁵⁶
VNN2	vanin 2(VNN2)	YES	NA	NA

Table Ap 4-1. IL-6 dependent gene signature literature review.

Appendices References

1. Houthuijzen, J. M. *et al.* CD26-negative and CD26-positive tissue-resident fibroblasts contribute to functionally distinct CAF subpopulations in breast cancer. *Nat. Commun.* **14**, 183 (2023).
2. Costa, A. *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **33**, 463-479.e10 (2018).
3. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163**, 506–519 (2015).
4. Nardone, A. *et al.* A Distinct Chromatin State Drives Therapeutic Resistance in Invasive Lobular Breast Cancer. *Cancer Res.* 3673–3686 (2022) doi:10.1158/0008-5472.can-21-3186.
5. Sikora, M. J. *et al.* Invasive lobular carcinoma cell lines are characterized by unique estrogen-mediated gene expression patterns and altered tamoxifen response. *Cancer Res.* **74**, 1463–1474 (2014).
6. Xu, D., Zhang, Y. & Jin, F. The role of AKR1 family in tamoxifen resistant invasive lobular breast cancer based on data mining. *BMC Cancer* **21**, 1–14 (2021).
7. Selli, C. *et al.* Molecular changes during extended neoadjuvant letrozole treatment of breast cancer: distinguishing acquired resistance from dormant tumours. *Breast Cancer Res.* **21**, 1–15 (2019).
8. Simões, B. M. *et al.* Targeting STAT3 signaling using stabilised sulforaphane (SFX-01) inhibits endocrine resistant stem-like cells in ER-positive breast cancer. *Oncogene* **39**, 4896–4908 (2020).
9. Taube, J. H. *et al.* Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15449–15454 (2010).
10. Emad, A. *et al.* Superior breast cancer metastasis risk stratification using an epithelial-mesenchymal-amoeboïd transition gene signature. *Breast Cancer Res.* **22**, 1–13 (2020).
11. Casneuf, T. *et al.* Interleukin-6 is a potential therapeutic target in interleukin-6 dependent, estrogen receptor- α - Positive breast cancer. *Breast Cancer Targets Ther.* **8**, 13–27 (2016).
12. Yousuf, U., Sofi, S., Makhdoomi, A. & Mir, M. A. Identification and analysis of dysregulated fatty acid metabolism genes in breast cancer subtypes. *Med. Oncol.* **39**, 1–15 (2022).
13. Bjørklund, S. S. *et al.* Expression of an estrogen-regulated variant transcript of the peroxisomal branched chain fatty acid oxidase ACOX2 in breast carcinomas. *BMC Cancer* **15**, 1–13 (2015).
14. Ikegami, M., Falcone, A. & Whitsett, J. A. STAT-3 regulates surfactant phospholipid homeostasis in normal lung and during endotoxin-mediated lung

- injury. *J. Appl. Physiol.* **104**, 1753–1760 (2008).
15. Huang, C. *et al.* AKR1B10 promotes breast cancer metastasis through integrin $\alpha 5/\delta$ -catenin mediated FAK/Src/Rac1 signaling pathway. *Oncotarget* **7**, 43779–43791 (2016).
16. van Weverwijk, A. *et al.* Metabolic adaptability in metastatic breast cancer by AKR1B10-dependent balancing of glycolysis and fatty acid oxidation. *Nat. Commun.* **10**, (2019).
17. Qu, J. *et al.* AKR1B10 promotes breast cancer cell proliferation and migration via the PI3K/AKT/NF- κ B signaling pathway. *Cell Biosci.* **11**, 1–13 (2021).
18. Cao, Z. *et al.* AKR1B10 as a Potential Novel Serum Biomarker for Breast Cancer: A Pilot Study. *Front. Oncol.* **12**, 1–13 (2022).
19. Huang, C. *et al.* AKR1B10 activates diacylglycerol (DAG) second messenger in breast cancer cells. *Mol. Carcinog.* **57**, 1300–1310 (2018).
20. Sárközy, M. *et al.* Transcriptomic alterations in the heart of non-obese type 2 diabetic Goto-Kakizaki rats. *Cardiovasc. Diabetol.* **15**, 1–21 (2016).
21. Popławski, P. *et al.* Restoration of type 1 iodothyronine deiodinase expression in renal cancer cells downregulates oncoproteins and affects key metabolic pathways as well as anti-oxidative system. *PLoS One* **12**, 1–26 (2017).
22. Liu, C. *et al.* AKR1B10 accelerates the production of proinflammatory cytokines via the NF- κ B signaling pathway in colon cancer. *J. Mol. Histol.* **53**, 781–791 (2022).
23. Shimoda, M. *et al.* Endocrine sensitivity of estrogen receptor-positive breast cancer is negatively correlated with aspartate- β -hydroxylase expression. *Cancer Sci.* **108**, 2454–2461 (2017).
24. Bai, X. *et al.* Bio-nanoparticle based therapeutic vaccine induces immunogenic response against triple negative breast cancer. *Am. J. Cancer Res.* **11**, 4141–4174 (2021).
25. Barboro, P. *et al.* The prognostic value of aspartate beta-hydroxylase in early breast cancer. *Int. J. Biol. Markers* **37**, 328–335 (2022).
26. Lin, Q. *et al.* Multi-organ metastasis as destination for breast cancer cells guided by biomechanical architecture. *Am. J. Cancer Res.* **11**, 2537–2567 (2021).
27. Lin, Q. *et al.* ASPH-notch Axis guided Exosomal delivery of Prometastatic Secretome renders breast Cancer multi-organ metastasis. *Mol. Cancer* **18**, 1–17 (2019).
28. Liao, C. *et al.* Identification of bbox1 as a therapeutic target in triple-negative breast cancer. *Cancer Discov.* **10**, 1707–1721 (2020).
29. Kim, Y. S. *et al.* Analysis of Genomic Alterations Associated with Recurrence in Early Stage HER2-Positive Breast Cancer. *Cancers (Basel)*. **14**, 1–13 (2022).
30. Czapiewski, P. *et al.* BCL3 expression is strongly associated with the occurrence of breast cancer relapse under tamoxifen treatment in a retrospective cohort

- study. *Virchows Arch.* (2022) doi:10.1007/s00428-021-03238-8.
31. Leyh, B., Dittmer, A., Lange, T., Martens, J. W. M. & Dittmer, J. Stromal cells promote anti-estrogen resistance of breast cancer cells through an insulin-like growth factor binding protein 5 (IGFBP5)/B-cell leukemia/lymphoma 3 (Bcl-3) axis. *Oncotarget* **6**, 39307–39328 (2015).
 32. Sflomos, G. *et al.* Intraductal xenografts show lobular carcinoma cells rely on their own extracellular matrix and LOXL1. *EMBO Mol. Med.* **13**, 1–19 (2021).
 33. Turnham, D. J. *et al.* Bcl-3 promotes multi-modal tumour cell migration via NF- κ B1 mediated regulation of Cdc42. *Carcinogenesis* **41**, 1432–1443 (2020).
 34. Wakefield, A. *et al.* Bcl3 selectively promotes metastasis of ERBB2-driven mammary tumors. *Cancer Res.* **73**, 745–755 (2013).
 35. Brocke-Heidrich, K. *et al.* BCL3 is induced by IL-6 via Stat3 binding to intronic enhancer HS4 and represses its own transcription. *Oncogene* **25**, 7297–7304 (2006).
 36. Siersbæk, R. *et al.* IL6/STAT3 Signaling Hijacks Estrogen Receptor α Enhancers to Drive Breast Cancer Metastasis. *Cancer Cell* **38**, 412-423.e9 (2020).
 37. McLaughlin, R. P. *et al.* A kinase inhibitor screen identifies a dual cdc7/CDK9 inhibitor to sensitise triple-negative breast cancer to EGFR-targeted therapy. *Breast Cancer Res.* **21**, 1–15 (2019).
 38. Iwai, K. *et al.* A CDC7 inhibitor sensitizes DNA-damaging chemotherapies by suppressing homologous recombination repair to delay DNA damage recovery. *Sci. Adv.* **7**, 1–15 (2021).
 39. Jia, R. *et al.* Identification of key genes unique to the luminal a and basal-like breast cancer subtypes via bioinformatic analysis. *World J. Surg. Oncol.* **18**, 1–11 (2020).
 40. Zografos, E. *et al.* Serum proteomic signatures of male breast cancer. *Cancer Genomics and Proteomics* **16**, 129–137 (2019).
 41. Parri, E., Kuusanmäki, H., van Adrichem, A. J., Kaustio, M. & Wennerberg, K. Identification of novel regulators of STAT3 activity. *PLoS One* **15**, 1–18 (2020).
 42. Palmieri, C. *et al.* Site-specific CpG methylation in the CCAAT / enhancer binding protein delta (CEBP d) CpG island in breast cancer is associated with metastatic relapse. 732–738 (2012) doi:10.1038/bjc.2012.308.
 43. Mendoza-Villanueva, D. *et al.* The C/EBP δ protein is stabilized by estrogen receptor α activity, inhibits SNAI2 expression and associates with good prognosis in breast cancer. *Oncogene* **35**, 6166–6176 (2016).
 44. Balamurugan, K. *et al.* C/EBP δ links IL-6 and HIF-1 signaling to promote breast cancer stem cell-associated phenotypes. *Oncogene* **38**, 3765–3780 (2019).
 45. Liu, T. *et al.* Tumor-associated macrophages in human breast cancer produce new monocyte attracting and pro-angiogenic factor YKL-39 indicative for

- increased metastasis after neoadjuvant chemotherapy. *Oncoimmunology* **7**, 1–17 (2018).
46. Litviakov, N. *et al.* Expression of M2 macrophage markers YKL-39 and CCL18 in breast cancer is associated with the effect of neoadjuvant chemotherapy. *Cancer Chemother. Pharmacol.* **82**, 99–109 (2018).
 47. Walia, V. *et al.* hCLCA2 is a p53-inducible inhibitor of breast cancer cell proliferation. *Cancer Res.* **69**, 6624–6632 (2009).
 48. Walia, V. *et al.* Loss of breast epithelial marker hCLCA2 promotes epithelial-to-mesenchymal transition and indicates higher risk of metastasis. *Oncogene* **31**, 2237–2246 (2012).
 49. Ramena, G., Yin, Y., Yu, Y., Walia, V. & Elble, R. C. CLCA2 interactor EVA1 is required for mammary epithelial cell differentiation. *PLoS One* **11**, 1–21 (2016).
 50. Zhao, G. *et al.* Identification of NDRG1-regulated genes associated with invasive potential in cervical and ovarian cancer cells. *Biochem. Biophys. Res. Commun.* **408**, 154–159 (2011).
 51. Hughes, K., Blanck, M., Pensa, S. & Watson, C. J. Stat3 modulates chloride channel accessory protein expression in normal and neoplastic mammary tissue. *Cell Death Dis.* **7**, 1–9 (2016).
 52. Min, J. Y. *et al.* Upregulation of CYP1B1 by hypoxia is mediated by ER α activation in breast cancer cells. *Am. J. Cancer Res.* **12**, 2798–2816 (2022).
 53. Lei, T. *et al.* ZNF276 promotes the malignant phenotype of breast carcinoma by activating the CYP1B1-mediated Wnt/ β -catenin pathway. *Cell Death Dis.* **13**, 1–15 (2022).
 54. Hollis, P. R. *et al.* CYP1B1 Augments the Mesenchymal, Claudin-Low, and Chemoresistant Phenotypes of Triple-Negative Breast Cancer Cells. *Int. J. Mol. Sci.* **23**, 1–17 (2022).
 55. Lin, Q. *et al.* CYP1B1-catalyzed 4-OHE2 promotes the castration resistance of prostate cancer stem cells by estrogen receptor α -mediated IL6 activation. *Cell Commun. Signal.* **20**, 1–16 (2022).
 56. Patel, S. A. A. *et al.* Interleukin-6 mediated upregulation of CYP1B1 and CYP2E1 in colorectal cancer involves DNA methylation, miR27b and STAT3. *Br. J. Cancer* **111**, 2287–2296 (2014).
 57. Litzenburger, U. M. *et al.* Constitutive IDO expression in human cancer is sustained by an autocrine signaling loop involving IL-6, STAT3 and the AHR. *Oncotarget* **5**, 1038–1051 (2014).
 58. Deng, J. *et al.* Decreased expression of C10orf10 and its prognostic significance in human breast cancer. *PLoS One* **9**, 1–8 (2014).
 59. Nisticò, C. *et al.* Lipid droplet biosynthesis impairment through dgat2 inhibition sensitizes mcf7 breast cancer cells to radiation. *Int. J. Mol. Sci.* **22**, 1–18 (2021).
 60. Almanza, A. *et al.* Regulated IRE1 α -dependent decay (RIDD)-mediated

- reprogramming of lipid metabolism in cancer. *Nat. Commun.* **13**, (2022).
61. Ahonen, M. A. *et al.* Human adipocyte differentiation and composition of disease-relevant lipids are regulated by miR-221-3p. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1866**, (2021).
 62. Jin, Y. *et al.* Comprehensive analysis of the expression, prognostic significance, and function of FAM83 family members in breast cancer. *World J. Surg. Oncol.* **20**, 1–18 (2022).
 63. Marino, N. *et al.* FAM83A is a potential biomarker for breast cancer initiation. *Biomark. Res.* **10**, 1–14 (2022).
 64. Liu, C., Jiang, Y. & Han, B. Mir-613 suppresses chemoresistance and stemness in triple-negative breast cancer by targeting fam83a. *Cancer Manag. Res.* **12**, 12623–12633 (2020).
 65. Boyer, A. P., Collier, T. S., Vidavsky, I. & Bose, R. Quantitative proteomics with siRNA screening identifies novel mechanisms of trastuzumab resistance in HER2 amplified breast cancers. *Mol. Cell. Proteomics* **12**, 180–193 (2013).
 66. Bartel, C. A. & Jackson, M. W. HER2-positive breast cancer cells expressing elevated FAM83A are sensitive to FAM83A loss. *PLoS One* **12**, 1–18 (2017).
 67. Lee, S. Y. *et al.* FAM83A confers EGFR-TKI resistance in breast cancer cells and in mice. *J. Clin. Invest.* **122**, 3211–3220 (2012).
 68. Susini, T. *et al.* Immunohistochemical evaluation of fgd3 expression: A new strong prognostic factor in invasive breast cancer. *Cancers (Basel)*. **13**, (2021).
 69. Willis, S. *et al.* High Expression of FGD3, a Putative Regulator of Cell Morphology and Motility, Is Prognostic of Favorable Outcome in Multiple Cancers. *JCO Precis. Oncol.* 1–13 (2017) doi:10.1200/po.17.00009.
 70. Ou Yang, T.-H., Cheng, W., Zheng, T., Maurer, M. A. & Anastassiou, D. Breast Cancer Prognostic Biomarker Using Attractor Metagenes and the FGD3–SUSD3 Metagene. *Cancer Epidemiol. Biomarkers Prev.* **23**, 2850–2856 (2014).
 71. Hayakawa, M. *et al.* Novel insights into FGD3, a putative GEF for Cdc42, that undergoes SCFFWD1/β-TrCP -mediated proteasomal degradation analogous to that of its homologue FGD1 but regulates cell morphology and motility differently from FGD1. *Genes to Cells* **13**, 329–342 (2008).
 72. Keller, M. *et al.* Inhibiting the glycerophosphodiesterase EDI3 in ER - HER2 + breast cancer cells resistant to HER2 - targeted therapy reduces viability and tumour growth. *J. Exp. Clin. Cancer Res.* 1–18 (2023) doi:10.1186/s13046-022-02578-w.
 73. Liu, Y. *et al.* Hypoxia-induced GPCPD1 depalmitoylation triggers mitophagy via regulating PRKN-mediated ubiquitination of VDAC1. *Autophagy* **00**, 1–21 (2023).
 74. Lesjak, M. S. *et al.* EDI3 links choline metabolism to integrin expression, cell adhesion and spreading. *Cell Adh. Migr.* **8**, 499–508 (2014).

75. Hou, J. *et al.* PD-L1-mediated gasdermin C expression switches apoptosis to pyroptosis in cancer cells and facilitates tumour necrosis. *Nat. Cell Biol.* **22**, 1264–1275 (2020).
76. Sun, K., Chen, R., Li, J. & Luo, Z. LINC00511 / hsa - miR - 573 axis - mediated high expression of Gasdermin C associates with dismal prognosis and tumor immune infiltration of breast cancer. *Sci. Rep.* 1–20 (2022) doi:10.1038/s41598-022-19247-9.
77. Xu, D., Ji, Z. & Qiang, L. Molecular Characteristics, Clinical Implication, and Cancer Immunity Interactions of Pyroptosis-Related Genes in Breast Cancer. *Front. Med.* **8**, 1–15 (2021).
78. Luo, J. & Lai, J. Pyroptosis-related molecular classification and immune microenvironment infiltration in breast cancer: A novel therapeutic target. *J. Cell. Mol. Med.* **26**, 2259–2272 (2022).
79. Memar, M., Farazmandfar, T., Sabaghian, A., Shahbazi, M. & Gosalipour, M. Transcriptome Profiling of Cisplatin Resistance in Triple-negative Breast Cancer: New Insight into the Role of PI3k/Akt Pathway. *Curr. Mol. Med.* **23**, 559–568 (2022).
80. Zhu, X. F. *et al.* Knockdown of heme oxygenase-1 promotes apoptosis and autophagy and enhances the cytotoxicity of doxorubicin in breast cancer cells. *Oncol. Lett.* **10**, 2974–2980 (2015).
81. Zhu, X. *et al.* HMOX-1 inhibits TGF- β -induced epithelial-mesenchymal transition in the MCF-7 breast cancer cell line. *Int. J. Mol. Med.* **40**, 411–417 (2017).
82. Yuan, L. *et al.* Upregulation of the ferroptosis-related STEAP3 gene is a specific predictor of poor triple-negative breast cancer patient outcomes. *Front. Oncol.* **13**, 1032364 (2023).
83. Zhu, L. *et al.* Genomic Analysis Uncovers Immune Microenvironment Characteristics and Drug Sensitivity of Ferroptosis in Breast Cancer Brain Metastasis. *Front. Genet.* **12**, 1–18 (2022).
84. Chen, L. *et al.* EIF2A promotes cell survival during paclitaxel treatment in vitro and in vivo. *J. Cell. Mol. Med.* **23**, 6060–6071 (2019).
85. Liu, X. *et al.* Acceleration of pancreatic tumorigenesis under immunosuppressive microenvironment induced by reg3g overexpression. *Cell Death Dis.* **8**, e3033-12 (2017).
86. Song, J. *et al.* 14-3-3 ζ inhibits heme oxygenase-1 (HO-1) degradation and promotes hepatocellular carcinoma proliferation: Involvement of STAT3 signaling. *J. Exp. Clin. Cancer Res.* **38**, 3 (2019).
87. Pawlik, T. M. *et al.* Proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein. *BMC Cancer* **6**, 68 (2006).
88. CINE, N. *et al.* Identification of ApoA1, HPX and POTEE genes by omic analysis in breast cancer. *Oncol. Rep.* **32**, 1078–1086 (2014).

89. Park, M. H., Kwon, H. J., Kim, J., Lee, B. & Lee, S. J. Elevated Interleukin-13 Receptor Alpha 1 Expression in Tumor Cells Is Associated with Poor Prognosis in Patients with Invasive Breast Cancer. *Ann. Surg. Oncol.* **24**, 3780–3787 (2017).
90. He, M., Hu, C., Deng, J., Ji, H. & Tian, W. Identification of a novel glycolysis-related signature to predict the prognosis of patients with breast cancer. *World J. Surg. Oncol.* **19**, 1–14 (2021).
91. Zhao, N. *et al.* Group 2 innate lymphoid cells promote TNBC lung metastasis via the IL-13-MDSC axis in a murine tumor model. *Int. Immunopharmacol.* **99**, 107924 (2021).
92. Feng, X. *et al.* Identification of diagnostic biomarkers and therapeutic targets in peripheral immune landscape from coronary artery disease. *J. Transl. Med.* **20**, 1–17 (2022).
93. Robinson, M. B. *et al.* IL-6 trans-signaling increases expression of airways disease genes in airway smooth muscle. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **309**, L129–L138 (2015).
94. Venmar, K. T., Carter, K. J., Hwang, D. G., Dozier, E. A. & Fingleton, B. IL4 receptor ILR4 α regulates metastatic colonization by mammary tumors through multiple signaling pathways. *Cancer Res.* **74**, 4329–4340 (2014).
95. Venmar, K. T., Kimmel, D. W., Cliffler, D. E. & Fingleton, B. IL4 receptor α mediates enhanced glucose and glutamine metabolism to support breast cancer growth. *Biochim. Biophys. Acta - Mol. Cell Res.* **1853**, 1219–1228 (2015).
96. Ma, R. Y. *et al.* Monocyte-derived macrophages promote breast cancer bone metastasis outgrowth. *J. Exp. Med.* **217**, (2020).
97. Parveen, S. *et al.* Therapeutic targeting with DABIL-4 depletes myeloid suppressor cells in 4T1 triple-negative breast cancer model. *Mol. Oncol.* **15**, 1330–1344 (2021).
98. Bian, Z. *et al.* Arginase-1 is neither constitutively expressed in nor required for myeloid-derived suppressor cell-mediated inhibition of T-cell proliferation. *Eur. J. Immunol.* **48**, 1046–1058 (2018).
99. Deimel, L. P., Li, Z., Roy, S. & Ranasinghe, C. STAT3 determines IL-4 signalling outcomes in naïve T cells. *Sci. Rep.* **11**, 1–14 (2021).
100. Gao, J. *et al.* Decitabine assists umbilical cord-derived mesenchymal stem cells in improving glucose homeostasis by modulating macrophage polarization in type 2 diabetic mice. *Stem Cell Res. Ther.* **10**, 1–15 (2019).
101. Gante, I. *et al.* One Step Nucleic Acid Amplification (OSNA) Lysate Samples Are Suitable to Establish a Transcriptional Metastatic Signature in Patients with Early Stage Hormone Receptors-Positive Breast Cancer. *Cancers (Basel)*. **14**, (2022).
102. He, L. *et al.* LRG1 mediated by ATF3 promotes growth and angiogenesis of gastric cancer by regulating the SRC/STAT3/VEGFA pathway. *Gastric Cancer* **25**, 527–541 (2022).

103. Jemmerson, R., Staskus, K., Higgins, L., Conklin, K. & Kelekar, A. Intracellular leucine - rich alpha - 2 - glycoprotein - 1 competes with Apaf - 1 for binding cytochrome c in protecting MCF - 7 breast cancer cells from apoptosis. *Apoptosis* **26**, 71–82 (2021).
104. Ramirez-Ardila, D. E. *et al.* LRG1 mRNA expression in breast cancer associates with PIK3CA genotype and with aromatase inhibitor therapy outcome. *Mol. Oncol.* **10**, 1363–1373 (2016).
105. Zhang, Y. S., Han, L., Yang, C., Liu, Y. J. & Zhang, X. M. Prognostic Value of LRG1 in Breast Cancer: A Retrospective Study. *Oncol. Res. Treat.* **44**, 36–41 (2021).
106. Zhong, B. *et al.* Colorectal cancer-associated fibroblasts promote metastasis by up-regulating LRG1 through stromal IL-6/STAT3 signaling. *Cell Death Dis.* **13**, (2022).
107. Dritsoula, A. *et al.* Angiopathic activity of LRG1 is induced by the IL-6/STAT3 pathway. *Sci. Rep.* **12**, 1–14 (2022).
108. Yu, B. *et al.* LRG1 facilitates corneal fibrotic response by inducing neutrophil chemotaxis via Stat3 signaling in alkali-burned mouse corneas. *Am. J. Physiol. - Cell Physiol.* **321**, C415–C428 (2021).
109. Ni, H. *et al.* Molecular prognostic factors for distant metastases in premenopausal patients with hr+/her2– early breast cancer. *J. Pers. Med.* **11**, (2021).
110. Byun, K. *et al.* Clusterin/ApoJ enhances central leptin signaling through Lrp2-mediated endocytosis. *EMBO Rep.* **15**, 801–808 (2014).
111. Gil, S. Y. *et al.* Clusterin and LRP2 are critical components of the hypothalamic feeding regulatory pathway. *Nat. Commun.* **4**, 1–10 (2013).
112. Ohki, K. *et al.* Characteristics of genetic alterations of peripheral T-cell lymphoma in childhood including identification of novel fusion genes: the Japan Children's Cancer Group (JCCG). *Br. J. Haematol.* **194**, 718–729 (2021).
113. Vecchi, M. *et al.* Breast cancer metastases are molecularly distinct from their primary tumors. *Oncogene* **27**, 2148–2158 (2008).
114. Hu, L. *et al.* Expression, purification, and breast cancer cell inhibiting effect of recombinant human lactoferrin C-lobe. *Biosci. Biotechnol. Biochem.* **80**, 257–263 (2016).
115. Iglesias-Figueroa, B. F. *et al.* Recombinant human lactoferrin induces apoptosis, disruption of F-actin structure and cell cycle arrest with selective cytotoxicity on human triple negative breast cancer cells. *Apoptosis* **24**, 562–577 (2019).
116. Cizkova, M. *et al.* Gene Expression Profiling Reveals New Aspects of PIK3CA Mutation in ERalpha-Positive Breast Cancer: Major Implication of the Wnt Signaling Pathway. *PLoS One* **5**, e15647 (2010).
117. Zhao, Q., Cheng, Y. & Xiong, Y. LTF Regulates the Immune Microenvironment of Prostate Cancer Through JAK/STAT3 Pathway. *Front. Oncol.* **11**, 1–12

- (2021).
118. Sun, X., Bartos, A., Whitsett, J. A. & Dey, S. K. Uterine deletion of Gp130 or Stat3 shows implantation failure with increased estrogenic responses. *Mol. Endocrinol.* **27**, 1492–1501 (2013).
 119. Cutone, A. *et al.* Native and iron-saturated bovine lactoferrin differently hinder migration in a model of human glioblastoma by reverting epithelial-to-mesenchymal transition-like process and inhibiting interleukin-6/STAT3 axis. *Cell. Signal.* **65**, 109461 (2020).
 120. Chea, C. *et al.* Molecular mechanism of inhibitory effects of bovine lactoferrin on the growth of oral squamous cell carcinoma. *PLoS One* **13**, 1–19 (2018).
 121. García, E. P. *et al.* MUC5B silencing reduces chemo-resistance of MCF-7 breast tumor cells and impairs maturation of dendritic cells. *Int. J. Oncol.* **48**, 2113–2123 (2016).
 122. Valque, H., Gouyer, V., Gottrand, F. & Desseyn, J. L. MUC5B Leads to Aggressive Behavior of Breast Cancer MCF7 Cells. *PLoS One* **7**, (2012).
 123. Yuan, Z. *et al.* Central Role of the Threonine Residue within the p+1 Loop of Receptor Tyrosine Kinase in STAT3 Constitutive Phosphorylation in Metastatic Cancer Cells. *Mol. Cell. Biol.* **24**, 9390–9400 (2004).
 124. Hao, Y. *et al.* Mycoplasma pneumoniae modulates STAT3-STAT6/EGFR-FOXA2 signaling to induce overexpression of airway mucins. *Infect. Immun.* **82**, 5246–5255 (2014).
 125. Dave, A. *et al.* The Breast Cancer Single-Cell Atlas: Defining cellular heterogeneity within model cell lines and primary tumors to inform disease subtype, stemness, and treatment options. *Cell. Oncol.* (2023) doi:10.1007/s13402-022-00765-7.
 126. Jia, L. *et al.* A 10-Gene Signature for Predicting the Response to Neoadjuvant Trastuzumab Therapy in HER2-Positive Breast Cancer. *Clin. Breast Cancer* **21**, e654–e664 (2021).
 127. Conley, S. J. *et al.* HER2 drives Mucin-like 1 to control proliferation in breast cancer cells. *Oncogene* **35**, 4225–4234 (2016).
 128. Li, Q. H. *et al.* Small breast epithelial mucin promotes the invasion and metastasis of breast cancer cells via promoting epithelial-to-mesenchymal transition. *Oncol. Rep.* **44**, 509–518 (2020).
 129. O'Brien, N. *et al.* Use of a panel of novel genes for differentiating breast cancer from non-breast tissues. *Tumor Biol.* **28**, 312–317 (2007).
 130. Zhang, Y., Lun, X. & Guo, W. Expression of TRPC1 and SBEM protein in breast cancer tissue and its relationship with clinicopathological features and prognosis of patients. *Oncol. Lett.* **20**, 1–1 (2020).
 131. Liu, L. *et al.* Small breast epithelial mucin tumor tissue expression is associated with increased risk of recurrence and death in triple-negative breast cancer patients. *Diagn. Pathol.* **8**, 1–10 (2013).

132. Juchem, K. W. *et al.* NFAM1 Promotes Pro-Inflammatory Cytokine Production in Mouse and Human Monocytes. *Front. Immunol.* **12**, 1–19 (2022).
133. Lenz, G. *et al.* t-Darpp activates IGF-1R signaling to regulate glucose metabolism in trastuzumab-resistant breast cancer cells. *Clin. Cancer Res.* **24**, 1216–1226 (2018).
134. Murad, R. *et al.* Transcriptome and chromatin landscape changes associated with trastuzumab resistance in HER2 + breast cancer cells. *Gene* **799**, 145808 (2021).
135. Kotecha, S. *et al.* Dopamine and cAMP-regulated phosphoprotein 32 kDa (DARPP-32) and survival in breast cancer: a retrospective analysis of protein and mRNA expression. *Sci. Rep.* **9**, 1–11 (2019).
136. Zhu, S. *et al.* Silencing of miR490–3p by H. pylori activates DARPP-32 and induces resistance to gefitinib. *Cancer Lett.* **491**, 87–96 (2020).
137. Zhu, S. *et al.* Activation of IGF1R by DARPP-32 promotes STAT3 signaling in gastric cancer cells. *Oncogene* **38**, 5805–5816 (2019).
138. Chen, Z. *et al.* Gastric tumour-derived ANGPT2 regulation by DARPP-32 promotes angiogenesis. *Gut* **65**, 925–934 (2016).
139. Banh, R. S. *et al.* PTP1B controls non-mitochondrial oxygen consumption by regulating RNF213 to promote tumour survival during hypoxia. *Nat. Cell Biol.* **18**, 803–813 (2016).
140. Bergenfelz, C. *et al.* S100A9 expressed in ER-PgR-breast cancers induces inflammatory cytokines and is associated with an impaired overall survival. *Br. J. Cancer* **113**, 1234–1243 (2015).
141. Zhang, S. *et al.* Distinct prognostic values of S100 mRNA expression in breast cancer. *Nat. Publ. Gr.* 1–11 (2017) doi:10.1038/srep39786.
142. Allaoui, R. *et al.* Cancer-associated fibroblast-secreted CXCL16 attracts monocytes to promote stroma activation in triple-negative breast cancers. *Nat. Commun.* **7**, (2016).
143. Olsson, A. *et al.* Tasquinimod triggers an early change in the polarization of tumor associated macrophages in the tumor microenvironment. *J. Immunother. Cancer* **3**, 1–14 (2015).
144. Lee, M. J. *et al.* Interleukin-6 Induces S100A9 Expression in Colonic Epithelial Cells through STAT3 Activation in Experimental Ulcerative Colitis. *PLoS One* **7**, (2012).
145. Chang, Q. *et al.* The IL-6/JAK/Stat3 feed-forward loop drives tumorigenesis and metastasis. *Neoplasia (United States)* **15**, 848–862 (2013).
146. Takagi, R. *et al.* S100A9 Increases IL-6 and RANKL Expressions through MAPKs and STAT3 Signaling Pathways in Osteocyte-Like Cells. *Biomed Res. Int.* **2020**, (2020).
147. Xu, X. *et al.* S100A9 promotes human lung fibroblast cells activation through

- receptor for advanced glycation end-product-mediated extracellular-regulated kinase 1/2, mitogen-activated protein-kinase and nuclear factor- κ B-dependent pathways. *Clin. Exp. Immunol.* **173**, 523–535 (2013).
148. Cheng, P. *et al.* Inhibition of dendritic cell differentiation and accumulation of myeloid-derived suppressor cells in cancer is regulated by S100A9 protein. *J. Exp. Med.* **205**, 2235–2249 (2008).
 149. Alkhateeb, T. *et al.* S100A9 maintains myeloid-derived suppressor cells in chronic sepsis by inducing miR-21 and miR-181b. *Mol. Immunol.* **112**, 72–81 (2019).
 150. Huyton, T., Göttmann, W., Bade-Döding, C., Paine, A. & Blasczyk, R. The T/NK cell co-stimulatory molecule SECTM1 is an IFN ‘early response gene’ that is negatively regulated by LPS in Human monocytic cells. *Biochim. Biophys. Acta - Gen. Subj.* **1810**, 1294–1301 (2011).
 151. Moy, I. *et al.* Estrogen-dependent sushi domain containing 3 regulates cytoskeleton organization and migration in breast cancer cells. *Oncogene* **34**, 362–372 (2015).
 152. Hao, X. lin *et al.* Identification of TC2N as a novel promising suppressor of PI3K-AKT signaling in breast cancer. *Cell Death Dis.* **10**, (2019).
 153. Wells, W. A. *et al.* Expression of ‘Spot 14’ (THRSP) predicts disease free survival in invasive breast cancer: Immunohistochemical analysis of a new molecular marker. *Breast Cancer Res. Treat.* **98**, 231–240 (2006).
 154. Hilvo, M. *et al.* Novel theranostic opportunities offered by characterization of altered membrane lipid metabolism in breast cancer progression. *Cancer Res.* **71**, 3236–3245 (2011).
 155. Song, J. *et al.* Pan-Cancer Analysis Reveals the Signature of TMC Family of Genes as a Promising Biomarker for Prognosis and Immunotherapeutic Response. *Front. Immunol.* **12**, 1–15 (2021).
 156. Gan, X.-Y. *et al.* Role of Up-Regulated Transmembrane Channel-Like Protein 5 in Pancreatic Adenocarcinoma. *Dig. Dis. Sci.* **68**, 1894–1912 (2023).